Quantum work statistics and resource theories: bridging the gap through Rényi divergences

G. Guarnieri,[1, *] N. H. Y. Ng,[2] K. Modi,[3] J. Eisert,[2] M. Paternostro,[4] and J. Goold[5]

[1]*Department of Optics, Palacký University, 17. Listopadu 1192/12, 771 46 Olomouc, Czech Republic*
[2]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[3]*School of Physics and Astronomy, Monash University, Victoria 3800, Australia*
[4]*Centre for Theoretical Atomic, Molecular and Optical Physics,*
Queen's University Belfast, Belfast BT7 1NN, United Kingdom
[5]*School of Physics, Trinity College Dublin, Dublin 2, Ireland*
(Dated: May 15, 2018)

The work performed on or extracted from a non-autonomous quantum system described by means of a two-point projective-measurement approach takes the form of a stochastic variable. We show that the cumulant generating function of work can be recast in the form of quantum Rényi-α divergences, and by exploiting convexity of this cumulant generating function, derive a single-parameter family of bounds for the first moment of work. Higher order moments of work can also be obtained from this result. In this way, we establish a link between quantum work statistics in stochastic approaches on the one hand and resource theories for quantum thermodynamics on the other hand, a theory in which Rényi-α divergences take a central role. To explore this connection further, we consider an extended framework involving a control switch and an auxiliary battery, which is instrumental to reconstruct the work statistics of the system. We compare and discuss our bounds on the work distribution to findings on deterministic work studied in resource theoretic settings.

Fundamental out-of-equilibrium fluctuation theorems have been formulated to characterize the full non-linear response of both classical and quantum systems, to the action of a time-dependent external perturbation [1–5]. Such theorems can be seen as refined statements of the second law of thermodynamics suitable for application at the nano-scale. As such, they play an important role in the characterization of quantum and classical thermodynamic processes and thermal machines [6–9]. In such a framework, the statistics of stochastic thermodynamic variables can be gathered through two-time projective-measurement protocols, where the fluctuating work done by or on a system driven out of equilibrium or the heat that it exchanges with an environment are defined in terms of the difference of energy eigenvalues observed at the start and the end of the dynamics [10, 11]. This approach is experimentally viable [12, 13], has been useful for the characterization of non-equilibrium features of quenched many-body systems [14–18] and there is strong evidence that it has a physically meaningful semi-classical limit [19–23].

A complementing and in many ways radically different formal approach to the stochastic thermodynamics of quantum systems is based on resource theories. These are mathematical frameworks that specify restrictions to the transformations that can be performed on a quantum system to identify sets of *free states* that can be prepared under such constraints. Any state ρ that is not free can then be consumed to create final states which are also not free, making ρ a useful resource [24]. Initially applied to entanglement and coherence [25], resource-theoretical approaches have also been applied to study thermodynamics of quantum processes [26–31], providing insights on the inter-convertibility of finite resource states and on extractable work [32–34] from a general, axiomatic perspective that does not rely on specific details of

a particular system. Despite some success in capturing fundamental aspects of the interplay between non-equilibrium thermodynamics and quantum dynamics, it is not yet ascertained whether resource theories are mutually compatible with fluctuation theorem settings, and if their predictive powers are equivalent. Moreover, their predictive statements are rarely phrased in a directly comparable manner, the only exception insofar being in the thermodynamic limit. In such case the fluctuations become negligible, and the optimal amount of work extracted is found to be given as a function of the non-equilibrium free energy of the system [26].

In this work, we make a step forward towards bridging the gap between the resource theoretic approach to thermodynamics of quantum systems and the quantum stochastic one by presenting a situation in which these two frameworks may be directly compared. In order to provide the foundations of our approach, we start by considering closed unitary dynamics for the system and, after showing that the cumulant generating function can always be recast in terms of quantum Rényi-α divergences, we derive a family of single parameter bounds on the average work. Moreover, the relation between the cumulant generating function and the Rényi divergences allows to relate higher moments of the work distribution to higher derivatives of the latter. The approach take is further motivated by recent insights into modified versions of the Crooks relation in the context of single-shot work extraction [35] and one-shot dissipated work from Renyi divergences [36]. We compare our result to findings from a resource theory perspective, where extractable work can also be phrased in terms of Rényi divergences, and discuss the similarities and distinctions of these two results.

We then consider an open quantum system scenario, where the system of interest interacts with a thermal bath and is attached to both an auxiliary quantum battery and a control switch (the latter describing the action of an external driving work protocol) so as to meet the usual conditions invoked in resource theory-based approaches. As a general prescription,

arXiv:1804.09962v2 [quant-ph] 14 May 2018

we perform the two-point measurement protocol on the auxiliary battery rather than on the system, thus preserving any initial coherence in the latter. Work can then be consistently defined in terms of the energy difference in the battery. This allows us to show that, when the system is initially prepared in a thermal state (which is anyway implicitly assumed in every explicit quantum stochastic approach [19, 37]) the cumulant generating function can still be cast in terms of quantum Rényi-α divergences. Analogue considerations and results as done in the first part of the work for closed systems are thus recovered in this extended scenario. In particular, this provides strong (or significant) evidence of the equivalence. Moreover, our results allow to attach a clear physical interpretation to the α-Rényi divergences by linking them to statistical quantities that are experimentally accessible.

Stochastic approach. Consider an isolated quantum system – initially prepared in an equilibrium state at inverse temperature $\beta > 0$ – subjected to an external force that changes a *work* parameter λ_t in time according to a generic finite-time protocol. The latter includes, at the initial time $t = 0$ and final time $t = \tau$, projective measurements of the energy of the system, which results in the values $E_n^{\lambda_0}$ and $E_m^{\lambda_\tau}$. Here, n and m labels the respective energy levels of the initial and final Hamiltonian $\mathcal{H}(\lambda_0)$, $\mathcal{H}(\lambda_\tau)$ of the system. Thermal and quantum randomness render the measured energy difference $E_m^{\lambda_\tau} - E_n^{\lambda_0}$, which can be interpreted as the work done on the system through the protocol, a stochastic variable whose values are provided by the trajectory-ensemble distribution [10]

$$p_\tau(W) = \sum_{n,m} P_\tau \left[E_m^{\lambda_\tau}, E_n^{\lambda_0} \right] \delta \left[W - (E_m^{\lambda_\tau} - E_n^{\lambda_0}) \right]. \quad (1)$$

Here, $P_\tau[E_m^{\lambda_\tau}, E_n^{\lambda_0}]$ is the joint probability density that the two-time energy measurements results in the values $E_m^{\lambda_\tau}$ and $E_n^{\lambda_0}$. The Fourier transform of p_τ gives a generating function $\Theta(\eta, \tau) := \langle e^{i\eta W} \rangle_\tau$ which, by derivation over the counting field parameter η, gives the nth-order moment of work. Another informative quantity that we shall consider in this work is the cumulant generating function

$$\Phi(\eta, \tau) := \ln \langle e^{-\eta W} \rangle_\tau = \ln \int dW \, p_\tau(W) e^{-\eta W}. \quad (2)$$

The quantity $(-1)^n \partial_\eta^n \Phi(\eta, \tau)|_{\eta=0}$, gives us the cumulants of work. Using the Hölder inequality, it is possible to demonstrate the convexity of Φ with respect to the first argument [38]. This property can be equivalently stated as [39]

$$\Phi(\eta, \tau) \geq \eta \partial_\eta \Phi(\eta, \tau)|_{\eta=0}, \quad (3)$$

and, as $\langle W \rangle_\tau = -\partial_\eta \Phi(\eta, \tau)|_{\eta=0}$, we immediately obtain a single-parameter family of lower bounds for the mean work,

$$\beta \langle W \rangle_\tau \geq -\frac{\beta}{\eta} \Phi(\eta, \tau), \quad \eta > 0. \quad (4)$$

A similar set of one-parameter bounds was recently derived in the context of Landauer erasure [40]. For negative values $\eta < 0$, a family of upper bounds $\beta \langle W \rangle_\tau \leq \beta \Phi(\eta, \tau)/|\eta|$ is obtained instead.

Connecting the bounds to Rényi divergences. It is well known in the field of full counting statistics [3] that, for an initial Gibbs state of the bath, the cumulant generating function can be recast as

$$\Phi(\eta, \tau) = \ln \text{Tr} \left[\rho_S(\eta, \tau) \right] \quad (5)$$

with $\rho_S(\eta, \tau) = U_{\eta/2}(\tau) \rho_S(0) U_{-\eta/2}^\dagger(\tau)$, with the operator $U_\eta(\tau) := e^{-\eta H(\lambda_\tau)} U(\tau) e^{\eta H(\lambda_0)}$ and $U(\tau)$ being the time-evolution operators of the system at time τ. Starting from this expression, the following identity can be derived (see Appendix A for details)

Observation 1 (Cumulant generating function) *The cumulant generating function for the moments of work is given by*

$$\Phi(\eta, \tau) = -\frac{\eta}{\beta} S_{1-\frac{\eta}{\beta}} \left(\rho_S(\tau) \,||\, \mathcal{G}_S(\lambda_\tau) \right) - \eta \Delta F, \quad (6)$$

where $\mathcal{G}_S(\lambda_t) := Z(\lambda_t)^{-1} e^{-\beta H(\lambda_t)}$ *denotes the canonical Gibbs state at time t and $\Delta F = F(\lambda_\tau) - F(\lambda_0)$ is the free energy difference between canonical Gibbs states at the initial and final points, with the free energy of Gibbs states at time t being $F(\lambda_t) = -\beta^{-1} \ln Z(\lambda_t)$ and $Z(\lambda_t)$ the partition function of $\mathcal{H}(\lambda_t)$.*

In this expression, the quantum α-Rényi divergences are defined as

$$S_\alpha(\rho||\sigma) := \frac{1}{\alpha - 1} \ln \text{Tr} \left[\rho^\alpha \sigma^{1-\alpha} \right], \quad \alpha \in (0,1) \cup (1, +\infty), \quad (7)$$

with ρ and σ being two generic density matrices [41]. The Rényi divergence of order $\alpha = 1$ reduces to the familiar quantum relative entropy, i.e., $\lim_{\alpha \to 1} S_\alpha(\rho||\sigma) = D(\rho||\sigma) = \text{tr}(\rho \ln \rho - \rho \ln \sigma)$. Eq. (7) has recently gained much attention due to its role in resource-theoretical formulations of thermodynamics [28] and the central role that it plays in the quantification of the irreversible entropy production resulting from non-equilibrium processes [41, 42]. Combining Eq. (6) with Eq. (4), one obtains an inequality on the irreversible entropy $\langle S_{\text{irr}} \rangle := \beta(\langle W \rangle - \Delta F)$ [43, 44],

$$\langle S_{\text{irr}} \rangle \geq S_{1-\frac{\eta}{\beta}} \left(\rho_S(\tau) \,||\, \mathcal{G}_S(\lambda_\tau) \right). \quad (8)$$

It is important to stress such a relation stems just from the convexity property of the cumulant generating function of the work distribution and Eq. (6). However, noting Refs. [45–47] and accepting that the condition on Rényi divergences is the tightest for $\alpha = 1$, $\langle S_{\text{irr}} \rangle = \lim_{\eta \to 0^+} S_{1-\frac{\eta}{\beta}} \left(\rho_S(\tau) \,||\, \mathcal{G}_S(\lambda_\tau) \right)$, we arrive at the following statement.

Observation 2 (Stochastic irreversible entropy) *The irreversible entropy in the stochastic approach is*

$$\langle S_{irr} \rangle = D \left(\rho_S(\tau) \,||\, \mathcal{G}_S(\lambda_\tau) \right). \quad (9)$$

Furthermore, Eq. (6) can be used to relate higher moments of the work distribution and higher derivatives of the Rényi-α divergence. In particular, the second cumulant $\text{Var}(W)$ is

$$\text{Var}(W) = \frac{2}{\beta^2} \frac{\partial S_\alpha(\rho_S(\tau)||\mathcal{G}_S(\lambda_\tau))}{\partial \alpha} \bigg|_{\alpha=1} \quad (10)$$

$$= \frac{1}{\beta^2} V(\rho_S(\tau) \| \mathcal{G}_S(\lambda_\tau)), \tag{11}$$

where $V(\rho\|\sigma) := \mathrm{tr}(\rho(\log\rho - \log\sigma)^2) - (\mathrm{tr}(\rho(\log\rho - \log\sigma)))^2$ is the *relative entropy variance* [48, 49]. Finally, the fact that this cumulant generating function can be expressed in terms of Rényi divergences allows us also to bound the fluctuations of work. By using a simple Chernoff bound, we have in fact that, for any $k > 0$,

$$\Pr\left[W \geq \langle W \rangle + k\sigma_W\right] \leq \frac{1}{k^2}, \tag{12}$$

where $\sigma_W = \sqrt{\mathrm{Var}(W)}$ is the standard deviation for the distribution of W. In a single instance of this thermodynamic process, W may exhibit arbitrary fluctuations. However, if one considers multiple identical processes and determines the overall accumulated work, then we know via central limit theorem that the total work assumes a sharply peaked normal distribution. More concretely, consider the amount of work W_n used to perform the stochastic process described above, for n identical copies of the initial system. In this picture, one can define the ε-deterministic work W_n^ε to be the amount of work such that $W_n \leq W_n^\varepsilon$, except with some failure probability ε. The final state will be $\rho_S(\tau)^{\otimes n}$, and furthermore note that both $D(\rho_S(\tau)\|\mathcal{G}_S(\lambda_\tau))$ and $V(\rho_S(\tau)\|\mathcal{G}_S(\lambda_\tau))$ are additive under tensor product. Thus, for any $\varepsilon > 0$, by substituting $k^{-2} = \varepsilon$, we conclude from Eq. (12) the following.

Observation 3 (Stochastic ε-deterministic work) *The work obtained in a setting failing with probability $\varepsilon > 0$ in the multi-copy stochastic approach is given by*

$$W_n^\varepsilon = n \left[\langle W \rangle + \sqrt{\frac{\mathrm{Var}(W)}{\varepsilon n}} \right], \tag{13}$$

with $\langle W \rangle = \beta^{-1}[\langle S_{irr} \rangle + \Delta F]$ where $\langle S_{irr} \rangle$ was derived in Eq. (9), and $\mathrm{Var}(W)$ in Eq. (11).

General scenario and connection to resource theory. All the results insofar have been obtained considering a closed quantum system subject to an external driving work protocol [37]. The present aim is to demonstrate that a fundamental relation like Eqs. (6) and (8) can be retrieved in a different scenario closer to the typcial scenario in resource theories. For this reason, alongside the system of interest S, let us consider a control switch C, modeling the action of an external driving protocol, and a battery B operating as a storage system for work, such that the total Hamiltonian is given by

$$H_{S,B,C} = H_S(\lambda_0) \otimes \Pi_{\lambda_0,C} + H_S(\lambda_\tau) \otimes \Pi_{\lambda_\tau,C} + H_B, \tag{14}$$

where $\Pi_{\lambda,C} = |\lambda\rangle\langle\lambda|_C$. The free Hamiltonian $H_B = X$ of the battery is taken to be a Hamiltonian given by the position operator.

Next, we specify the class of interactions allowed to take place between the systems S, C and B. Here, we consider operations that satisfy the following constraints: **(1)** *Unitarity of the dynamics of the whole compound, governed by some*

$\mathcal{U}$. **(2)** *Energy-preserving nature*, i.e., $[\mathcal{U}, H_{S,B,C}] = 0$, with $H_{S,C,B}$ denoting the Hamiltonian of the overall system. **(3)** *Invariance under displacements of the battery*, i.e., $[\Delta_B, \mathcal{U}] = 0$, with Δ_B being the Weyl displacement operator shifting positions [50]. This set of operations closely resemble the set of thermal operations described in a resource theory setting [51], being a special case that in Eq. (14) there is no additional thermal bath, but instead the system itself is initialized in a Gibbs state. The inclusion of both the control switch and the battery are necessary in order to model an explicit time-dependent external work protocol into a time-independent, energy-preserving transformation [27, 32, 52]. While condition **(1)** is comes simply from quantum mechanics, request **(2)** is equivalent to asking that the Gibbs state of the global system is preserved. Finally, constraint **(3)** ensures that the battery acts only as a system that stores/provides work, instead of acting as an additional resource for coherence, or as a entropy sink. Work is defined as energy difference on the battery [53],

$$W := -\left(E_n^B - E_m^B\right), \tag{15}$$

where this is again a fluctuating work variable. The crucial difference brought along by Eq. (15) lies in the fact that the statistics of work is reconstructed by performing the two-projective-measurement scheme on the battery, rather than on the system.

As a final constraint, similar to that of Ref. [51], we require that the unitary $\mathcal{U}(\tau)$ acting on the systems S, B, C, perfectly produces the desired change on the system Hamiltonian from $H_S(\lambda_0)$ to $H_S(\lambda_\tau)$. This means that, if the initial state of the control switch is taken as $\rho_C(0) = \Pi_{\lambda_0,C}$, we want to have

$$\mathcal{U}(\tau)\left(\rho_{S,B}(0) \otimes \Pi_{\lambda_0,C}\right)\mathcal{U}^\dagger(\tau) = \rho_{S,B}(\tau) \otimes \Pi_{\lambda_\tau,C}. \tag{16}$$

To satisfy this, we require that $\mathcal{U}(\tau) = \mathcal{U}_{S,B,1}(\tau) \otimes |\lambda_\tau\rangle\langle\lambda_0|_C + \mathcal{U}_{S,B,2}(\tau) \otimes |\lambda_0\rangle\langle\lambda_\tau|_C$, where $\mathcal{U}_{S,B,(1,2)}(\tau)$ are generic unitary transformations on the joint system S, B. This ensures that when $\rho_C(0) = \Pi_{\lambda_0,C}$, then $\mathcal{U}(\tau)$ effectively induces a unitary transformation on system S, B,

$$\rho_{S,B}(\tau) = \mathcal{U}_{S,B,1}(\tau)\rho_{S,B}(0)\mathcal{U}_{S,B,1}^\dagger(\tau). \tag{17}$$

Condition **(2)** expressing the energy conservation of the global system S, B, C implies that

$$\mathcal{U}_{S,B,1}(\tau)\left(H_S(\lambda_0) + H_B\right)\mathcal{U}_{S,B,1}^\dagger(\tau) = H_S(\lambda_\tau) + H_B. \tag{18}$$

From this, we see that $\mathcal{U}_{S,B,1}(\tau)$ does not necessarily preserve the energy of S, B, and whatever energy difference incurred on S, B is stored in the state of the switch C.

We demonstrate that a relation akin to Eq. (6) can be derived also in this extended scenario. Let us consider the initial system to be prepared in $\rho_S(0) = \mathcal{G}_S(\lambda_0)$, i.e. an equilibrium Gibbs at inverse temperature β relative to the initial Hamiltonian $H_S(\lambda_0)$. Furthermore, let the initial state of the battery is a pure state $\rho_B(0)$, in a Gaussian state that well approximates a state with definite position. Keeping in mind that after the unitary transformation, the two-point measurement protocol will be performed on the battery. Therefore, the cumulant

generating function of work defined in Eq. (15) is given by

$$\Phi(\eta, \tau) = \ln \mathrm{Tr}\left[e^{\eta H_B}\mathcal{U}_{S,B,1}(\tau)e^{-\eta H_B}\rho_{S,B}(0)\mathcal{U}_{S,B,1}^\dagger(\tau)\right],$$

where $\rho_{S,B}(0) = \mathcal{G}_S(\lambda_0) \otimes |x\rangle\langle x|_B$, and $\mathcal{U}_{S,B,1}(\tau)$ satisfies Eq. (18). Using this property of $\mathcal{U}_{S,B,1}(\tau)$, we show that (see Appendix A 2 for details)

$$\Phi(\eta, \tau) = -\frac{\eta}{\beta}S_{1-\frac{\eta}{\beta}}\left(\tilde{\rho}_S(\tau) \,\|\, \mathcal{G}_S(\lambda_\tau)\right) - \eta\Delta F, \quad (19)$$

where $\tilde{\rho}_S(\tau)$ is defined as

$$\tilde{\rho}_S(\tau) = \left(\mathrm{Tr}_B\left[\mathcal{U}_{S,B}(\tau)\left(\rho_S(0)^\gamma \otimes \rho_B(0)\right)\mathcal{U}_{S,B}^\dagger(\tau)\right]\right)^{1/\gamma} \quad (20)$$

with $\gamma := 1 - \eta/\beta$. A comparison between Eq. (19) with Eq. (6) shows the mutual similarity, the only difference being in the first argument of the quantum Rényi-α divergence, namely in the $\tilde{\rho}_S(\tau)$ in place of $\rho_S(\tau)$. The former in fact now depends on the state of the newly introduced battery and on the operation $\mathcal{U}_{S,B}(\tau)$ performed on the S, B compound and keeps track of the fact that the energy statistics is measured and reconstructed on the battery rather than on the system. If one substitutes $\gamma = 1$, one recovers Eq. (9) exactly. However, the generic γ dependence implies that the second order correction terms might take on a more complicated form, when $\rho_{S,B}(\tau)$ contains correlations. It is interesting to note that a similar observation has been made in Ref. [51] (Section IV, Eq. (30)), where higher order moments of the work distribution could not be directly analyzed due to correlating terms between system and battery energy.

In order to further compare this result with that of the resource theory setting, let us assume that the final joint state $\rho_{S,B}(\tau)$ in Eq. (17) is a product state $\rho_S(\tau) \otimes \rho_B(\tau)$. This implies that $\tilde{\rho}_S(\tau) = \rho_S(\tau)$, and thus one recovers the identity in Eq. (6), the set of lower bounds on $\langle S_{\mathrm{irr}}\rangle$ in Eq. (8), and also the second moment of work distribution given by Eq. (10).

Comparison of the two approaches. In Ref. [28], a family of generalized second laws has been derived in the resource theory setting. These laws form a set of necessary and sufficient conditions for single-shot state transformations, on a single-copy of $\rho_S(0) \to \rho_S(\tau)$ via thermal operations. Furthermore, one may utilize these conditions to calculate the amount of deterministic work required for this process. This is modelled by requiring that the state transition $\rho_S(0) \otimes |E_0\rangle\langle E_0|_B \to \rho_S(\tau) \otimes |E_\tau\rangle\langle E_\tau|_B$ satisfies all the generalized second laws, and the amount of work invested is given by $W_{\mathrm{det}} = E_0 - E_\tau$. Applying the generalized second laws to this scenario tells us that the amount of deterministic work used in bringing the system from $\rho_S(0)$ to $\rho_S(\tau)$ is ([28], Appendix I)

$$W_{\mathrm{det}} \geq \mathcal{F}_\alpha\left(\rho_S(\tau), \mathcal{G}_S(\lambda_\tau)\right) - \mathcal{F}_\alpha\left(\rho_S(0), \mathcal{G}_S(\lambda_0)\right), \quad (21)$$

for all $\alpha \geq 0$, with

$$\mathcal{F}_\alpha\left(\rho(t), \mathcal{G}(\lambda_t)\right) := -\beta^{-1}\left[\ln Z(t) - S_\alpha\left(\rho(t) \,\|\, \mathcal{G}_S(t)\right)\right]. \quad (22)$$

If the initial state is assumed to thermal i.e., $\rho_S(0) = \mathcal{G}_S(\lambda_0)$, then $\mathcal{F}_\alpha\left(\rho_S(0), \mathcal{G}_S(\lambda_0)\right) = -\beta^{-1}\ln Z_S(\lambda_0)$. Therefore, if one defines the quantity $S_{\mathrm{irr}}^{\mathrm{det}} := \beta(W_{\mathrm{det}} - \Delta F)$, then Eq. (21) reads as follows.

Observation 4 (Resource-theoretic irreversible entropy)
The irreversible entropy in a resource-theoretic approach is lower bounded by

$$S_{irr}^{\mathrm{det}} \geq S_\alpha\left(\rho_S(\tau) \,\|\, \mathcal{G}_S(\lambda_\tau)\right), \qquad \alpha \geq 0. \quad (23)$$

Comparing Eq. (23) and (8), we observe a direct connection for $\eta \in (0, \beta)$, corresponding to the range $\alpha \in (0, 1)$. In this parameter regime, we see that $\langle S_{\mathrm{irr}}\rangle$ and $S_{\mathrm{irr}}^{\mathrm{det}}$ are bounded identically. Therefore, the stochastic approach sheds some light on the significance of these α-free energies, due to their relation with the physically accessible quantity $\Phi(\eta, \tau)$ [12, 13]. In contrast, the qualitative difference between S_{irr} and $S_{\mathrm{irr}}^{\mathrm{det}}$ is captured by the regime of $\eta < 0$ (corresponding to $\alpha > 1$). In this regime, we have that $S\langle S_{\mathrm{irr}}\rangle \leq S_\alpha\left(\rho_S(\tau) \,\|\, \mathcal{G}_S(\lambda_\tau)\right)$ [40], while for $S_{\mathrm{irr}}^{\mathrm{det}}$, Eq. (23) still holds. This difference is largely due to the fact that Eq. (8) deals with mean work, thus considered as a fluctuating quantity, while Eq. (23) bounds the deterministic work W_{det}. The best estimate for $S_{\mathrm{irr}}^{\mathrm{det}}$ is given by the $\infty-$Rényi divergence instead, and in general $S_\infty \geq S_1$.

A second reconciliation. A second reconciliation point between the stochastic approach and the resource theory approach can be reached when one compares the quantity W_n^ε in Eq. (13) to the ε-deterministic work of formation $W_{\mathrm{F},n}^\varepsilon$ derived in Ref. [54]. The analysis of $W_{\mathrm{F},n}^\varepsilon$ adapts also a resource theoretic approach, namely it considers work for a single-shot process. However, this process may be a global operation occuring on n copies of identical systems, for finite but large n. In particular, one considers the amount of work $W_{\mathrm{F},n}^\varepsilon$ required in order to prepare n identical copies of some final target state, $\rho_S^{\otimes n}$, with fidelity at least $1 - \varepsilon$. The quantity of interest $W_{\mathrm{F},n}^\varepsilon$ is defined as follows: For some fixed energy value E, and some parameter n, consider the minimum integer m such that the following transition on $m + n$ systems with identical, time-independent Hamiltonians,

$$|E\rangle\langle E|^{\otimes m} \otimes \mathcal{G}^{\otimes n} \to \sigma, \quad \hat{F}(\sigma, \mathcal{G}^{\otimes m} \otimes \rho^{\otimes n}) \geq 1 - \varepsilon, \quad (24)$$

is possible via thermal operations with a bath at inverse temperature β, $\mathcal{G}$ being the Gibbs state, and $\hat{F}(\rho, \sigma)$ denoting Uhlmann's fidelity [55]. The amount of work is given by $W_{\mathrm{F},n}^\varepsilon = mE$. In Ref. [54], it is shown that

$$W_{\mathrm{F},n}^\varepsilon \approx \beta^{-1}\left[nD(\rho\|\mathcal{G}) + \sqrt{nV(\rho\|\mathcal{G})}f(\varepsilon)\right], \quad (25)$$

where $f(\varepsilon) > 0$. Comparing the expressions for W_n^ε captured in Observation 3 (taking into account that when the initial and final Hamiltonian coincide, $\Delta F = 0$) and $W_{\mathrm{F},n}^\varepsilon$, which were defined using very different approaches, we see that nevertheless they are in qualitative agreement with one another.

Outlook. In this work, we have brought two approaches to quantum thermodynamics significantly closer to each other. While the approaches taken are radically different in mindset, they give rise to expressions formally providing similar or identical predictions, specifically when this line of thought is applied to notions of work extraction in quantum thermodynamics. It is the hope that this reconciling work can significantly contribute to the emerging theory of quantum thermodynamics.

[1] C. Jarzynski, Annu. Rev. Condens. Matter Phys. **2**, 329 (2011).

[2] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012).

[3] M. Esposito, U. Harbola, and S. Mukamel, Rev. Mod. Phys. **81**, 1665 (2009).

[4] M. Campisi, P. Hänggi, and P. Talkner, Rev. Mod. Phys. **83**, 771 (2011).

[5] P. Hänggi and P. Talkner, Nature Phys. **11**, 108 (2015).

[6] J. Millen and A. Xuereb, New J. Phys. **18**, 011002 (2016).

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

[8] S. Vinjanampathy and J. Anders, Contemp. Phys. **57**, 545 (2016).

[9] F. Binder, S. Vinjanampathy, K. Modi, and J. Goold, Phys. Rev. E **91**, 032119 (2015).

[10] P. Talkner, E. Lutz, and P. Hänggi, Phys. Rev. E **75**, 050102 (2007).

[11] P. Talkner, M. Campisi, and P. Hänggi, J. Stat. Mech. **P02025** (2009), 10.1088/1742-5468/2009/02/P02025.

[12] T. B. Batalhão, A. M. Souza, L. Mazzola, R. Auccaise, R. S. Sarthour, I. S. Oliveira, J. Goold, G. De Chiara, M. Paternostro, and R. M. Serra, Phys. Rev. Lett. **113**, 140601 (2014).

[13] S. An, J.-N. Zhang, M. Um, D. Lv, Y. Lu, J. Zhang, Z.-Q. Yin, H. Quan, and K. Kim, Nature Phys. **11**, 193 (2015).

[14] A. Silva, Phys. Rev. Lett. **101**, 120603 (2008).

[15] A. Gambassi and A. Silva, Phys. Rev. Lett. **109**, 250602 (2012).

[16] E. Mascarenhas, H. Bragança, R. Dorner, M. F. Santos, V. Vedral, K. Modi, and J. Goold, Phys. Rev. E **89**, 062103 (2014).

[17] A. Sindona, J. Goold, N. L. Gullo, and F. Plastina, New J. Phys. **16**, 045013 (2014).

[18] L. Fusco, S. Pigeon, T. J. G. Apollaro, A. Xuereb, L. Mazzola, M. Campisi, A. Ferraro, M. Paternostro, and G. De Chiara, Phys. Rev. X **4**, 031029 (2014).

[19] C. Jarzynski, H. T. Quan, and S. Rahav, Phys. Rev. X **5**, 031038 (2015).

[20] L. Zhu, Z. Gong, B. Wu, H. Quan, *et al.*, Phys. Rev. E **93**, 062108 (2016).

[21] Q. Wang and H. Quan, Phys. Rev. E **95**, 032113 (2017).

[22] I. García-Mata, A. J. Roncaglia, and D. A. Wisniacki, Phys. Rev. E **95**, 050102 (2017).

[23] I. García-Mata, A. J. Roncaglia, and D. A. Wisniacki, arXiv:1709.05115 (2017).

[24] F. G. Brandão and G. Gour, Phys. Rev. Lett. **115**, 070503 (2015).

[25] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).

[26] M. Horodecki and J. Oppenheim, Nat. Comm. **4**, 2059 (2013).

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

[28] F. Brandao, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, Proc. Natl. Ac. Sci. **112**, 3275 (2015).

[29] P. Faist, F. Dupuis, J. Oppenheim, and R. Renner, Nature Comm. **6**, 7669 (2015).

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

[31] N. H. Y. Ng, L. Mancinska, C. Cirstoiu, J. Eisert, and S. Wehner, New J. Phys. **17**, 085004 (2015).

[32] J. Aberg, Nature Comm. **4**, 1925 (2013).

[33] D. Egloff, O. C. O. Dahlsten, R. Renner, and V. Vedral, New J. Phys. **17**, 073001 (2015).

[34] R. Gallego, J. Eisert, and H. Wilming, New J. Phys. **18**, 103017 (2016).

[35] N. Y. Halpern, A. J. P. Garner, O. O. Dahlsten, and V. Vedral, New J. Phys. **17**, 095003 (2015).

[36] N. Y. Halpern, A. J. P. Garner, O. O. Dahlsten, and V. Vedral, ArXiv:1505.06217.

[37] M. Campisi, P. Hänggi, and P. Talkner, Rev. Mod. Phys. **83**, 771 (2011).

[38] H. Touchette, Phys. Rep. **478**, 1 (2009).

[39] R. T. Rockafellar, *Convex analysis* (Princeton University Press, 2015).

[40] G. Guarnieri, S. Campbell, J. Goold, S. Pigeon, B. Vacchini, and M. Paternostro, New J. Phys. **19**, 103038 (2017).

[41] B.-B. Wei and M. Plenio, New J. Phys. **19**, 023002 (2017).

[42] J. P. Santos, G. Landi, and M. Paternostro, Phys. Rev. Lett. **118**, 220601 (2017).

[43] M. J. Donald, J. Stat. Phys. **49**, 81 (1987).

[44] S. Deffner and E. Lutz, Phys. Rev. Lett. **105**, 170402 (2010).

[45] H.-P. Breuer, Phys. Rev. A **68**, 032105 (2003).

[46] J. M. R. Parrondo, C. V. den Broeck, and R. Kawai, New J. Phys. **11**, 073008 (2009).

[47] S. Deffner and E. Lutz, Phys. Rev. Lett. **107**, 140404 (2011).

[48] M. Tomamichel and M. Hayashi, IEEE Trans. Inf. Th. **59**, 7693 (2013).

[49] S. M. Lin and M. Tomamichel, Quant. Inf. Proc. **14**, 1501 (2015).

[50] L. Cohen, *The Weyl operator and its generalization* (Springer Science & Business Media, 2012).

[51] Á. M. Alhambra, L. Masanes, J. Oppenheim, and C. Perry, Phys. Rev. X **6**, 041017 (2016).

[52] J. Aberg, (2016), arXiv:1601.01302.

[53] P. Skrzypczyk, A. J. Short, and S. Popescu, Nat. Commun. **5**, 1 (2014).

[54] C. T. Chubb, M. Tomamichel, and K. Korzekwa, ArXiv e-prints (2017), arXiv:1711.01193.

[55] A. Uhlmann, Rep. Math. Phys. **9**, 273 (1976).

Appendix A: Details of calculations

1. Proof of Eq. (6)

In the following we will explicitly derive the identity in Eq. (6), which connects the cumulant generating function of work statistics to an opportune $\alpha-$Rényi divergence. Here, we make use of the fact that the initial state is given by

$$\rho_S(0) = \mathcal{G}_S(\lambda_0) = \frac{1}{Z(\lambda_0)} e^{-\beta H(\lambda_0)}. \tag{A1}$$

For all values of $\eta \in (-\infty, \infty)\backslash\{0, \beta\}$, the cumulant generating function then reads as

$$\begin{aligned}
\Phi(\eta, \tau) &= \ln \mathrm{Tr}_S \left[e^{-(\eta/2)H(\lambda_\tau)} U(\tau) e^{(\eta/2)H(\lambda_0)} \rho_S(0) e^{(\eta/2)H(\lambda_0)} U^\dagger(\tau) e^{-(\eta/2)H(\lambda_\tau)} \right] \\
&= \ln \mathrm{Tr}_S \left[e^{-\eta H(\lambda_\tau)} U(t) e^{\eta H(\lambda_0)} \frac{e^{-\beta H(\lambda_0)}}{Z(\lambda_0)} U^\dagger(\tau) \right] \\
&= \ln \mathrm{Tr}_S \left[e^{-\eta H(\lambda_\tau)} U(\tau) \frac{e^{-(\beta-\eta)H(\lambda_0)}}{Z(\lambda_0)} U^\dagger(\tau) \right] \\
&= \ln \mathrm{Tr}_S \left[\left(\frac{e^{-\beta H(\lambda_\tau)}}{Z(\lambda_\tau)} \right)^{\eta/\beta} U(\tau) \left(\frac{e^{-\beta H(\lambda_0)}}{Z(\lambda_0)} \right)^{1-\eta/\beta} U^\dagger(\tau) \right] + \ln \left[\frac{Z(\lambda_\tau)}{Z(\lambda_0)} \right]^{\eta/\beta} \\
&= \ln \mathrm{Tr}_S \left[(\mathcal{G}_S(\lambda_\tau))^{\eta/\beta} (\rho_S(\tau))^{1-\eta/\beta} \right] + \ln \left[\frac{Z(\lambda_\tau)}{Z(\lambda_0)} \right]^{\eta/\beta} \\
&= \left(\frac{\eta}{\beta} - 1 \right) S_{\frac{\eta}{\beta}} (\mathcal{G}_S(\lambda_\tau) \| \rho_S(\tau)) + \frac{\eta}{\beta} \ln \frac{Z(\lambda_\tau)}{Z(\lambda_0)} \\
&= -\frac{\eta}{\beta} S_{1-\frac{\eta}{\beta}} (\rho_S(\tau) \| \mathcal{G}_S(\lambda_\tau)) - \eta \Delta F. \tag{A2}
\end{aligned}$$

In the last line, a skew-symmetry property of the $\alpha-$ Rényi divergence has been used, namely that for $\forall \alpha \neq 0, 1$, we have

$$S_\alpha(\rho\|\sigma) = \frac{\alpha}{1-\alpha} S_{1-\alpha}(\sigma\|\rho). \tag{A3}$$

Alternatively, one can also achieve this by applying the cyclic property of the trace operation in the third last line, i.e. by using the fact that $\mathrm{tr}_S(AB) = \mathrm{tr}_S(BA)$.

2. Proof of Eq. (19)

We know that if the global unitary $\mathcal{U}$ on joint systems S, B, C the requirement $\mathcal{U}(\tau) = \mathcal{U}_{S,B,1}(\tau) \otimes |\lambda_\tau\rangle\langle\lambda_0|_C + \mathcal{U}_{S,B,2}(\tau) \otimes |\lambda_0\rangle\langle\lambda_\tau|_C$, then it satisfies Eq. (16) and also gives an effective unitary $U_{S,B}(\tau) = U_{S,B,1}(\tau)$ on the reduced state of the system and battery. This unitary, in particular, changes the Hamiltonian on the system from $\mathcal{H}_S(\lambda_0)$ to $\mathcal{H}_S(\lambda_\tau)$, as described in Eq. (18). Moreover, note that since $\mathcal{H}_S(\lambda)$ always commutes with $\mathcal{H}_B$, we see how the two point measurement scheme acting on the battery can be directly related to the measurement statistics done on the system. In particular,

$$\begin{aligned}
&\mathcal{U}_{S,B}(\tau) [H_S(\lambda_0) + H_B] \mathcal{U}_{S,B}^\dagger(\tau) = H_S(\lambda_\tau) + H_B, \\
\Rightarrow \quad &\mathcal{U}_{S,B}(\tau) e^{-\eta[H_B + H_S(\lambda_0)]} \mathcal{U}_{S,B}^\dagger(\tau) = e^{-\eta[H_S(\lambda_\tau) + H_B]}, \\
\Rightarrow \quad &\mathcal{U}_{S,B}(\tau) e^{-\eta H_B} \mathcal{U}_{S,B}^\dagger(\tau) \mathcal{U}_{S,B}(\tau) e^{-\eta H_S(\lambda_0)} \mathcal{U}_{S,B}^\dagger(t) = e^{-\eta H_S(\lambda_\tau)} e^{-\eta H_B}, \\
\Rightarrow \quad &\mathcal{U}_{S,B}(\tau) e^{-\eta H_B} = e^{-\eta H_S(\lambda_\tau)} e^{-\eta H_B} \mathcal{U}_{S,B}(\tau) e^{\eta H_S(\lambda_0)}. \tag{A4}
\end{aligned}$$

Let us consider the two-time measurement protocol on the battery B as explained in the main text. The cumulant generating function is

$$\begin{aligned}
\Phi(\eta, \tau) &= \ln \mathrm{Tr}_{S,B,C} \left[e^{\eta H_B} \mathcal{U}(\tau) e^{-\eta H_B} \rho_{S,B,C}(0) \mathcal{U}^\dagger(\tau) \right] \\
&= \ln \mathrm{Tr}_{S,B} \left[e^{\eta H_B} \mathcal{U}_{S,B}(\tau) e^{-\eta H_B} \rho_{S,B}(0) \mathcal{U}_{S,B}^\dagger(\tau) \right], \tag{A5}
\end{aligned}$$

where we have substituted $\rho_C(0) = \Pi_{\lambda_0,C}$ and used Eq. (16), and traced out system C. We first note that, in order to evaluate $\Phi(\eta, \tau)$, the relation $\rho_B(0)^\alpha \propto \rho_B(0)$ is needed in our calculations, in order to simplify the cumulant generating function down to terms that only involve the system S. One can understand this intuitively: if the initial battery starts out with some non-trivial energy distribution, and if one defines the work statistics via TPM on the battery, this will depend not only on work fluctuations from the system but also on the prior distribution on the battery. Furthermore, $\rho_B(0)^\alpha \propto \rho_B(0)$ is satisfied if and only if

$$\rho_B(0) = \frac{1}{d_R}\Pi_R \tag{A6}$$

is a maximally mixed state on a subspace R of dimension d_R. In particular, if ρ_B is of such a form, then

$$\rho_B(0)^\alpha = d_R^{1-\alpha}\rho_B(0). \tag{A7}$$

Later on, we shall see that in order to maximize the cumulant generating function, $d_R = 1$, which means that the battery initial state $\rho_B(0)$ is a pure Gaussian state approximating a state with definite position. Now, by evaluating Eq. (A5) by making use of Eq. (A4), we have

$$\begin{aligned}
\Phi(\eta, \tau) &= \ln \mathrm{Tr}_{S,B}\left[e^{\eta H_B}\mathcal{U}_{S,B}(\tau)e^{-\eta H_B}\,\rho_{S,B}(0)\mathcal{U}_{S,B}^\dagger(\tau)\right] \\
&= \ln \mathrm{Tr}_{S,B}\left[e^{\eta H_B}\left(e^{-\eta H_S(\lambda_\tau)}e^{-\eta H_B}\mathcal{U}_{S,B}(\tau)e^{\eta H_S(\lambda_0)}\right)\rho_{S,B}(0)\mathcal{U}_{S,B}^\dagger(\tau)\right] \\
&= \ln \mathrm{Tr}_{S,B}\left[e^{-\eta H_S(\lambda_\tau)}\mathcal{U}_{S,B}(t)e^{\eta H_S(\lambda_0)}\left[\mathcal{G}_S(\lambda_0) \otimes \rho_B(0)\right]\mathcal{U}_{S,B}^\dagger(t)\right] \\
&= \ln \mathrm{Tr}_{S,B}\left[\left(\frac{e^{-\beta H_S(\lambda_\tau)}}{Z(\lambda_\tau)}\right)^{\eta/\beta} \otimes \mathbb{1}_B\,\mathcal{U}_{S,B}(t)\left(\frac{e^{-\beta H_S(\lambda_0)}}{Z(\lambda_0)}\right)^{1-\eta/\beta} \otimes \rho_B(0)\mathcal{U}_{S,B}^\dagger(t)\right] + \frac{\eta}{\beta}\ln\frac{Z(\lambda_\tau)}{Z(\lambda_0)} \\
&= \ln \mathrm{Tr}_{S,B}\left[\left(\mathcal{G}_S(\lambda_\tau) \otimes \mathbb{1}_B\right)^{\eta/\beta}\mathcal{U}_{S,B}(t)\left(\mathcal{G}_S(\lambda_0) \otimes \rho_B(0)\right)^{1-\eta/\beta}\mathcal{U}_{S,B}^\dagger(t)\right] + \frac{\eta}{\beta}\ln\frac{Z(\lambda_\tau)}{Z(\lambda_0)} - \frac{\eta}{\beta}\ln d_R. \tag{A8}
\end{aligned}$$

Besides other manipulations which are straightforward, the last line is obtained by assuming Eq. (A7). Observing Eq. (A8), since the last term is always non-positive for $\eta > 0$, in order to maximize it we take $d_R = 1$, and this term vanishes. We also observe that within the first term, we have an expression $\mathcal{U}_{S,B}(t)\rho_{S,B}(0)^{1-\eta/\beta}\mathcal{U}_{S,B}(t)^\dagger$. To simplify this, note that the final state on S, B is given by

$$\rho_{S,B}(t) = \mathcal{U}_{S,B}(t)\rho_{S,B}(0)\mathcal{U}_{S,B}(t)^\dagger = \sum_i p_i|f_i\rangle\langle f_i|_{S,B}, \tag{A9}$$

where $\rho_{S,B}(0)$ has the diagonal form $\rho_{S,B}(0) = \sum_i p_i|e_i\rangle\langle e_i|$, and $\mathcal{U}_{S,B}$ performs a unitary transformation from the ordered basis $\{|e_i\rangle_{S,B}\}$ to $\{|f_i\rangle_{S,B}\}$. In other words, for all i, $\mathcal{U}_{S,B}|e_i\rangle = |f_i\rangle$. This, together with the fact that for any $\alpha \in \mathbb{R}$, $\rho_{S,B}(0)^\alpha = \sum_i p_i^\alpha|e_i\rangle\langle e_i|$ gives us that

$$\mathcal{U}_{S,B}(\tau)\rho_{S,B}(0)^{1-\eta/\beta}\mathcal{U}_{S,B}(\tau)^\dagger = \rho_{S,B}(\tau)^{1-\eta/\beta}, \tag{A10}$$

which we may now substitute back into Eq. (A8). Finally, we have

$$\begin{aligned}
\Phi(\eta, \tau) &= \ln \mathrm{Tr}_{S,B}\left[\left(\mathcal{G}_S(\lambda_\tau) \otimes \mathbb{1}_B\right)^{\eta/\beta}\left(\rho_{S,B}(\tau)\right)^{1-\eta/\beta}\right] + \frac{\eta}{\beta}\ln\frac{Z(\lambda_\tau)}{Z(\lambda_0)} \\
&= \ln \mathrm{Tr}_S\left[\left(\mathcal{G}_S(\lambda_\tau)\right)^{\eta/\beta}\,\mathrm{Tr}_B\left[\left(\rho_{S,B}(\tau)\right)^{1-\eta/\beta}\right]\right] + \frac{\eta}{\beta}\ln\frac{Z(\lambda_\tau)}{Z(\lambda_0)} \\
&= \left(\frac{\eta}{\beta} - 1\right)S_{\frac{\eta}{\beta}}\left(\mathcal{G}_S(\lambda_\tau)\,||\,\tilde{\rho}_S(\tau)\right) + \frac{\eta}{\beta}\ln\frac{Z(\lambda_\tau)}{Z(\lambda_0)} \\
&= -\frac{\eta}{\beta}S_{1-\frac{\eta}{\beta}}\left(\tilde{\rho}_S(\tau)\,||\,\mathcal{G}_S(\lambda_\tau)\right) + \frac{\eta}{\beta}\ln\frac{Z(\lambda_\tau)}{Z(\lambda_0)}, \tag{A11}
\end{aligned}$$

where the idempotency of $\rho_B(0)$ and of $\mathbb{1}_B$ have been used (an operator A is idempotent iff for any $\gamma \neq 0$, we have $A^\gamma = A$). Furthermore, we have defined the quantity

$$\tilde{\rho}_S(\tau) = \left(\mathrm{Tr}_B\left[\mathcal{U}_{S,B}(\tau)\left(\rho_S(0)^\gamma \otimes \rho_B(0)\right)\mathcal{U}_{S,B}^\dagger(\tau)\right]\right)^{1/\gamma}, \qquad \gamma := 1 - \eta/\beta. \tag{A12}$$

Similarly, Eq. (A3) has been used to obtain the last line in Eq. (A11).

Appendix B: Details of the consistency between the family of lower bounds Eq. (8) and [28]

According to Ref. [28], a transition from a generic state ρ to another state ρ' is possible if and only if the family of generalized second laws applies, namely *iff*

$$\mathcal{F}_\alpha\left(\rho, \mathcal{G}\right) \geq \mathcal{F}_\alpha\left(\rho', \mathcal{G}'\right), \qquad \forall \alpha \geq 0, \tag{B1}$$

where $\mathcal{F}_\alpha$ is defined in Eq. (22) and $\mathcal{G}$ denotes the Gibbs state of the system. According to Ref. [28], we assume to start an initially uncorrelated state

$$\rho_{S,B}(0) = \rho_S(0) \otimes \rho_B(0), \tag{B2}$$

where, in line with our choice throughout the main text, we take the initial state of the battery to be $\rho_B(0) = |E_0\rangle\langle E_0|_B$, corresponding to a pure energy eigenstate with energy $E_B(0)$. Moreover, in the weak coupling regime assumed in Ref. [28] (see their Appendix I for details) implies

$$\rho_{S,B}(\tau) := \rho' = \rho_S(\tau) \otimes \rho_B(\tau). \tag{B3}$$

In what follows, $\mathcal{G}_{S(B)}$ denotes the Gibbs state of the system (battery). Furthermore, for deterministic work, the final state of $\rho_B(\tau) = |E_B(\tau)\rangle\langle E_B(\tau)|_B$ is a pure energy eigenstate as well. Using the definition of $\mathcal{F}_\alpha$ as given by Eq. (22), and noting that $\mathcal{F}_\alpha$ is additive under tensor product, we have therefore that Eq. (B1) reduces to

$$\mathcal{F}_\alpha\left(\rho_S(0) \otimes \rho_B(0), \mathcal{G}_S \otimes \mathcal{G}_B\right) \geq \mathcal{F}_\alpha\left(\rho_S(\tau) \otimes \rho_B(\tau), \mathcal{G}'_S \otimes \mathcal{G}'_B\right)$$
$$\mathcal{F}_\alpha\left(\rho_B(0), \mathcal{G}_B\right) - \mathcal{F}_\alpha\left(\rho_B(\tau), \mathcal{G}'_B\right) \geq \mathcal{F}_\alpha\left(\rho_S(\tau), \mathcal{G}'_S\right) - \mathcal{F}_\alpha\left(\rho_S(0), \mathcal{G}_S\right)$$
$$\beta^{-1}\left[S_\alpha(\rho_B(0)||\mathcal{G}_B) - \ln Z_B(0) - S_\alpha(\rho_B(\tau)||\mathcal{G}'_B) + \ln Z_B(\tau)\right] \geq \mathcal{F}_\alpha\left(\rho_S(\tau), \mathcal{G}'_S\right) - \mathcal{F}_\alpha\left(\rho_S(0), \mathcal{G}_S\right).$$

To further simplify these expressions, let us note that $S_\alpha(\rho_B(0)||\mathcal{G}_B)$ can be further evaluated. Since $\rho_B(0)$ is diagonal in the energy eigenbasis by construction, $\rho_B(0)$ and $\mathcal{G}_B$ also commute. Their $\alpha-$Rényi divergence is simply given by

$$S_\alpha(\rho_B(0)||\mathcal{G}_B) = \frac{1}{\alpha-1}\ln\mathrm{tr}\left[(|E_B(0)\rangle\langle E_B(0)|)^\alpha \left(\frac{e^{-\beta\mathcal{H}_B}}{Z_B(0)}\right)^{1-\alpha}\right] = \frac{1}{\alpha-1}\ln\left(\frac{e^{-\beta E_B(0)}}{Z_B(0)}\right)^{1-\alpha} = \beta E_B(0) + \ln Z_B(0).$$

This expression holds for a pure battery state, and since one is concerned with deterministic work here, a similar expression holds for $S_\alpha(\rho_B(\tau)||\mathcal{G}_B)$ as well. Substituting both expressions into the previous result, we end up with

$$E_B(0) - E_B(\tau) \geq \mathcal{F}_\alpha\left(\rho_S(\tau), \mathcal{G}'_S\right) - \mathcal{F}_\alpha\left(\rho_S(0), \mathcal{G}_S\right),$$

from which, by simply using the definition of work given by the energy difference in the battery $W = E_B(0) - E_B(\tau)$, we obtain Eq. (21).