Self-consistency of the two-point energy measurement protocol

M. Hamed Mohammady[1, *]

[1]*RCQI, Institute of Physics, Slovak Academy of Sciences, Dúbravská cesta 9, Bratislava 84511, Slovakia*

A thermally isolated quantum system undergoes unitary evolution by interacting with an external work source. The two-point energy measurement (TPM) protocol defines the work exchanged between the system and the work source by performing ideal energy measurements on the system before, and after, the unitary evolution. However, the ideal energy measurements used in the TPM protocol ultimately result from a unitary coupling with a measurement apparatus, which requires an interaction with an external work source. For the TPM protocol to be self-consistent, we must be able to perform the TPM protocol on the compound of system plus apparatus, thus revealing the total work distribution, such that when ignoring the apparatus degrees of freedom, we recover the original TPM work distribution for the system of interest. In the present manuscript, we show that such self-consistency is satisfied so long as the apparatus is initially prepared in an energy eigenstate. Moreover, we demonstrate that if the apparatus Hamiltonian is equivalent to the "pointer observable", then: (i) the total work distribution will satisfy the "unmeasured" first law of thermodynamics for all system states and system-only unitary processes; and (ii) the total work distribution will be identical to the system-only work distribution, for all system states and system-only unitary processes, if and only if the unmeasured work due to the unitary coupling between system and apparatus is zero for all system states.

I. INTRODUCTION

The definition of work for quantum systems is one of the most contentious issues in quantum thermodynamics, and continues to be a subject of heated debate [1–12]. The paradigmatic scenario is the work done on a thermally isolated system: a system which is only mechanically manipulated, by means of inducing time-dependence on its Hamiltonian, and thus evolves unitarily. Such mechanical manipulation results from an interaction with an external work source, and is hence generally accompanied by an exchange of work. In the limiting case where the system starts and ends in a classical mixture of energy eigenstates, in any given realization the work done on the system is well defined, and is the difference in energy eigenvalues. By performing ideal energy measurements before, and after, the unitary evolution, one can therefore observe which particular value of work obtains in any given realization without disturbing the system. Furthermore, the average work done, given by the observed probability distribution over work, will be equivalent to the difference in average energies evaluated before, and after, the unitary evolution; the "unmeasured" first law of thermodynamics is satisfied. The two-point energy measurement (TPM) protocol extends this procedure for determining the work distribution, namely, performing ideal energy measurements before and after the unitary evolution, to general unitary processes and general states [13, 14]. However, in general if the initial state does not commute with the Hamiltonian, the unmeasured first law will be violated; the average work obtained by the TPM protocol will not coincide with the difference in average energies. Indeed, as shown in Ref. [15], no measurement procedure exists which simultaneously recovers the work distribution for systems in a classical mixture of energy eigenstates, and recovers the

average work as the unmeasured work, for all states and unitary processes.

That the TPM protocol cannot always satisfy the unmeasured first law ultimately rests on one of the central maxims of quantum measurement theory: no information without disturbance [16]. To be sure, ideal measurements are the least disturbing measurements available [17, 18], but only insofar as there are *some* states that are undisturbed by such measurements. This perceived failure of the TPM definition has lead to alternative formulations of work, such as defining work as the unmeasured work *simpliciter* [19–21], and the Margenau-Hill method and related approaches using quasi-probability distributions [22–26].

Of course, there is another issue raised by the TPM protocol, or indeed any method which uses measurement as part of the definition for work: can such a method be self-consistent? The ideal energy measurements used in the TPM protocol must ultimately result from a *physical* interaction between the system and a measurement apparatus. The quantum theory of measurement allows for the measurement of any observable to be physically modeled as a *normal measurement scheme*, which involves a unitary interaction between the system and a measurement apparatus which is initially prepared in a fixed pure state – a condition which is possible to satisfy in principle, thermodynamic limitations on preparing pure states notwithstanding [27–30] – followed by measurement of the apparatus by a sharp *pointer observable* [31]. Normal measurement schemes have been used to "indirectly" measure work [5, 32]. Of course, such unitary interactions between the system and apparatus themselves result from mechanically manipulating this compound system, and hence are generally accompanied with an exchange of work with an external work source. If the TPM definition of work is valid for the system of interest, therefore, it stands to reason that it is valid for the compound of system plus apparatus; by performing ideal energy measurements on both system and apparatus, before and

after the total unitary evolution of the compound system, we thus obtain the total work distribution. We shall say that a given measurement scheme for the TPM protocol is self-consistent if the marginal work distribution for the system, obtained when ignoring the apparatus degrees of freedom, is identical to the original system-only TPM work distribution, for all system states and system-only unitary processes. This idea is inspired by the theory of quantum incompatibility, where two observables are said to be compatible if there exists a third observable, the marginals of which recover the original observables in question [33]. In the present manuscript, we show that such self-consistency is always achieved if the apparatus is initially prepared in an energy eigenstate.

Interestingly, if we further restrict the measurement scheme such that the apparatus Hamiltonian is equivalent to the pointer observable used to measure the apparatus, then the total work distribution will always satisfy the unmeasured first law; the average total work will be the difference in average energy given the total unitary evolution, for all system states and system-only unitary processes (i.e. excluding the apparatus state, and the unitary interaction between system and apparatus, which are fixed by the chosen measurement scheme). This is a consequence of the strong repeatability of ideal energy measurements [18], which implies that given the unitary interaction between system and apparatus, followed by measurement of the apparatus by the pointer observable, "directly" performing an ideal energy measurement on the system is superfluous. Of course, this statement should not be taken as a refutation of Ref. [15], since the initial state of the apparatus is always fixed, and commutes with the Hamiltonian by construction. But this observation does illustrate that it is possible for the TPM work distribution to satisfy the unmeasured first law for a large class of initial states that do not commute with the Hamiltonian.

Finally, in the case where the apparatus is initially prepared in an energy eigenstate, and the apparatus Hamiltonian is equivalent to the pointer observable, we show that the total work distribution will be identical with the system-only work distribution, for all system states and system-only unitary processes, if and only if the subspace of the apparatus which is involved during the measurement process corresponds with a single degenerate subspace of the apparatus Hamiltonian. This condition is further shown to be equivalent to the statement that the unmeasured work, due to the unitary interaction between system and apparatus, vanishes for all system states.

II. TPM PROTOCOL

We consider systems with a separable Hilbert space $\mathcal{H}$, with $\mathcal{L}(\mathcal{H})$ the algebra of bounded operators on $\mathcal{H}$, $\mathcal{T}(\mathcal{H}) \subseteq \mathcal{L}(\mathcal{H})$ the space of trace-class operators, and $\mathcal{S}(\mathcal{H}) \subset \mathcal{T}(\mathcal{H})$ the space of positive unit-trace operators (states), respectively. Moreover, we shall assume that the system is thermally isolated, with a bounded, time-dependent Hamilto-

nian $H(t) = H + H_I(t)$. Here, H is the system's "bare" Hamiltonian, describing it when it is fully isolated, i.e., isolated both thermally and mechanically. We assume this Hamiltonian to have a discrete spectrum, and may thus write it as

$$H = \sum_m \epsilon_m P_m. \tag{1}$$

Here, ϵ_m are energy eigenvalues, and $P_m \geqslant \mathbb{O}$ the corresponding spectral projections such that $P_m P_n = \delta_{m,n} P_m$ and $\sum_m P_m = \mathbb{1}$. By the spectral theorem, the bare Hamiltonian H is associated with a discrete, sharp observable $P := \{P_m\}$, where m are the measurement outcomes which, given a state preparation $\rho \in \mathcal{S}(\mathcal{H})$, are observed with the probability $\mathrm{tr}[P_m \rho]$ [34].

The time-dependence of $H(t)$ is entirely due to the term $H_I(t)$, which results from mechanically coupling the system with an external work source. If we assume that the system is only coupled with the work source for times $t \in (t_0, t_1)$, such that $H_I(t) = \mathbb{O}$ for all $t \leqslant t_0$ and $t \geqslant t_1$, then the system's time evolution due to its interaction with the work source will be described by the unitary operator $V := \overleftarrow{T} \exp(-i \int_{t_0}^{t_1} dt\, H(t))$, where we note that throughout this manuscript we use $\hbar = 1$ [23]. Given an initial state preparation ρ, the unmeasured work is thus

$$W := \mathrm{tr}[(V^\dagger H V - H)\rho]. \tag{2}$$

The TPM protocol, for revealing the distribution of work due to the interaction between the system and the work source, is given by the following sequence of operations:

(i) At time $t = t_0$, perform an ideal measurement of the bare Hamiltonian on the system, which is initially in an arbitrary state ρ. Given that outcome m is observed, the system will be prepared in the (unnormalized) state

$$P_m \rho P_m. \tag{3}$$

(ii) Between time t_0 and t_1, let the system evolve unitarily, given its interaction with the external work source. The system will thus be prepared in the (unnormalized) state

$$V P_m \rho P_m V^\dagger. \tag{4}$$

(iii) At time $t = t_1$, perform an ideal measurement of the bare Hamiltonian on the system. Given that outcome n is observed, the system will be prepared in the (unnormalized) state

$$P_n V P_m \rho P_m V^\dagger P_n. \tag{5}$$

The sequence of energy measurement outcomes $x := (m, n)$ thus corresponds with the work done $w(x) := \epsilon_n - \epsilon_m$, and its probability is given by the Born rule as

the trace of the final unnormalized state Eq. (5), which reads

$$p_\rho^V(x) := \mathrm{tr}[P_m V^\dagger P_n V P_m \rho]. \tag{6}$$

Therefore the probability distribution for the work done, w, given the initial state ρ and unitary operator V, is

$$p_\rho^V(w) := \sum_x \delta(w - w(x)) p_\rho^V(x), \tag{7}$$

where $\delta(a - b) = 1$ if $a = b$, and is zero otherwise. The average work can thus be computed to be

$$
\begin{aligned}
\langle w \rangle_\rho^V &:= \sum_w p_\rho^V(w) w \equiv \sum_x p_\rho^V(x) w(x), \\
&= \mathrm{tr}[(\mathcal{I}_\mathcal{M}^L(V^\dagger H V) - H)\rho], \\
&\equiv \mathrm{tr}[(V^\dagger H V - H)\mathcal{I}_\mathcal{M}^L(\rho)], \tag{8}
\end{aligned}
$$

where $\mathcal{I}_\mathcal{M}^L(\cdot) := \sum_m P_m(\cdot)P_m$ is the Lüders channel for the bare Hamiltonian H. Given that for any $A \in \mathcal{L}(\mathcal{H})$, $\mathcal{I}_\mathcal{M}^L(A) = A$ if and only if $[H, A] = \mathbb{O}$ [35], it follows that $\langle w \rangle_\rho^V = \mathrm{tr}[(V^\dagger H V - H)\rho]$ for all V (for all ρ) only if $[H, \rho] = \mathbb{O}$ ($[H, V^\dagger H V] = \mathbb{O}$). In other words, the unmeasured first law Eq. (2) cannot be satisfied for all states and all unitary processes.

A. Introducing the measurement apparatus in the TPM protocol

As shown above, the TPM protocol relies on performing ideal energy measurements on the system of interest both before, and after, the unitary evolution V. Such measurements are physically realized by an appropriate interaction between the system of interest and a measurement apparatus. The quantum theory of measurement allows all measurements on the system of interest to be modeled as a *normal measurement scheme* [31, 36]. Here, the system of interest first interacts with a quantum "probe" of a measurement apparatus, initially prepared in a fixed pure state, by an appropriate unitary operator. Subsequently, the probe is measured by an appropriate pointer observable, and the measurement outcome observed indicates that the corresponding outcome has been observed for the desired system observable.

Since two energy measurements are performed on the system during the TPM protocol, we can generally consider the apparatus to be composed of two probes, one of which interacts with the system at time $t = t_0$, and the other at time $t = t_1$. As such, for the ideal energy measurement performed at time t_j, we may mathematically describe the normal measurement scheme by the tuple $(\mathcal{H}_\mathcal{A}^{(j)}, |\xi^{(j)}\rangle, U^{(j)}, Z^{(j)})$, where: $\mathcal{H}_\mathcal{A}^{(j)}$ is the Hilbert space for the probe used, which is initially prepared in the pure state $|\xi^{(j)}\rangle$; $Z^{(j)} := \{Z_m^{(j)}\}$ is a sharp pointer observable, which has the same outcomes as the system observable $P := \{P_m\}$; and $U^{(j)}$ is a joint unitary operator on the

compound Hilbert space $\mathcal{H} \otimes \mathcal{H}_\mathcal{A}^{(j)}$. This normal measurement scheme will realize an ideal measurement of the bare Hamiltonian on the system of interest if, for all $T \in \mathcal{T}(\mathcal{H})$ and m, we have

$$\mathrm{tr}_{\mathcal{H}_\mathcal{A}^{(j)}}[(\mathbb{1} \otimes Z_m^{(j)})U^{(j)}(T \otimes P[\xi^{(j)}])U^{(j)\dagger}] = P_m T P_m, \tag{9}$$

where $P[\xi^{(j)}] \equiv |\xi^{(j)}\rangle\langle\xi^{(j)}|$ is a projection on the unit vector $|\xi^{(j)}\rangle \in \mathcal{H}_\mathcal{A}^{(j)}$, and $\mathrm{tr}_{\mathcal{H}_\mathcal{A}^{(j)}} : \mathcal{T}(\mathcal{H} \otimes \mathcal{H}_\mathcal{A}^{(j)}) \to \mathcal{T}(\mathcal{H})$ is the partial trace over the probe [37, 38]. It is simple to verify that in order for the unitary $U^{(j)}$ to satisfy Eq. (9), it must satisfy

$$U^{(j)}(|\psi\rangle \otimes |\xi^{(j)}\rangle) = \sum_m P_m|\psi\rangle \otimes |\phi_m^{(j)}\rangle \tag{10}$$

for all $|\psi\rangle \in \mathcal{H}$, where $|\phi_m^{(j)}\rangle$ are eigenvalue-1 eigenstates of the projection operators $Z_m^{(j)}$, i.e., $Z_n^{(j)}|\phi_m^{(j)}\rangle = \delta_{m,n}|\phi_m^{(j)}\rangle$ [39].

The TPM protocol can now be performed as follows:

(i) At time $t = t_0$, bring the system, initially prepared in an arbitrary state ρ, in contact with probe $\mathcal{H}_\mathcal{A}^{(0)}$. The state of the compound system is thus $\rho \otimes P[\xi^{(0)}]$. Subsequently let the system interact with the probe by the unitary operator $U^{(0)}$, which prepares the state

$$
\begin{aligned}
&U^{(0)}(\rho \otimes P[\xi^{(0)}])U^{(0)\dagger} \\
&= \sum_{m,m'} P_m \rho P_{m'} \otimes |\phi_m^{(0)}\rangle\langle\phi_{m'}^{(0)}|. \tag{11}
\end{aligned}
$$

Finally, perform a measurement of the probe by the pointer observable $Z^{(0)}$. Given that outcome m is observed, the system will be prepared in the (unnormalized) state

$$P_m \rho P_m. \tag{12}$$

(ii) Between time t_0 and t_1, let the system evolve unitarily, given its interaction with the external work source. The system will thus be prepared in the (unnormalized) state

$$V P_m \rho P_m V^\dagger. \tag{13}$$

(iii) At time $t = t_1$, bring the system in contact with probe $\mathcal{H}_\mathcal{A}^{(1)}$. The (unnormalized) state of the compound system is thus $V P_m \rho P_m V^\dagger \otimes P[\xi^{(1)}]$. Subsequently let the system interact with the probe by the unitary operator $U^{(1)}$, which prepares the (unnormalized) state

$$
\begin{aligned}
&U^{(1)}(V P_m \rho P_m V^\dagger \otimes P[\xi^{(1)}])U^{(1)\dagger} \\
&= \sum_{n,n'} P_n V P_m \rho P_m V^\dagger P_{n'} \otimes |\phi_n^{(1)}\rangle\langle\phi_{n'}^{(1)}|. \tag{14}
\end{aligned}
$$

Finally, perform a measurement of the probe by the pointer observable $Z^{(1)}$. Given that outcome n is observed, the system will be prepared in the (unnormalized) state

$$P_n V P_m \rho P_m V^\dagger P_n. \tag{15}$$

It is evident that the measurement scheme described above is identical to the original TPM protocol involving "direct" measurements on the system.

III. CONSISTENTLY APPLYING THE TPM PROTOCOL TO BOTH SYSTEM AND APPARATUS

The measurement scheme introduced in Sec. II A does not make any assumptions regarding the Hamiltonian of the apparatus probes, nor the time it takes for the unitary operators $U^{(j)}$ to be generated; indeed, these were assumed to be implemented instantaneously. However, for the unitary operator $U^{(j)}$ on $\mathcal{H} \otimes \mathcal{H}_\mathcal{A}^{(j)}$ to be physical, it must also result from mechanically manipulating the Hamiltonian of this composite system, and thus requires an interaction with an external work source for a finite duration [40]. Let us therefore write the total time-dependent Hamiltonian as $H_{\mathrm{tot}}(t) = H_{\mathrm{tot}} + H_I(t) + H_{\mathrm{int}}^{(0)}(t) + H_{\mathrm{int}}^{(1)}(t)$, where $H_{\mathrm{tot}} = H + H_\mathcal{A}^{(0)} + H_\mathcal{A}^{(1)}$ is the additive, total bare Hamiltonian of system plus apparatus, and $H_I(t)$ is the system-only interaction Hamiltonian introduced in Sec. II. We shall denote the bare Hamiltonian of each probe in the spectral form as

$$H_\mathcal{A}^{(j)} = \sum_\mu \lambda_\mu^{(j)} Q_\mu^{(j)}, \tag{16}$$

where $\lambda_\mu^{(j)}$ are energy eigenvalues and $Q_\mu^{(j)}$ the spectral projections. The interaction Hamiltonian for the composite system $\mathcal{H} \otimes \mathcal{H}_\mathcal{A}^{(j)}$, due to coupling with an external work source, is denoted $H_{\mathrm{int}}^{(j)}(t)$. Moreover, $H_{\mathrm{int}}^{(0)}(t) = \mathbb{0}$ for all $t \leqslant t_0'$ and $t \geqslant t_0$, and similarly $H_{\mathrm{int}}^{(1)}(t) = \mathbb{0}$ for all $t \leqslant t_1$ and $t \geqslant t_1'$, where $t_0' < t_0 < t_1 < t_1'$. In other words, the interaction Hamiltonian $H_{\mathrm{int}}^{(0)}(t)$ is non-vanishing only for a finite duration before the system undergoes its isolated unitary evolution V, and similarly $H_{\mathrm{int}}^{(1)}(t)$ is non-vanishing only for a finite duration after the system undergoes its isolated unitary evolution V. Therefore, by choosing the interaction Hamiltonians $H_{\mathrm{int}}^{(j)}(t)$ appropriately so that

$$\overleftarrow{T} \exp\left(-i \int_{t_0'}^{t_0} dt \left[H + H_\mathcal{A}^{(0)} + H_{\mathrm{int}}^{(0)}(t)\right]\right) = U^{(0)},$$

$$\overleftarrow{T} \exp\left(-i \int_{t_1}^{t_1'} dt \left[H + H_\mathcal{A}^{(1)} + H_{\mathrm{int}}^{(1)}(t)\right]\right) = U^{(1)}, \tag{17}$$

the total unitary operator which describes the compound system's evolution during the extended period $t \in (t_0', t_1')$ will be

$$V_{\mathrm{tot}} := \overleftarrow{T} \exp\left(-i \int_{t_0'}^{t_1'} dt\, H_{\mathrm{tot}}(t)\right),$$

$$= U^{(1)}(V \otimes e^{-i\theta_0 H_\mathcal{A}^{(0)}} \otimes e^{-i\theta_1 H_\mathcal{A}^{(1)}}) U^{(0)},$$

$$\equiv e^{-i\theta_0 H_\mathcal{A}^{(0)}} U^{(1)} V U^{(0)} e^{-i\theta_1 H_\mathcal{A}^{(1)}}. \tag{18}$$

Here, $e^{-i\theta_j H_\mathcal{A}^{(j)}}$, where $\theta_0 = t_1' - t_0$ and $\theta_1 = t_1 - t_0'$, describes the contribution to the total unitary evolution from the bare Hamiltonian of probe $\mathcal{H}_\mathcal{A}^{(j)}$, i.e., for the time period where the interaction Hamiltonian $H_{\mathrm{int}}^{(j)}(t)$ vanishes. Note that the final line of Eq. (18) is obtained because the unitary operators $e^{-i\theta_0 H_\mathcal{A}^{(0)}}$ and $e^{-i\theta_1 H_\mathcal{A}^{(1)}}$ commute with $U^{(1)}$ and $U^{(0)}$, respectively, since they act on different Hilbert spaces. Given an initial state preparation $\rho \in \mathcal{S}(\mathcal{H})$, the total unmeasured work will thus read

$$W_{\mathrm{tot}} := \mathrm{tr}[(V_{\mathrm{tot}}^\dagger H_{\mathrm{tot}} V_{\mathrm{tot}} - H_{\mathrm{tot}})\rho \otimes P[\xi]], \tag{19}$$

where $|\xi\rangle := |\xi^{(0)}\rangle \otimes |\xi^{(1)}\rangle$ is the initial state of the apparatus $\mathcal{H}_\mathcal{A} := \mathcal{H}_\mathcal{A}^{(0)} \otimes \mathcal{H}_\mathcal{A}^{(1)}$, composed of both probes.

Now we may perform the TPM protocol on the total compound system so as to determine the total work distribution given the total unitary operator in Eq. (18). For this to be consistent with the original TPM protocol on the system alone, however, we require that when averaging out the energy measurements performed on the apparatus, we must obtain the probability distribution given in Eq. (6), for all system states $\rho \in \mathcal{S}(\mathcal{H})$ and system-only unitary operators V. In order for this to be satisfied, we demand that $|\xi^{(j)}\rangle$ be an eigenstate of the probe Hamiltonian $H_\mathcal{A}^{(j)}$, with eigenvalue $\lambda_0^{(j)}$. This will ensure that the initial ideal energy measurement of the apparatus will not disturb it, so that the unitary interaction between system and apparatus by the unitary operators $U^{(j)}$ will result in the same state transformation as discussed in Sec. II A.

Let us first note that, given the assumption that the apparatus is initially prepared in an energy eigenstate, and using Eq. (10) and Eq. (18), we can show that for all $|\psi\rangle \in \mathcal{H}$,

$$V_{\mathrm{tot}}(|\psi\rangle \otimes |\xi\rangle)$$
$$= e^{-i\theta_1 \lambda_0^{(1)}} \sum_{m,n} P_n V P_m |\psi\rangle \otimes e^{-i\theta_0 H_\mathcal{A}^{(0)}} |\phi_m^{(0)}\rangle \otimes |\phi_n^{(1)}\rangle. \tag{20}$$

Here, we have used the fact that $|\xi^{(1)}\rangle$ is an energy eigenstate with eigenvalue $\lambda_0^{(1)}$ to infer that the component of V_{tot} given by $e^{-i\theta_1 H_\mathcal{A}^{(1)}}$ only induces a constant phase factor $e^{-i\theta_1 \lambda_0^{(1)}}$, which is not physically observable. Using this, we may now examine the extended TPM protocol, which will be as follows:

(i) At time $t = t_0'$, perform an ideal energy measurement on the compound system $\mathcal{H} \otimes \mathcal{H}_\mathcal{A}$, initially prepared in

the state $\rho \otimes P[\xi]$. Since the probes $\mathcal{H}_{\mathcal{A}}^{(j)}$ are initially prepared in an energy eigenstate with energy eigenvalue $\lambda_0^{(j)}$, only outcomes $(m, 0, 0)$ are observed with non-zero probability, which result in the compound system being prepared in the (unnormalized) state

$$P_m \rho P_m \otimes Q_0^{(0)} P[\xi^{(0)}] Q_0^{(0)} \otimes Q_0^{(1)} P[\xi^{(1)}] Q_0^{(1)}$$
$$= P_m \rho P_m \otimes P[\xi^{(0)}] \otimes P[\xi^{(1)}]. \tag{21}$$

(ii) Between time t_0' and t_1', let the compound system evolve according to the total unitary operator V_{tot} defined in Eq. (18) and Eq. (20). This prepares the (unnormalized) state

$$V_{\text{tot}}(P_m \rho P_m \otimes P[\xi^{(0)}] \otimes P[\xi^{(1)}]) V_{\text{tot}}^\dagger$$
$$= \sum_{n,n'} P_n V P_m \rho P_m V^\dagger P_{n'} \otimes \tilde{P}[\phi_m^{(0)}] \otimes |\phi_n^{(1)}\rangle\langle\phi_{n'}^{(1)}|, \tag{22}$$

where $\tilde{P}[\phi_m^{(0)}] := e^{-i\theta_0 H_{\mathcal{A}}^{(0)}} P[\phi_m^{(0)}] e^{i\theta_0 H_{\mathcal{A}}^{(0)}}$.

(iii) At time $t = t_1'$, perform an ideal energy measurement on the compound system $\mathcal{H} \otimes \mathcal{H}_{\mathcal{A}}$. Given the outcomes (n, μ, ν), this prepares the (unnormalized) state

$$P_n V P_m \rho P_m V^\dagger P_n \otimes Q_\mu^{(0)} \tilde{P}[\phi_m^{(0)}] Q_\mu^{(0)} \otimes Q_\nu^{(1)} P[\phi_n^{(1)}] Q_\nu^{(1)}. \tag{23}$$

The full sequence of measurement outcomes is thus $X := (x, (0, \mu), (0, \nu))$, where $x := (m, n)$ is the sequence of outcomes for the system, while $(0, \mu)$ and $(0, \nu)$ are the sequences of outcomes for probes $\mathcal{H}_{\mathcal{A}}^{(0)}$ and $\mathcal{H}_{\mathcal{A}}^{(1)}$, respectively. The sequence X corresponds with the total work done $\mathcal{W}(X) := w(x) + w_{\mathcal{A}}^{(0)}(\mu) + w_{\mathcal{A}}^{(1)}(\nu)$, where $w(x) := \epsilon_n - \epsilon_m$ is the contribution to the total work from the system, while $w_{\mathcal{A}}^{(j)}(\mu) := \lambda_\mu^{(j)} - \lambda_0^{(j)}$ is the contribution to the total work from probe $\mathcal{H}_{\mathcal{A}}^{(j)}$. The probability of observing sequence X, meanwhile, is given by the trace of the final unnormalized state Eq. (23), which is

$$p_{\rho,\xi}^{V_{\text{tot}}}(X) := p_\rho^V(x) \text{tr}[Q_\mu^{(0)} P[\phi_m^{(0)}]] \text{tr}[Q_\nu^{(1)} P[\phi_n^{(1)}]], \tag{24}$$

where we recall that $p_\rho^V(x)$ is defined in Eq. (6). Note that here, we have used the fact that $Q_\mu^{(0)}$ is a spectral projection of $H_{\mathcal{A}}^{(0)}$ to infer that $\text{tr}[Q_\mu^{(0)} \tilde{P}[\phi_m^{(0)}]] = \text{tr}[Q_\mu^{(0)} e^{-i\theta_0 H_{\mathcal{A}}^{(0)}} P[\phi_m^{(0)}] e^{i\theta_0 H_{\mathcal{A}}^{(0)}}] = \text{tr}[Q_\mu^{(0)} P[\phi_m^{(0)}]]$.

Given that $\sum_\mu Q_\mu^{(0)} = \sum_\nu Q_\nu^{(1)} = \mathbb{1}$, the marginal probability distribution for the system-only work will read as

$$\sum_{\mu,\nu} p_{\rho,\xi}^{V_{\text{tot}}}(X) = p_\rho^V(x) \text{tr}[P[\phi_m^{(0)}]] \text{tr}[P[\phi_n^{(1)}]] = p_\rho^V(x), \tag{25}$$

and so the extended TPM protocol on the compound of system plus apparatus is self-consistent.

A. Satisfying the unmeasured first law for the total work

Since the apparatus is initially prepared in an energy eigenstate, the average total work, for the total unitary process discussed in the previous section, clearly satisfies

$$\langle \mathcal{W} \rangle_{\rho,\xi}^{V_{\text{tot}}} := \sum_X p_{\rho,\xi}^{V_{\text{tot}}}(X) \mathcal{W}(X),$$
$$= \text{tr}[(V_{\text{tot}}^\dagger H_{\text{tot}} V_{\text{tot}} - H_{\text{tot}}) \mathcal{I}_{\mathcal{M}}^L(\rho) \otimes P[\xi]]. \tag{26}$$

To see this, simply compare with Eq. (8). As before, if ρ does not commute with the Hamiltonian H, the total work is not guaranteed to satisfy the unmeasured first law, i.e., it is possible for some ρ and V to have Eq. (26) differ from Eq. (19) (note that both the apparatus state $|\xi\rangle$, and the contribution to V_{tot} from the system-apparatus coupling, i.e., the unitaries $U^{(j)}$, are always fixed). However, as we shall show below, if additionally the probe Hamiltonians $H_{\mathcal{A}}^{(j)}$ are equivalent to the pointer observables $Z^{(j)}$, i.e., if we have

$$H_{\mathcal{A}}^{(j)} = \sum_m \lambda_m^{(j)} Z_m^{(j)}, \tag{27}$$

the unmeasured first law is guaranteed to be satisfied for the total work.

Let us re-examine the TPM protocol on the compound of system plus apparatus once more in detail, this time assuming that Eq. (27) holds:

(i) At time $t = t_0'$ perform an ideal energy measurement on the compound system $\mathcal{H} \otimes \mathcal{H}_{\mathcal{A}}$, initially prepared in the state $\rho \otimes P[\xi]$. As before, we assume that $|\xi^{(j)}\rangle$ are energy eigenstates, with eigenvalues $\lambda_0^{(j)}$, and hence only the outcomes $(m, 0, 0)$ are observed with non-zero probability, resulting in the compound system to be prepared in the (unnormalized) state

$$P_m \rho P_m \otimes Z_0^{(0)} P[\xi^{(0)}] Z_0^{(0)} \otimes Z_0^{(1)} P[\xi^{(1)}] Z_0^{(1)}$$
$$= P_m \rho P_m \otimes P[\xi^{(0)}] \otimes P[\xi^{(1)}]. \tag{28}$$

(ii) Between time t_0' and t_1', let the compound system evolve according to the total unitary operator V_{tot} defined in Eq. (18) and Eq. (20). This prepares the (unnormalized) state

$$V_{\text{tot}}(P_m \rho P_m \otimes P[\xi^{(0)}] \otimes P[\xi^{(1)}]) V_{\text{tot}}^\dagger$$
$$= \sum_{n,n'} P_n V P_m \rho P_m V^\dagger P_{n'} \otimes P[\phi_m^{(0)}] \otimes |\phi_n^{(1)}\rangle\langle\phi_{n'}^{(1)}|. \tag{29}$$

Note that since $|\phi_m^{(0)}\rangle$ are eigenvalue-1 eigenstates of the projection operators $Z_m^{(0)}$, which clearly commute with the Hamiltonian, it follows that $e^{-i\theta_0 H_{\mathcal{A}}^{(0)}} P[\phi_m^{(0)}] e^{i\theta_0 H_{\mathcal{A}}^{(0)}} = P[\phi_m^{(0)}]$.

(iii) At time $t = t'_1$, perform an ideal energy measurement on the compound system $\mathcal{H} \otimes \mathcal{H}_\mathcal{A}$. Given outcomes (n, n', n''), this prepares the (unnormalized) state

$$P_n V P_m \rho P_m V^\dagger P_n \otimes Z_{n'}^{(0)} P[\phi_m^{(0)}] Z_{n'}^{(0)} \otimes Z_{n''}^{(1)} P[\phi_n^{(1)}] Z_{n''}^{(1)}$$
$$= \delta_{m,n'} \delta_{n,n''} P_n V P_m \rho P_m V^\dagger P_n \otimes P[\phi_m^{(0)}] \otimes P[\phi_n^{(1)}], \tag{30}$$

where the final line follows from the fact that $Z_n^{(j)} |\phi_m^{(j)}\rangle = \delta_{m,n} |\phi_m^{(j)}\rangle$.

Equation (30) implies that the only sequences of energy measurement outcomes that are observed with non-zero probability are $X := (x, (0, m), (0, n))$, where we recall that $x := (m, n)$. In other words, the energy transitions of the apparatus fully determine the energy transitions of the system, and vice versa. As such, let us remove some of the redundancy and write $X := ((0,0), x)$, where $(0,0)$ denotes the energy measurement outcomes on the apparatus at time t'_0, and $x = (m, n)$ denotes both the energy measurement outcomes on the apparatus at time t'_1, as well as the sequence of energy measurement outcomes on the system at times t'_0, t'_1. The total work done given the sequence X is thus $\mathcal{W}(X) := w(x) + w_\mathcal{A}^{(0)}(m) + w_\mathcal{A}^{(1)}(n)$, where $w_\mathcal{A}^{(j)}(m) := \lambda_m^{(j)} - \lambda_0^{(j)}$, with the probability

$$p_{\rho,\xi}^{V_\mathrm{tot}}(X) = p_\rho^V(x). \tag{31}$$

Note that this is equivalent to Eq. (24) when we replace $Q_\mu^{(j)}$ with $Z_m^{(j)}$, which gives $\mathrm{tr}[Z_m^{(j)} P[\phi_m^{(j)}]] = 1$.

As shown in Appendix (A), the average total work will now read as

$$\langle \mathcal{W} \rangle_{\rho,\xi}^{V_\mathrm{tot}} := \sum_X p_{\rho,\xi}^{V_\mathrm{tot}}(X) \mathcal{W}(X),$$
$$= \mathrm{tr}[(V_\mathrm{tot}^\dagger H_\mathrm{tot} V_\mathrm{tot} - H_\mathrm{tot})(\rho \otimes P[\xi])] \tag{32}$$

for all $\rho \in \mathcal{S}(\mathcal{H})$ and V. As such, comparing with Eq. (19) we see that so long as the apparatus is initially prepared in an energy eigenstate, and Eq. (27) is satisfied, then not only will the TPM protocol for the total work be self-consistent, but the total work will always satisfy the unmeasured first law, even for initial system states ρ that do not commute with the Hamiltonian. As a final observation, note that Eq. (26) must be equivalent to Eq. (32) when Eq. (27) is satisfied. Consequently, in such a case the following equality holds:

$$\mathrm{tr}[(V_\mathrm{tot}^\dagger H_\mathrm{tot} V_\mathrm{tot} - H_\mathrm{tot})(\rho \otimes P[\xi])]$$
$$= \mathrm{tr}[(V_\mathrm{tot}^\dagger H_\mathrm{tot} V_\mathrm{tot} - H_\mathrm{tot})(\mathcal{I}_\mathcal{M}^L(\rho) \otimes P[\xi])] \tag{33}$$

for all $\rho \in \mathcal{S}(\mathcal{H})$ and V. This is a consequence of the strong repeatability of ideal energy measurements [18], which implies that directly performing ideal energy measurements on the system is redundant; the structure of the unitary operators $U^{(j)}$, together with the fact that we measure the probes by the pointer observables $Z^{(j)}$, ensures that the system automatically undergoes an ideal energy measurement.

B. Necessary and sufficient conditions for the total work distribution to be equal to the system-only work distribution

If the apparatus is initially prepared in an energy eigenstate, and the apparatus probe Hamiltonians are equivalent to the pointer observable, the probability distribution for the total work $\mathcal{W}$, given an initial total state $\rho \otimes P[\xi]$ and total unitary operator V_tot, is

$$p_{\rho,\xi}^{V_\mathrm{tot}}(\mathcal{W}) := \sum_X \delta(\mathcal{W} - \mathcal{W}(X)) p_{\rho,\xi}^{V_\mathrm{tot}}(X),$$
$$\equiv \sum_x \delta(\mathcal{W} - (w(x) + w_\mathcal{A}(x))) p_\rho^V(x), \tag{34}$$

where we have used Eq. (31), together with the definition $w_\mathcal{A}(x) := w_\mathcal{A}^{(0)}(m) + w_\mathcal{A}^{(1)}(n)$.

It is simple to see that, in general, the total work probability distribution Eq. (34) is different to the system-only work probability distribution Eq. (7). In order for these distributions to be the same, for all system states ρ and system-only unitary operators V, we must have $w_\mathcal{A}(x) = 0$ for all x such that $p_\rho^V(x) > 0$ for some ρ and V. This ensures that for all ρ and V,

$$\sum_x \delta(w - w(x)) p_\rho^V(x) = \sum_x \delta(w - \mathcal{W}(X)) p_\rho^V(x). \tag{35}$$

Recall that $w_\mathcal{A}^{(j)}(m) := \lambda_m^{(j)} - \lambda_0^{(j)}$, where $\lambda_0^{(j)}$ is a fixed energy eigenvalue, and that $p_\rho^V(x) := \mathrm{tr}[P_m V^\dagger P_n V P_m \rho]$. Consequently, the condition $w_\mathcal{A}(x) = 0$ for all x such that $p_\rho^V(x) > 0$ for some ρ and V is equivalent to the condition $\lambda_m^{(j)} = \lambda_0^{(j)}$ for all m such that $P_m > \mathbb{0}$, i.e., if only a single degenerate energy subspace of the probe is involved during the measurement process. Interestingly, we shall see that this condition is equivalent to the statement that the unmeasured work given the unitary operator $U^{(j)}$ vanishes for all system states.

The unmeasured work, given the measurement unitary coupling between system and probe $\mathcal{H}_\mathcal{A}^{(j)}$, is given as

$$W_\mathrm{meas}^{(j)} := \mathrm{tr}[(U^{(j)\dagger} H_\mathrm{tot}^{(j)} U^{(j)} - H_\mathrm{tot}^{(j)}) \rho \otimes P[\xi^{(j)}]], \tag{36}$$

where we define $H_\mathrm{tot}^{(j)} := H + H_\mathcal{A}^{(j)}$ as the additive Hamiltonian of the composite system $\mathcal{H} \otimes \mathcal{H}_\mathcal{A}^{(j)}$. This can equivalently be written as

$$W_\mathrm{meas}^{(j)} = \mathrm{tr}[\Gamma_{\xi^{(j)}} \left(U^{(j)\dagger} H_\mathrm{tot}^{(j)} U^{(j)} - H_\mathrm{tot}^{(j)} \right) \rho], \tag{37}$$

where $\Gamma_{\xi^{(j)}} : \mathcal{L}(\mathcal{H} \otimes \mathcal{H}_\mathcal{A}^{(j)}) \to \mathcal{L}(\mathcal{H})$ is the restriction map for $|\xi^{(j)}\rangle$, defined by the identity $\mathrm{tr}[\Gamma_{\xi^{(j)}}(B) T] = \mathrm{tr}[B(T \otimes P[\xi^{(j)}])]$ for all $B \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H}_\mathcal{A}^{(j)})$ and $T \in \mathcal{T}(\mathcal{H})$ [41]. Recalling that the unitary operator $U^{(j)}$ always satisfies Eq. (10), we thus have

$$\Gamma_{\xi^{(j)}} \left(U^{(j)\dagger} H_\mathrm{tot}^{(j)} U^{(j)} - H_\mathrm{tot}^{(j)} \right) = \sum_m w_\mathcal{A}^{(j)}(m) P_m. \tag{38}$$

For a detailed proof, refer to Appendix (B). The right hand side of Eq. (38) vanishes if for each m, either $P_m = \mathbb{O}$, or $w_{\mathcal{A}}^{(j)}(m) = 0$. Consequently, $w_{\mathcal{A}}^{(j)}(m) = 0$ for all m such that $P_m > \mathbb{O}$ is necessary and sufficient for the left hand side of Eq. (38) to vanish. But by Eq. (37) this implies that $W_{\mathrm{meas}}^{(j)} = 0$ for all $\rho \in \mathcal{S}(\mathcal{H})$; given that the apparatus is in the state $|\xi^{(j)}\rangle$, then irrespective of what state the system is prepared in, the unmeasured work given the unitary operator $U^{(j)}$ will vanish. We refer to this as $U^{(j)}$ satisfying "weak" energy conservation, which is a weaker condition than full energy conservation, i.e., $[H_{\mathrm{tot}}^{(j)}, U^{(j)}] = \mathbb{O}$, which implies that $W_{\mathrm{meas}}^{(j)} = 0$ for all choices of the apparatus state $|\xi^{(j)}\rangle$.

We note that while a fully degenerate probe Hamiltonian, $H_{\mathcal{A}}^{(j)} = \lambda_0^{(j)} \mathbb{1}$, or a fully energy conserving unitary, $[H_{\mathrm{tot}}^{(j)}, U^{(j)}] = \mathbb{O}$, are sufficient conditions for the total work distribution Eq. (34) to equal the system-only work distribution Eq. (7), they are not necessary.

To illustrate the first point, consider the system Hamiltonian $H = \epsilon_1 P_1 + \epsilon_2 P_2$, where $P_1, P_2 > \mathbb{O}$. However, this is equivalent to $H = \epsilon_1 P_1 + \epsilon_2 P_2 + \epsilon_3 P_3$ such that $P_3 = \mathbb{O}$. Therefore, the ideal measurement of H can be realized by the normal measurement scheme $(\mathcal{H}_{\mathcal{A}}, |\xi\rangle, U, Z)$, with the three-valued pointer observable $Z := \{Z_1, Z_2, Z_3\}$, $Z_m > \mathbb{O}$, and the unitary operator U which satisfies

$$U(|\psi\rangle \otimes |\xi\rangle) = \sum_{m=1}^{3} P_m |\psi\rangle \otimes |\phi_m\rangle \qquad (39)$$

for all $|\psi\rangle \in \mathcal{H}$, where $|\phi_m\rangle$ are eigenvalue-1 eigenstates of Z_m. Note that the term for $m = 3$ vanishes, since $P_3|\psi\rangle = \mathbb{O}|\psi\rangle = 0$ for all $|\psi\rangle$; the apparatus is never taken to the state $|\phi_3\rangle$. Let the apparatus have the Hamiltonian $H_{\mathcal{A}} = \lambda(Z_1 + Z_2) + \lambda' Z_3$, where $\lambda \neq \lambda'$, so that $H_{\mathcal{A}}$ is not fully degenerate. Notwithstanding, if $|\xi\rangle$ is in the support of $Z_1 + Z_2$, we still have $w_{\mathcal{A}}(m) = 0$ for $m = 1, 2$, i.e., for all m corresponding to $P_m > \mathbb{O}$. As stated previously, it is only necessary that a single degenerate energy subspace of the apparatus be "involved" during the measurement process; for all measurement outcomes that are observed, the state of the apparatus starts and ends in the support of $Z_1 + Z_2$.

To illustrate that full energy conservation by the unitary is also not necessary, consider the simple case where $\mathcal{H} \simeq \mathbb{C}^2$, with orthonormal basis $\{|0\rangle, |1\rangle\}$, and Hamiltonian $H = \epsilon|1\rangle\langle 1|$, $\epsilon > 0$. A normal measurement scheme for an ideal measurement of H can be given as $(\mathcal{H}_{\mathcal{A}}, |0\rangle, U, Z)$, where

$\mathcal{H}_{\mathcal{A}} \simeq \mathbb{C}^2$, $Z := \{|0\rangle\langle 0|, |1\rangle\langle 1|\}$, and

$$U : \begin{cases} |m, 0\rangle \mapsto |m, m\rangle \\ |m, 1\rangle \mapsto |m \oplus_2 1, m\rangle \end{cases}, \qquad (40)$$

where $m = 0, 1$ and $\oplus_2$ denotes addition modulo 2. Note that only the transformation $|m, 0\rangle \mapsto |m, m\rangle$ is ever utilized, since the apparatus is initially prepared in state $|0\rangle$. If the apparatus Hamiltonian is fully degenerate, $H_{\mathcal{A}} = \lambda \mathbb{1}$, then U will satisfy weak energy conservation; given $H_{\mathrm{tot}} = H + H_{\mathcal{A}}$, then for any $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, we have $\langle \psi, 0|U^\dagger H_{\mathrm{tot}} U|\psi, 0\rangle = |\beta|^2\epsilon + \lambda = \langle \psi, 0|H_{\mathrm{tot}}|\psi, 0\rangle$. However, $[U, H_{\mathrm{tot}}] \neq \mathbb{O}$, since $U|1, 1\rangle = |0, 1\rangle$, and hence $\langle 1, 1|U^\dagger H_{\mathrm{tot}} U|1, 1\rangle = \lambda \neq \langle 1, 1|H_{\mathrm{tot}}|1, 1\rangle = \epsilon + \lambda$.

IV. CONCLUSIONS

A definition for work which relies on measurements is self-consistent if it can account for the contribution to work by the measurement process itself, at least in principle. More precisely, for self-consistency we demand that the marginal of the total work distribution for system and measurement apparatus, obtained by ignoring the apparatus degrees of freedom, recovers the original work distribution for the system alone. In the case of the two-point energy measurement (TPM) protocol, we have shown that this is possible so long as the measurement apparatus is initially prepared in an energy eigenstate. Furthermore, if the apparatus Hamiltonian is chosen to be equivalent to the pointer observable, then the total work distribution will always satisfy the unmeasured first law; the average total work will equal the change in average energy given the total unitary evolution. This is a consequence of the strong repeatability of ideal energy measurements, which implies that directly performing energy measurements on the system is redundant. Finally, we have shown that the total work distribution will be identical to the system-only work distribution if and only if the unmeasured work, given the unitary interaction between system and apparatus, vanishes for all system states. Extending the present framework of analysis to other definitions of work remain as open questions for further research.

[1] A. E. Allahverdyan and T. M. Nieuwenhuizen, Phys. Rev. E **71**, 066102 (2005).

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

[3] P. Skrzypczyk, A. J. Short, and S. Popescu, Nat. Commun. **5**, 4185 (2014).

[4] G. Watanabe, B. P. Venkatesh, and P. Talkner, Phys. Rev. E **89**, 052116 (2014).

[5] A. J. Roncaglia, F. Cerisola, and J. P. Paz, Phys. Rev. Lett. **113**, 250601 (2014).

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

[7] J. Gemmer and J. Anders, New J. Phys. **17**, 085006 (2015).

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

[9] M. Hayashi and H. Tajima, Phys. Rev. A **95**, 032132 (2017).

[10] P. Faist and R. Renner, Phys. Rev. X **8**, 021011 (2018).

[11] W. Niedenzu, M. Huber, and E. Boukobza, Quantum **3**, 195 (2019).

[12] A. Sone, Y.-X. Liu, and P. Cappellaro, Phys. Rev. Lett. **125**, 060602 (2020).

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

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

[15] M. Perarnau-Llobet, E. Bäumer, K. V. Hovhannisyan, M. Huber, and A. Acin, Phys. Rev. Lett. **118**, 070601 (2017).

[16] P. Busch, in Quantum Reality, Relativ. Causality, Closing Epistemic Circ. (Springer, Dordrecht, 2009) pp. 229–256.

[17] P. J. Lahti, P. Busch, and P. Mittelstaedt, J. Math. Phys. **32**, 2770 (1991).

[18] P. Busch, M. Grabowski, and P. J. Lahti, Found. Phys. **25**, 1239 (1995).

[19] J. Anders and V. Giovannetti, New J. Phys. **15**, 033022 (2013).

[20] S. Deffner, J. P. Paz, and W. H. Zurek, Phys. Rev. E **94**, 010103 (2016).

[21] P. Strasberg, Phys. Rev. E **100**, 022127 (2019).

[22] H. Margenau and R. N. Hill, Prog. Theor. Phys. **26**, 722 (1961).

[23] A. E. Allahverdyan, Phys. Rev. E **90**, 032137 (2014).

[24] H. J. D. Miller and J. Anders, New J. Phys. **19**, 062001 (2017).

[25] M. Lostaglio, Phys. Rev. Lett. **120**, 040602 (2018).

[26] M. G. Díaz, G. Guarnieri, and M. Paternostro, (2020), arXiv:2007.00042.

[27] L. Masanes and J. Oppenheim, Nat. Commun. **8**, 14538 (2017).

[28] T. Debarba, G. Manzano, Y. Guryanova, M. Huber, and N. Friis, New J. Phys. **21**, 113002 (2019).

[29] M. H. Mohammady and A. Romito, Quantum **3**, 175 (2019).

[30] Y. Guryanova, N. Friis, and M. Huber, Quantum **4**, 222 (2020).

[31] M. Ozawa, J. Math. Phys. **25**, 79 (1984).

[32] G. De Chiara, P. Solinas, F. Cerisola, and A. J. Roncaglia, in Thermodyn. Quantum Regime. Fundam. Theor. Phys. (Springer, 2018) pp. 337–362.

[33] T. Heinosaari, T. Miyadera, and M. Ziman, J. Phys. A Math. Theor. **49**, 123001 (2016).

[34] T. Heinosaari and M. Ziman, The Mathematical language of Quantum Theory (Cambridge University Press, Cambridge, 2011).

[35] L. Weihua and W. Junde, J. Phys. A Math. Theor. **43**, 395206 (2010).

[36] P. Busch, M. Grabowski, and P. J. Lahti, Operational Quantum Physics, Lecture Notes in Physics Monographs, Vol. 31 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1995).

[37] P. Busch, P. J. Lahti, and Peter Mittelstaedt, The Quantum Theory of Measurement, Lecture Notes in Physics Monographs, Vol. 2 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1996).

[38] P. Busch, P. Lahti, J.-P. Pellonpää, and K. Ylinen, Quantum Measurement, Theoretical and Mathematical Physics (Springer International Publishing, Cham, 2016).

[39] P. Mittelstaedt, The Interpretation of Quantum Mechanics and the Measurement Process (Cambridge University Press, Cambridge, 1997).

[40] P. Strasberg, K. Modi, and M. Skotiniotis, (2020), arXiv:2011.13254.

[41] L. Loveridge, T. Miyadera, and P. Busch, Found. Phys. **48**, 135 (2018).

Appendix A: Proof of Eq. (32)

Let us introduce the operation (completely positive and trace non-increasing map) $\mathcal{J}_{x',x} : \mathcal{T}(\mathcal{H} \otimes \mathcal{H}_\mathcal{A}) \to \mathcal{T}(\mathcal{H} \otimes \mathcal{H}_\mathcal{A})$, defined as

$$\mathcal{J}_{x',x}(\cdot) := (\mathbb{1} \otimes Z_x) V_{\text{tot}} (\mathbb{1} \otimes Z_{x'})(\cdot)(\mathbb{1} \otimes Z_{x'}) V_{\text{tot}}^\dagger (\mathbb{1} \otimes Z_x). \tag{A1}$$

Here, $x := (m,n)$ and $x' := (m',n')$, so that $Z_x := Z_m^{(0)} \otimes Z_n^{(1)}$ and $Z_{x'} := Z_{m'}^{(0)} \otimes Z_{n'}^{(1)}$, and V_{tot} is defined in Eq. (18). First, let us show that the state transformation given the TPM protocol can be fully described by the operation $\mathcal{J}_{x',x}$. Denoting $(0,0) \equiv 0$, we find that for any $T \in \mathcal{T}(\mathcal{H})$, $P[\xi] := P[\xi^{(0)}] \otimes P[\xi^{(1)}]$, and x, the following:

$$\begin{aligned}
\mathcal{J}_{0,x}(T \otimes P[\xi]) &= (\mathbb{1} \otimes Z_x) V_{\text{tot}} (\mathbb{1} \otimes Z_0)(T \otimes P[\xi])(\mathbb{1} \otimes Z_0) V_{\text{tot}}^\dagger (\mathbb{1} \otimes Z_x), \\
&= (\mathbb{1} \otimes Z_x) V_{\text{tot}} (T \otimes P[\xi]) V_{\text{tot}}^\dagger (\mathbb{1} \otimes Z_x), \\
&= \sum_{m',m'',n',n''} (\mathbb{1} \otimes Z_x) \left(P_{n'} V P_{m'} T P_{m''} V^\dagger P_{n''} \otimes |\phi_{m'}^{(0)}\rangle\langle\phi_{m''}^{(0)}| \otimes |\phi_{n'}^{(1)}\rangle\langle\phi_{n''}^{(1)}| \right) (\mathbb{1} \otimes Z_x), \\
&= P_n V P_m T P_m V^\dagger P_n \otimes P[\phi_m^{(0)}] \otimes P[\phi_n^{(1)}].
\end{aligned} \tag{A2}$$

In the second line, we have used the fact that $|\xi\rangle$ is an eigenvalue-1 eigenstate of $Z_0 = Z_0^{(0)} \otimes Z_0^{(1)}$. In the third line, we have used Eq. (18) and Eq. (20). In the final line, we have used the fact that $|\phi_m^{(j)}\rangle$ are eigenvalue-1 eigenstates of $Z_m^{(j)}$.

Similarly, it is easy to show that $\mathcal{J}_{x',x}(T \otimes P[\xi]) = \mathbb{O}$ for all $x' \neq (0,0)$.

Recall that, given the sequence $X := (0,x)$, the TPM work done is

$$
\begin{aligned}
\mathcal{W}(X) &:= w(x) + w_{\mathcal{A}}^{(0)}(m) + w_{\mathcal{A}}^{(1)}(n), \\
&= (\epsilon_n + \lambda_m^{(0)} + \lambda_n^{(1)}) - (\epsilon_m + \lambda_0^{(0)} + \lambda_0^{(1)}).
\end{aligned}
\tag{A3}
$$

Using Eq. (A2), recalling that $H_{\text{tot}} = H + H_{\mathcal{A}}^{(0)} + H_{\mathcal{A}}^{(1)}$, and that $|\xi\rangle$ and $|\phi_m^{(j)}\rangle$ are eigenstates of $H_{\mathcal{A}}^{(0)} + H_{\mathcal{A}}^{(1)}$, we may verify that

$$
\begin{aligned}
\operatorname{tr}[H_{\text{tot}}\mathcal{J}_{0,x}(\rho \otimes P[\xi])] &= (\epsilon_n + \lambda_m^{(0)} + \lambda_n^{(1)})\operatorname{tr}[\mathcal{J}_{0,x}(\rho \otimes P[\xi])], \\
\operatorname{tr}[\mathcal{J}_{0,x}(H_{\text{tot}}(\rho \otimes P[\xi]))] &= (\epsilon_m + \lambda_0^{(0)} + \lambda_0^{(1)})\operatorname{tr}[\mathcal{J}_{0,x}(\rho \otimes P[\xi])].
\end{aligned}
\tag{A4}
$$

Consequently, we may express Eq. (A3) as

$$
\mathcal{W}(X) = \frac{\operatorname{tr}[H_{\text{tot}}\mathcal{J}_{0,x}(\rho \otimes P[\xi])]}{\operatorname{tr}[\mathcal{J}_{0,x}(\rho \otimes P[\xi])]} - \frac{\operatorname{tr}[\mathcal{J}_{0,x}(H_{\text{tot}}\rho \otimes P[\xi])]}{\operatorname{tr}[\mathcal{J}_{0,x}(\rho \otimes P[\xi])]}.
\tag{A5}
$$

Recalling that $p_{\rho,\xi}^{V_{\text{tot}}}(X) = \operatorname{tr}[\mathcal{J}_{0,x}(\rho \otimes P[\xi])]$, we may therefore write the average work as

$$
\begin{aligned}
\langle \mathcal{W}\rangle_{\rho,\xi}^{V_{\text{tot}}} &:= \sum_X p_{\rho,\xi}^{V_{\text{tot}}}(X)\mathcal{W}(X), \\
&= \sum_{x',x}\operatorname{tr}[H_{\text{tot}}\mathcal{J}_{x',x}(\rho \otimes P[\xi])] - \sum_{x',x}\operatorname{tr}[\mathcal{J}_{x',x}(H_{\text{tot}}\rho \otimes P[\xi])].
\end{aligned}
\tag{A6}
$$

Noting that $\sum_{x',x}\mathcal{J}_{x',x}$ is a trace-preserving operation, it follows that

$$
\sum_{x',x}\operatorname{tr}[\mathcal{J}_{x',x}(H_{\text{tot}}\rho \otimes P[\xi])] = \operatorname{tr}[H_{\text{tot}}(\rho \otimes P[\xi])].
\tag{A7}
$$

Similarly, noting that $\sum_{x'} Z_{x'} H_{\text{tot}} Z_{x'} = H_{\text{tot}}$, while $\sum_x Z_x P[\xi] Z_x = P[\xi]$, we have

$$
\sum_{x',x}\operatorname{tr}[H_{\text{tot}}\mathcal{J}_{x',x}(\rho \otimes P[\xi])] = \operatorname{tr}[H_{\text{tot}}V_{\text{tot}}(\rho \otimes P[\xi])V_{\text{tot}}^\dagger].
\tag{A8}
$$

Therefore, the average total work reads

$$
\langle \mathcal{W}\rangle_{\rho,\xi}^{V_{\text{tot}}} = \operatorname{tr}[(V_{\text{tot}}^\dagger H_{\text{tot}} V_{\text{tot}} - H_{\text{tot}})(\rho \otimes P[\xi])]
\tag{A9}
$$

for all $\rho \in \mathcal{S}(\mathcal{H})$ and V.

Appendix B: Proof of Eq. (38)

Let $H = \sum_m \epsilon_m P_m$ be the Hamiltonian of system $\mathcal{H}$, and $H_{\mathcal{A}} = \sum_m \lambda_m Z_m$ the Hamiltonian of system $\mathcal{H}_{\mathcal{A}}$, such that $H_{\text{tot}} = H + H_{\mathcal{A}}$ is the total, additive Hamiltonian of the compound system $\mathcal{H} \otimes \mathcal{H}_{\mathcal{A}}$. Moreover, let $|\xi\rangle \in \mathcal{H}_{\mathcal{A}}$ be a unit vector which is an eigenstate of $H_{\mathcal{A}}$ with eigenvalue λ_0. Finally, let U be a unitary operator on $\mathcal{H} \otimes \mathcal{H}_{\mathcal{A}}$ such that, for all $|\psi\rangle \in \mathcal{H}$,

$$
|U(\psi \otimes \xi)\rangle = \sum_m |P_m\psi \otimes \phi_m\rangle,
\tag{B1}
$$

where $|\psi \otimes \xi\rangle \equiv |\psi\rangle \otimes |\xi\rangle$, $|P_m\psi\rangle \equiv P_m|\psi\rangle$, and $|\phi_m\rangle$ are eigenstates of $H_{\mathcal{A}}$ with eigenvalue λ_m. It follows that for all $|\psi\rangle \in \mathcal{H}$, we have

$$
\begin{aligned}
\langle\psi \otimes \xi|H_{\text{tot}}|\psi \otimes \xi\rangle &= \langle\psi|H|\psi\rangle\langle\xi|\xi\rangle + \langle\psi|\psi\rangle\langle\xi|H_{\mathcal{A}}|\xi\rangle, \\
&= \langle\psi|H|\psi\rangle + \langle\psi|\psi\rangle\lambda_0, \\
&= \langle\psi|(H + \lambda_0\mathbb{1})|\psi\rangle.
\end{aligned}
\tag{B2}
$$

In the first line we use the additivity of H_{tot}, in the second line we use the fact that $\langle\xi|\xi\rangle = 1$ and $\langle\xi|H_A|\xi\rangle = \lambda_0$, and in the final line we use the fact that $\langle\psi|\psi\rangle\lambda_0 = \langle\psi|\lambda_0\mathbb{1}|\psi\rangle$. Similarly, for all $|\psi\rangle \in \mathcal{H}$ we have

$$
\begin{aligned}
\langle\psi\otimes\xi|U^\dagger H_{\text{tot}}U|\psi\otimes\xi\rangle &= \langle U(\psi\otimes\xi)|H_{\text{tot}}|U(\psi\otimes\xi)\rangle, \\
&= \sum_{m,n}\langle P_m\psi\otimes\phi_m|H_{\text{tot}}|P_n\psi\otimes\phi_n\rangle, \\
&= \sum_{m,n}\langle\psi|P_mHP_n|\psi\rangle\langle\phi_m|\phi_n\rangle + \langle\psi|P_mP_n|\psi\rangle\langle\phi_m|H_A|\phi_n\rangle, \\
&= \sum_m\langle\psi|P_mHP_m|\psi\rangle + \langle\psi|P_m|\psi\rangle\lambda_m, \\
&= \langle\psi|\big(H + \sum_m\lambda_mP_m\big)|\psi\rangle.
\end{aligned}
\tag{B3}
$$

In the first line we use the definition of the adjoint of a unitary operator U, in the second line we use Eq. (B1), in the third line we use the additivity of H_{tot} and the fact that the projection operators P_m are self-adjoint, in the fourth line we use $\langle\phi_m|\phi_n\rangle = \delta_{m,n}$ together with $P_mP_n = \delta_{m,n}P_m$ and $\langle\phi_m|H_A|\phi_m\rangle = \lambda_m$, and in the final line we use the fact that $\sum_m P_mHP_m = H$.

Combining Eq. (B2) with Eq. (B3) implies that, for all $|\psi\rangle \in \mathcal{H}$, we have

$$
\begin{aligned}
\langle\psi\otimes\xi|(U^\dagger H_{\text{tot}}U - H_{\text{tot}})|\psi\otimes\xi\rangle &= \langle\psi|\big(\sum_m\lambda_mP_m - \lambda_0\mathbb{1}\big)|\psi\rangle, \\
&= \langle\psi|\big(\sum_m(\lambda_m - \lambda_0)P_m\big)|\psi\rangle,
\end{aligned}
\tag{B4}
$$

where in the final line we use the fact that $\sum_m P_m = \mathbb{1}$. From the above equation, it follows that

$$
\Gamma_\xi(U^\dagger H_{\text{tot}}U - H_{\text{tot}}) = \sum_m(\lambda_m - \lambda_0)P_m,
\tag{B5}
$$

where $\Gamma_\xi : \mathcal{L}(\mathcal{H}\otimes\mathcal{H}_A) \to \mathcal{L}(\mathcal{H})$ is the restriction map for $|\xi\rangle \in \mathcal{H}_A$ defined as $\text{tr}[\Gamma_\xi(B)T] = \text{tr}[B(T\otimes P[\xi])]$ for all $B \in \mathcal{L}(\mathcal{H}\otimes\mathcal{H}_A)$ and $T \in \mathcal{T}(\mathcal{H})$.