

Heat and Work in Quantum Thermodynamics: a Cybernetic Approach

William Rupush* and Oscar Grånäs†

Department of Physics and Astronomy, Uppsala University

(Dated: March 28, 2024)

We present a new proposal for distinguishing heat from work based on a control-theoretic observability decomposition. We derive a Hermitian operator representing instantaneous dissipation of observable energy, and suggest a generalization of the von-Neumann entropy which can account for the model-uncertainty also present in pure states if the measured observables are informationally incomplete. In this view, the transition from a fundamental to a thermodynamic model consists in mapping the fundamental density matrix to an effective one, generally of lower dimension, encoding only what is observable given the constraints of our sensor and actuator capabilities. The generalized entropy captures the information loss incurred in this mapping. The theory is illustrated for the central spin model, where we show that the application of external controls can increase the size of thermal fluctuations and lower the entropy.

I. INTRODUCTION: HEAT AND WORK

The First Law of Thermodynamics, aside from being a statement of energy conservation, claims that there are two qualitatively different ways to alter the energy of a thermal system: by the application (or extraction) of *work*, or the addition (or subtraction) of *heat*. How to make this distinction precise is a major conceptual issue in thermodynamics, with fundamental ramifications. A satisfactory solution is still lacking, and this presents a roadblock towards the development of a mature science of quantum thermodynamics [1, 2].

A common approach is to proceed in a heuristic manner and make sure that the problem under consideration contains a subsystem which, whatever the exact definitions of heat and work, is one which would surely be recognized as either a heat or a work reservoir, or in other words, as a bath or a battery [3]. The bath is usually taken to be an infinite collection of quantum harmonic oscillators, and the battery as any system whose energy can be read off from measurements, and efficiently stored and extracted on demand. In this way the issue of precise definitions is sidelined for particular applications. Alternatively, in order to obtain a heat-work decomposition for interactions with a more general environment, one may follow R. Alicki [4] and consider the following application of the Leibniz product rule

$$\frac{d}{dt}\langle\hat{H}\rangle = \langle\frac{d}{dt}\hat{H},\hat{\rho}\rangle + \langle\hat{H},\frac{d}{dt}\hat{\rho}\rangle, \quad (1)$$

and decide that the first term corresponds to "work", and the second to "heat". This method is applied widely in the research literature to reduced density matrices in open systems [5, 6]. It can be shown that under certain constraints imposed on the Hamiltonian, only the second term contributes to an increase in the local von-Neumann

entropy. A slightly more general method which relaxes said constraints is obtained using a so called *local effective measurement basis* (LEMBAS) which is determined by the experimental set-up [7]. In either of the aforementioned views, only open quantum systems can thermalize, and any heat-work decomposition requires a partitioning of the total system into a "working medium" of interest, and an "external bath". But recent experiments [8–11] have indicated that closed systems thermalize as well, stimulating new interest in alternative methods, such as the eigenstate thermalization hypothesis [12]. In the view proposed in this paper, both closed and open systems can thermalize, and instead of proclaiming a system-bath split by fiat it is derived based on the measurement and control resources at hand.

The two central notions of control theory are *controllability* and *observability*. The first deals with the extent to which state transformations can be actuated by varying externally controllable parameters, and the second the extent to which the state of the system can be inferred from measurements. Control-theorists are interested in separating out degrees of freedom which are controllable and/or observable, from those which are not [13]. We claim that it is only when the system is not fully observable, that is, when at least a part of it is a "black box", that thermodynamics as typically conceived is applicable. "Entropy" appears when degrees of freedom are hidden from us. This is effectively the case in classical thermodynamics where we only have access to course-grained macroscopic observables, and fundamentally so for an event horizon surrounding a black hole, obscuring the interior for outside observers. In our view, the transition from a fundamental to a thermodynamical model amounts to a model-reduction procedure where the unobservable degrees of freedom have been discarded, here described as a mapping to an *effective density matrix*. Using this mapping we obtain an entropy function that is non-vanishing also for pure states, provided at least a part of the system is unobservable, and is in general time-varying for an isolated system.

We claim that the control-theoretic notion of observ-

* william.rupush@physics.uu.se

† oscar.granas@physics.uu.se

ability is needed for a proper understanding the heat-work decomposition: *heat is energy flow into unobservable degrees of freedom, and work into those which are observable*. Or in other words, *work is energy transfer into information-bearing degrees of freedom*. While the use of the observability decomposition to this end is novel the general idea is to some extent already familiar to physicists, who have recognized the connection between information and thermodynamics ever since Maxwell formulated his famous thought experiment with the demon [14], and even more so since Shannon introduced his information-theoretic entropy [15], and Landauer's discovery of the link between heat production and information erasure [16]. Our theory solidifies this connection and brings together the fields of quantum control, information and thermodynamics in a very explicit manner.

We stress that the concepts of heat, work and entropy presented here are not universal, but explicitly context-dependent and specific for any given quantum control system. The heat-work decomposition cannot be performed for a dynamical system without specifying our means for controlling and observing it. This philosophy claims that thermodynamics is as much a theory of engineering as of physics per se. This should not be surprising to the historian of science who is aware of the fact that thermodynamics was developed during the industrial revolution with the expressed purpose of optimizing the performance of heat engines.

This formalism by itself does not force the usual thermodynamical laws to hold; this would require assumptions about the actual physics of the system under study. There is no universal law stating that energy must tend to flow into unobservable degrees of freedom (i.e. for entropy to increase), but this becomes highly probable if we assume that the dimensionality of the observable degrees of freedom to be much smaller than that of the unobservable, as well as sufficient coupling between the two. Its value therefore lies in clarifying fundamental concepts, rather than replacing the other approaches. For example, statistical mechanics is still required to understand the route towards and conditions for equilibrium. Paradigmatic applications we have in mind include the calculation of the energy dissipated in quantum computation processes, and the construction of impurity models.

After giving an overview of the relevant concepts from the quantum control theory literature, where D. D'Alessandro's observability decomposition [17] is at center stage, we formulate our proposal explicitly and illustrate its application in the central spin model, which is experimentally relevant as it can be used to model decoherence in NV centers [18] and quantum dots [19, 20]. Along the way we also prove a proposition connecting the observability decomposition to a measure of Fisher information.

II. QUANTUM CONTROL THEORY

An up to date overview of the current state of quantum control theory can be found in [21]. Consider a quantum system in the density matrix formalism governed by the Liouville-von-Neumann equation

$$\frac{d}{dt}\hat{\rho} = -i[\hat{H}(u), \hat{\rho}], \quad (2)$$

where the Hamiltonian depends on a function u freely selected by the control engineer from a set of admissible controls \mathbb{U} , generally assumed to be piece-wise constant. The Hamiltonian is assumed to be of bilinear form

$$\hat{H}(t) = \hat{H}_0 + \sum_{j=1}^k u_j(t)\hat{H}_j, \quad (3)$$

where \hat{H}_0 is referred to as the *drift Hamiltonian*, and \hat{H}_i for $i \in \{1, \dots, k\}$ as *control Hamiltonians*. The measured output is the expectation value of a Hermitian operator \hat{S} , i.e. $y(t) = \langle \hat{S}, \hat{\rho} \rangle$. If \mathcal{H} is the Hilbert space of states under consideration, we will refer to the quadruple $\Sigma = (\mathcal{H}, \hat{H}(\cdot), \mathbb{U}, \hat{S})$ as a *quantum control system*. Let n denote $\dim\mathcal{H}$.

A. Controllability

The *operator controllability problem* consists of determining the *reachable set*

$$\mathcal{R} \triangleq \{\hat{U} \in U(n) \mid \hat{U} = \hat{U}_u(t) \text{ for some } t \geq 0, u \in \mathbb{U}\},$$

where \hat{U}_u satisfies the operator Schrödinger equation

$$i\frac{\partial}{\partial t}\hat{U}_u = \hat{H}(u)\hat{U}_u. \quad (4)$$

\mathcal{R} is the set of all unitary transformations which can be achieved by varying u over \mathbb{U} . If $\mathcal{R} = U(n)$ then any unitary matrix can be obtained by a suitable choice of controls u_1, \dots, u_k and time t , in which case the system is said to be *operator controllable*, the strongest of the controllability notions considered in the literature. [22]

A central theorem of quantum control theory states that the reachable set can be obtained as the exponential image of the *dynamical Lie algebra* [23], i.e.

$$\mathcal{R} = e^{\mathcal{L}}, \quad (5)$$

where \mathcal{L} is the Lie algebra generated by the set $\{-i\hat{H}_0, -i\hat{H}_1, \dots, -i\hat{H}_k\}$. It is the vector space of all real linear combinations of elements in this set, as well as of their nested commutators. It follows from the above that the system is operator controllable if $\mathcal{L} = \mathfrak{u}(n)$ or $\mathcal{L} = \mathfrak{su}(n)$. The difference between the two cases amounts to whether one can control the overall phase of the system, a difference generally considered to lack physical significance.

B. Observability

A quantum control system is said to be *observable* if its internal state can be determined from measurements of the chosen observable(s). Observability is defined formally with reference to the equivalence relation of *indistinguishability*. Let $\hat{\rho}(t, u, \hat{\rho}_0)$ denote the solution to Eq. 2 for the initial state $\hat{\rho}_0$ and applied control u , and assume that we are measuring an observable \hat{S} .

Definition 1 A pair of states $(\hat{\rho}_0, \hat{\rho}'_0)$ are *indistinguishable*, denoted by $\hat{\rho}_0 \sim \hat{\rho}'_0$, if for any control $u \in \mathbb{U}$ we have

$$\langle \hat{S}, \hat{\rho}(t, u, \hat{\rho}_0) \rangle = \langle \hat{S}, \hat{\rho}(t, u, \hat{\rho}'_0) \rangle, \quad \forall t \geq 0. \quad (6)$$

The quantum control system Σ is *observable in one step* if

$$\hat{\rho}_0 \sim \hat{\rho}'_0 \iff \hat{\rho}_0 = \hat{\rho}'_0. \quad (7)$$

This definition has been generalized to both multiple and non-projective measurements [24], but here we assume a single projective measurement. A necessary and sufficient condition for observability (in one step), due to D. D'Alessandro [17], makes use of an object called the *observability space*, defined

$$\mathcal{V} \triangleq \bigoplus_{j=0}^{\infty} \text{ad}_{\mathcal{L}}^j \text{span}\{i\hat{S}'\}, \quad (8)$$

where $\hat{S}' = \hat{S} - \frac{\text{Tr}(\hat{S})}{n} \hat{I}$. If the measured observable \hat{S} is an element of $i\mathcal{L}$, then \mathcal{V} is the ideal of \mathcal{L} generated by \hat{S} . A quantum control system is observable in one step if and only if

$$\mathcal{V} = \mathfrak{su}(n). \quad (9)$$

Note that if $\mathcal{L} = \mathfrak{su}(n)$, then since $\mathfrak{su}(n)$ is simple (i.e. contains no non-trivial ideals), Σ is observable for any \hat{S} not proportional to the identity operator. If the system is not observable, then there exists a decomposition $\hat{\rho} = \hat{\rho}_o + \hat{\rho}_u$, where $\hat{\rho}_o \in i\mathcal{V}$ and $\hat{\rho}_u \in i\mathcal{V}^\perp$, such that the measured output depends only on $\hat{\rho}_o$:

$$y(t) = \frac{1}{n} \text{Tr}(\hat{S}) + \langle \hat{S}, \hat{\rho}_o \rangle. \quad (10)$$

The orthogonal complement \mathcal{V}^\perp is taken inside of $\mathfrak{su}(n)$ with respect to the Hilbert-Schmidt inner product. If $\hat{S} \in i\mathcal{L}$ then the reachable set splits into a semi-direct product of Lie groups

$$e^{\mathcal{L}} \simeq e^{\mathcal{V}} \rtimes e^{\mathcal{L}/\mathcal{V}}. \quad (11)$$

This yields a dynamical decomposition where the action of quotient group leaves the observed output unaffected. The condition $\hat{S} \in i\mathcal{L}$ is obtained in all cases where one is measuring a local observable $\hat{S} \otimes \hat{I}$, on a locally controllable system, meaning a system where $\mathfrak{su}(\tilde{N}) \otimes \hat{I} \subset \mathcal{L}$

and $\tilde{N} = \dim(\hat{S})$.

To obtain the observability decomposition we project the density matrix onto each of the basis elements of $i\mathcal{V}$. Provided the basis is orthogonal, we have

$$\hat{\rho}_o(t) = - \sum_{i=1}^{\dim \mathcal{V}} \langle \hat{B}_i, \hat{\rho}(t) \rangle \frac{\hat{B}_i}{\|\hat{B}_i\|^2}. \quad (12)$$

\mathcal{V} is not itself necessarily a Lie algebra, but when it is, which includes the important case when $\hat{S} \in i\mathcal{L}$, work can be distinguished into qualitatively different kinds using any of the known Lie algebra decompositions. Notably, using the Levi decomposition [21] we can split \mathcal{V} into a direct sum $\mathcal{Z} \otimes \mathcal{S}$, where \mathcal{Z} is abelian and \mathcal{S} semi-simple. \mathcal{V} then assumes the form of a set of parallel subsystems that are mutually commuting and orthogonal, plus a classical subsystem \mathcal{Z} consisting of only commuting observables. In this way we can differentiate "classical" from "quantum" work depending on whether the energy goes into the classical system \mathcal{Z} , or any of the simple quantum systems in \mathcal{S} .

Remark 1 The key property which characterizes isolated thermal states is that observables have expectation values proportional to the trace of the corresponding Hermitian operator, i.e. $y(t) \propto \text{Tr}(\hat{S})$. This is referred to as the *microcanonical ensemble*. Note that this condition, for a specified observable \hat{S} , is equivalent to the condition that $\hat{\rho} \in i\mathcal{V}^\perp$ by Eq. 10. This implies that the maximally thermalized state is the maximally unobservable one, and therefore reveals a new framework for thinking about thermalization.

C. Observability and Fisher Information

The *Fisher information* is a key concept in both classical and quantum estimation theory which represents the extent to which measurements on random variables provides information on some unknown parameter of interest [25, 26]. The Fisher information can be defined in a variety of ways, but here we'll focus on a particular kind defined using *symmetric logarithmic derivative (SLD) operators* [27].

Consider a density matrix $\hat{\rho}(\theta)$ dependent on r parameters $(\theta_1, \dots, \theta_r)$. The matrix components of the Fisher information are then defined by the equation

$$F_{ij} \triangleq \frac{1}{2} \langle \hat{\rho}, [\hat{L}_{\theta_i}, \hat{L}_{\theta_j}]_+ \rangle, \quad (13)$$

where the SLD operators \hat{L}_i , for $i \in \{1, \dots, r\}$ are defined by the equations

$$\frac{\partial}{\partial \theta_i} \hat{\rho}(\theta) = \frac{1}{2} [\hat{L}_{\theta_i}, \hat{\rho}]_+. \quad (14)$$

A key property of the Fisher information is that it provides a bound on covariances of unbiased estimators $\hat{\theta}$ of

θ , a result called the Cramer-Rao inequality:

$$\text{Cov}(\tilde{\theta}) \geq F^{-1}. \quad (15)$$

The Fisher information is typically a complicated time-dependent function, but when the parameter θ is a component of the state along a basis vector \hat{B}_j of $i\mathcal{V}$ it assumes a particularly simple form. \hat{B}_j is then an unbiased estimator of θ_j for every $j \in \{1, \dots, \dim\mathcal{V}\}$.

Proposition 1 *Let Σ be a quantum control system with $\{\hat{B}_i\}$ the basis of its observability space, and let the estimated parameters be the projections $\theta_i = \langle \hat{\rho}, \hat{B}_i \rangle$. If the quantum state is pure, then the SLD Fisher information is the metric tensor of the observability space induced by the Hilbert-Schmidt inner-product:*

$$F_{ij} = \langle \hat{B}_i, \hat{B}_j \rangle. \quad (16)$$

A proof is found in the appendix.

III. QUANTUM DISSIPATION CHANNELS

We define a *quantum dissipation channel* to be any component $i\hat{H}_u \in \mathcal{V}^\perp$, of an element $i\hat{H} \in \mathcal{L}$, such that $[\hat{H}, \hat{H}_u] \neq 0$. In the particular case where $i\hat{S} \in \mathcal{L}$, when \mathcal{V} is an ideal of \mathcal{L} , we have the short exact sequence

$$\mathcal{V} \xhookrightarrow{i} \mathcal{L} \xrightarrow{\pi} \mathcal{L}/\mathcal{V}, \quad (17)$$

where i is the inclusion map and π the canonical projection. Then a set of dissipation channels can be viewed as a non-abelian Lie algebra extension of \mathcal{V} in \mathcal{L} . In this case a quantum system has a dissipation channel if the observability space is embedded in a strictly larger dynamical Lie algebra such that \mathcal{L}/\mathcal{V} does not commute with \mathcal{V} . This means that energy can be transferred to and from unobservable degrees of freedom. Once \mathcal{V} has been calculated one can straightforwardly identify the dissipation channel as that part of the system Hamiltonian which lies in $i\mathcal{V}^\perp$ and does not commute with the whole.

Performing the observability decomposition on the Hamiltonian gives us $\hat{H} = \hat{H}_u + \hat{H}_o$, where $\hat{H}_u \in i\mathcal{V}^\perp$ and $\hat{H}_o \in i\mathcal{V}$. The total energy is then decomposed as

$$E = \mathcal{O} + \mathcal{U}, \quad \mathcal{O} \triangleq \langle \hat{H}_o, \hat{\rho} \rangle, \quad \mathcal{U} \triangleq \langle \hat{H}_u, \hat{\rho} \rangle. \quad (18)$$

The main proposal of this paper is that differential heat and work flows can be identified as

$$\dot{Q} \triangleq \frac{d\mathcal{U}}{dt}, \quad \dot{W} \triangleq \frac{d\mathcal{O}}{dt}, \quad (19)$$

and that the component $\hat{H}_u \in i\mathcal{V}^\perp$ can be interpreted as a dissipation channel for the quantum control system under consideration provided that $[\hat{H}_u, \hat{H}_o] \neq 0$.

Consider for a moment the heat flow when the Hamiltonian is time-independent,

$$\dot{Q} = \langle \hat{H}_u, \frac{d}{dt}\hat{\rho} \rangle = -i\langle \hat{H}_u, [\hat{H}, \hat{\rho}] \rangle. \quad (20)$$

Employing a trace identity we obtain

$$-i\langle [\hat{H}_u, \hat{H}], \hat{\rho} \rangle = -i\langle [\hat{H}_u, \hat{H}_o], \hat{\rho} \rangle. \quad (21)$$

If we define the operator

$$\hat{\mathcal{D}} \triangleq i[\hat{H}_o, \hat{H}_u], \quad (22)$$

then $\langle \hat{\mathcal{D}}, \hat{\rho} \rangle = \dot{Q}$. In other words, the expectation value of $\hat{\mathcal{D}}$ is equal to the instantaneous transfer of energy from the observable to the unobservable component. This justifies interpreting it as a dissipation operator.

\dot{Q} can also be calculated from quantities of \mathcal{V} , in case we don't know $\hat{\mathcal{D}}$. For a pure state we have

$$\dot{Q} = - \sum_{i=1}^{\dim\mathcal{V}} \sum_{j=1}^{\dim\mathcal{V}} (\theta_i \dot{h}_j + \theta_j \dot{h}_i) F_{ij}, \quad (23)$$

where we've introduced the notation

$$h_j(t) \triangleq \langle \hat{H}(t), \hat{B}_j \rangle, \quad (24)$$

and the θ_i s and F are defined as in Proposition 1. If we define the matrix $\mathbb{H}_{ij} \triangleq \theta_i \dot{h}_j$ we obtain

$$\dot{Q} = -\langle \dot{\mathbb{H}}, F \rangle. \quad (25)$$

The right-hand side contains fluxes of observable quantities as well as a Fisher information, and the left is a dissipation measure. This opens up the possibility of using the Cramer-Rao inequality to obtain results similar to the Thermodynamic Uncertainty Relations discussed in the literature [28].

Remark 2 *The attentive reader will note that heat and work are traditionally viewed as path functions, while here they are functions of state. We believe this depends on whether the unobservable state component is explicitly modelled. If it is modelled, then the unobservable energy can be calculated at any initial and final times t_f and t_i to obtain the heat flow $Q = \mathcal{U}(t_f) - \mathcal{U}(t_i)$. However, if we are only modelling the observable part, then the transfer of energy between $i\mathcal{V}$ and $i\mathcal{V}^\perp$ have to be accounted for at each time step, making the heat flow an integral over the dissipation $Q = \int_{t_i}^{t_f} \langle \hat{\mathcal{D}} \rangle dt$.*

A. A Generalized Entropy

Having defined a heat flow as a transfer of energy into $i\mathcal{V}^\perp$, we note that all such transfers are not qualitatively equal. Roughly speaking, we want to differentiate between heat flows associated with the state becoming "more unobservable", and those associated with a

change in the occupied unobservable energy levels. This is essentially the difference between increasing the entropy and temperature respectively. We propose using the von-Neumann entropy S_{vN} under a substitution to an output-equivalent state depending only on $\hat{\rho}_o$. This "effective" state is $\hat{\rho}_o + \frac{1}{n}\hat{I}$, and is the unique density matrix depending only on $\hat{\rho}_o$ which is indistinguishable from $\hat{\rho}$ in the sense of Definition 1. Explicitly:

$$\mathcal{S}(\hat{\rho}) \triangleq -\langle \hat{\rho}_o + \frac{1}{n}\hat{I}, \log(\hat{\rho}_o + \frac{1}{n}\hat{I}) \rangle. \quad (26)$$

It reduces to the ordinary von-Neumann entropy when $\mathcal{V} = su(n)$, and when $\mathcal{V} \neq su(n)$ it converges to the maximum value of $\log n$ as $\hat{\rho}_o \rightarrow 0$. The difference $\mathcal{S}(\hat{\rho}) - S_{vN}(\hat{\rho})$ quantifies the information loss incurred in mapping to the effective state.

Define two operations \mathfrak{D} and \mathfrak{T}_t by their actions:

$$\mathfrak{D} : \hat{\rho} \mapsto \frac{1}{n}\hat{I} - \sum_{i=1}^{\dim \mathcal{V}} \langle \hat{\rho}, \hat{B}_i \rangle \frac{\hat{B}_i}{\|\hat{B}_i\|^2}, \quad (27)$$

$$\mathfrak{T}_t : \hat{\rho} \mapsto \hat{U}_t \hat{\rho} \hat{U}_t^\dagger. \quad (28)$$

They represent a projection onto the effective state (it is easy to verify that $\mathfrak{D}^2 = \mathfrak{D}$), and unitary time-evolution respectively. In general

$$\mathfrak{T}_t(\mathfrak{D}(\hat{\rho})) \neq \mathfrak{D}(\mathfrak{T}_t(\hat{\rho})). \quad (29)$$

Since $\mathcal{S} = S_{vN}(\mathcal{O}(\hat{\rho}))$, and noting that the von-Neumann entropy is unitary invariant, it follows that if the commutator is zero then the time derivative of the entropy is zero as well. To see when this holds, set the commutator to zero and rearrange using the cyclic property of

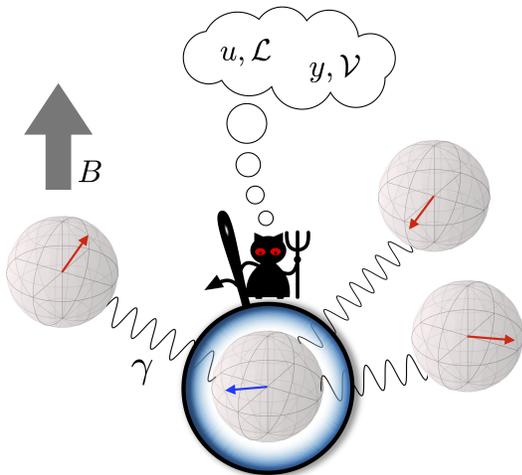


FIG. 1. An imperfect demon pondering his means for controlling and observing a central spin system. As full knowledge of the state is lacking, energy can leak out into unobservable degrees of freedom and thereby become "degraded", or "entropic".

the trace to obtain

$$\sum_{i=1}^{\dim \mathcal{V}} \langle \hat{\rho}, \hat{B}_i \rangle \frac{\hat{U}_t \hat{B}_i \hat{U}_t^\dagger}{\|\hat{B}_i\|^2} = \sum_{i=1}^{\dim \mathcal{V}} \langle \hat{\rho}, \hat{U}_t \hat{B}_i \hat{U}_t^\dagger \rangle \frac{\hat{B}_i}{\|\hat{B}_i\|^2}. \quad (30)$$

The above equation holds if it does so elementwise, i.e. if for every $i = 1, \dots, \dim \mathcal{V}$ we have

$$\langle \hat{\rho}, \hat{B}_i \rangle \mathfrak{T}_t(\hat{B}_i) = \langle \hat{\rho}, \mathfrak{T}_t(\hat{B}_i) \rangle \hat{B}_i. \quad (31)$$

This is true if $[\hat{H}, \hat{B}_i] = 0$ for all $\hat{H} \in i\mathcal{L}$ (which can be seen by taking the limit $t \rightarrow 0$), or if \hat{B}_i is a simultaneous eigenvector for \mathcal{L} . If $\mathcal{V} \subset \mathcal{L}$, then the existence of such an eigenvector is guaranteed if \mathcal{L} is solvable, by Lie's theorem. So a sufficient condition for the commutativity of \mathfrak{D} and \mathfrak{T}_t is that every basis vector of \mathcal{V} either commutes with the whole of \mathcal{L} or is a simultaneous eigenvector.

If we take the time-derivative of the entropy, assuming unitary dynamics, we get

$$\dot{\mathcal{S}} = -\langle \dot{\hat{\rho}}_o, \log(\hat{\rho}_o + \frac{1}{n}\hat{I}) \rangle = \quad (32)$$

$$\left\langle \sum_{j=1}^{\dim \mathcal{V}} \langle \hat{\rho}, [\hat{H}, \hat{B}_j] \rangle \frac{i\hat{B}_j}{\|\hat{B}_j\|^2}, \log(\hat{\rho}_o + \frac{1}{n}\hat{I}) \right\rangle. \quad (33)$$

We again see that $\dot{\mathcal{S}} \neq 0$ requires that at least some of the basis vectors of \mathcal{V} are time-varying (i.e. lie outside the center of \mathcal{L}). We can unpack this equation further if we assume that \mathcal{V} is an abelian, in which case $[\hat{H}, \hat{B}_j] = [\hat{H}_u, \hat{B}_j]$ for every j . Split the dissipation operator into a sum of dissipation operators, one for each basis vector of \mathcal{V} :

$$\hat{\mathcal{D}} = \sum_{j=1}^{\dim \mathcal{V}} \hat{\mathcal{D}}_j, \quad \hat{\mathcal{D}}_j \triangleq -ih_j \frac{[\hat{H}_u, \hat{B}_j]}{\|\hat{B}_j\|^2}, \quad (34)$$

where h_j is defined as in Equation 24. Then we can write the entropy change as

$$\dot{\mathcal{S}} = -\sum_{j=1}^{\dim \mathcal{V}} \frac{\langle \hat{\mathcal{D}}_j \rangle}{h_j} \left\langle \hat{B}_j, \log(\hat{\rho}_o + \frac{1}{n}\hat{I}) \right\rangle, \quad (35)$$

which clearly illuminates its relation to dissipation, in a manner reminiscent of Clausius' theorem.

IV. THE CENTRAL SPIN MODEL

The central spin model consists of a distinguished spin 1/2 particle, called the *central spin*, connected to a bath of N mutually non-interacting spin 1/2 particles. The interaction is of Heisenberg XXX type with a ferromagnetic ground state. We have a product Hilbert space $\mathcal{H} = \mathcal{H}_2 \otimes \mathcal{H}_2 \otimes \dots \otimes \mathcal{H}_2$ of $N+1$ factors, where the first factor represents the central spin, and $\dim \mathcal{H} = 2^{N+1}$. The whole system is immersed in a constant magnetic field B

aligned to the z -axis and interacting equally with the central and bath spins, and the system is controlled by a tunable magnetic field along the y -axis acting on the central spin only. The Hamiltonian is a sum $\hat{H}(t) = \hat{H}_0 + \hat{H}_c(t)$ of drift and control terms

$$\hat{H}_0 = -B(\hat{\sigma}_z \otimes \hat{I}_b + \hat{I}_2 \otimes \sum_j \hat{\sigma}_z^{(j)}) + \quad (36)$$

$$\sum_{j=1}^N \gamma_j (\hat{\sigma}_x \otimes \hat{\sigma}_x^{(j)} + \hat{\sigma}_y \otimes \hat{\sigma}_y^{(j)} + \hat{\sigma}_z \otimes \hat{\sigma}_z^{(j)}), \quad (37)$$

$$\hat{H}_c(t) = -u(t)\hat{\sigma}_y \otimes \hat{I}_b, \quad (38)$$

where $\gamma_j < 0$ for all j , and the spin and identity operators on the bath are

$$\hat{\sigma}_i^{(j)} \triangleq \underbrace{\hat{I}_2 \otimes \dots \otimes \hat{I}_2}_{j-1 \text{ of them}} \otimes \hat{\sigma}_i \otimes \underbrace{\hat{I}_2 \otimes \dots \otimes \hat{I}_2}_{N-j \text{ of them}}, \quad (39)$$

$$\hat{I}_b = \underbrace{\hat{I}_2 \otimes \dots \otimes \hat{I}_2}_{N \text{ of them}} \quad (40)$$

We are measuring the magnetization of the central spin in the x -direction:

$$\hat{S} = \hat{\sigma}_x \otimes \hat{I}_b. \quad (41)$$

Controllability, and therefore also observability, is strongly dependent on the relative values of the coupling parameters γ_j . If all of them are distinct the full system is both operator controllable and observable. We consider the opposite case where all of them are equal ($\gamma_j = \gamma_k \forall j, k$) and the entire bath is behaving like a single collective spin. We refer the reader to [29] for an extensive analysis of the controllability of the central spin model, and here just note that for equal coupling the system is uncontrollable and the dynamical Lie algebra satisfies

$$\dim(\mathcal{L}) = \begin{cases} \frac{1}{6}(2+N)(9+4N(4+N)) & \text{for } N \text{ even} \\ \frac{1}{6}(1+N)(3+2N)(7+2N) & \text{for } N \text{ odd} \end{cases}$$

The central spin, however, is locally controllable, which together with \hat{S} being a local observable, implies that $\mathcal{V} \subset \mathcal{L}$ is an ideal.

A. Numerical analysis

We let the number of bath spins (N) be equal to 3, giving us a Hilbert space dimension of $2^4 = 16$. We consider an initial state with all spins perfectly aligned to the constant magnetic field along the z direction. This is a pure state, and it will remain pure throughout the simulation. A control field, here a variable magnetic field along the y direction applied to the central spin, is obtained using the pulse-optimizer function available in Qutip [30, 31] to drive the central spin towards alignment with the x axis at terminal time τ . In other words, this is an optimal control problem of Mayer type with target functional

$$J[u] = \langle \hat{S}, \hat{\rho}(\tau) \rangle. \quad (42)$$

Afterwards the system evolves freely under a Gaussian control field for another time period of τ . During both processes we compute the observability decomposition of the energy (\mathcal{U}, \mathcal{O}), the instantaneous dissipation (\hat{D}), the entropy \mathcal{S} and the measured output (\hat{S}).

We are working in Hartree atomic units, and moreover, calculate the energy in units of $\frac{\text{Ha}}{\gamma}$, giving time units of $\frac{\hbar\gamma}{\text{Ha}}$. In these units, we set the magnetic field B equal to 10 and the interaction constant γ equal to -3 . The terminal time τ is set to unity, and the interval is discretized into 1000 timeslots for the optimization procedure and time-evolution. During the free evolution phase the control field is set equal to a Gaussian centered at zero with a standard deviation of 0.1.

The dynamical Lie algebra and the observability space are calculated numerically using the algorithm presented in Appendix B. We find that $\dim(\mathcal{L}) = 78$, in agreement with the results in [29], and has a maximum depth of 11. For the observability space we find $\dim(\mathcal{V}) = 41$ with a maximal depth of 1. The basis of the observability space is orthogonalized using the Gram-Schmidt procedure.

Figure 2 shows the evolution of the controlled system in the interval $[0, \tau]$. The heat, work and entropy changes are:

$$Q_c = 0.00, \quad W_c = 19.00, \quad \Delta\mathcal{S}_c = -0.29. \quad (43)$$

During the controlled phase the entropy begins at its maximal value, which is to be expected since the system is initialized in a state with zero output. It oscillates with the evolution of the system, attaining lower values the larger the absolute value of the output. Initially there is no dissipation of energy, but $\langle \hat{D} \rangle$ increases with the amplitude of oscillation of the observable output as the system is controlled towards the desired state. Overall, \mathcal{O} increases substantially during, while \mathcal{U} is almost constant.

Figure 3 shows the subsequent free evolution in the interval $[\tau, 2\tau]$, during which energy flows sinusoidally between the two subspaces. These oscillations may be interpreted as thermal fluctuations. This behavior is expected as isolated quantum systems undergo periodic motion. It also shows the free evolution of the state which is obtained at $t = 0.6$ during the controlled phase. The purpose is to illustrate that when we control the system to reduce the entropy, the more we reduce the entropy the greater the amplitude of the oscillations afterwards. The amplitudes of oscillations of the other quantities are similarly affected. In conclusion, with external control we can increase the size of thermal fluctuations.

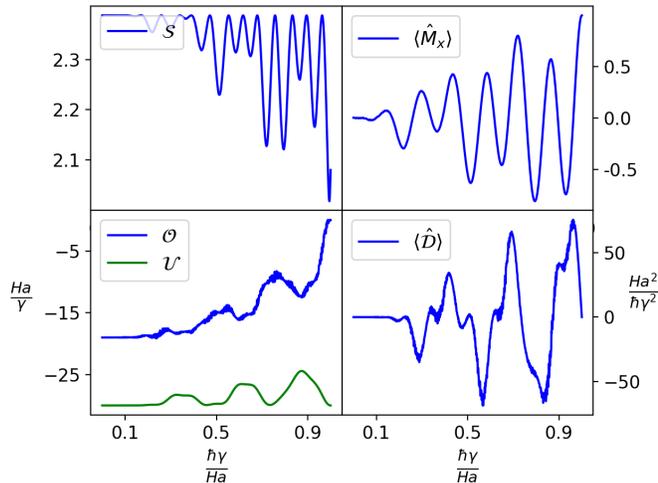


FIG. 2. Evolution of the central spin model during the controlled phase. Top left: generalized entropy \mathcal{S} . Top right: magnetization $\langle \hat{M}_x \rangle$ along the x -axis for the central spin. Bottom left: observable and unobservable energies \mathcal{O} and \mathcal{U} . Bottom right: instantaneous dissipation $\langle \hat{D} \rangle$.

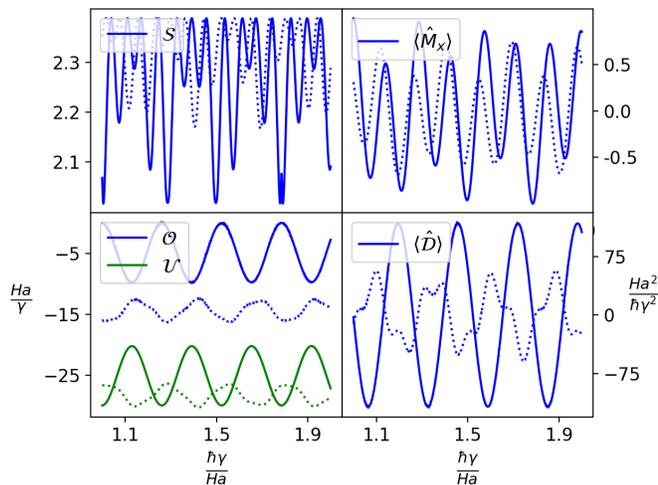


FIG. 3. Evolution of the central spin model during the free evolution phase. The quantities shown are the same as in Figure 2. For comparison, the dotted lines show a system that evolves freely from the state obtained at $t = 0.6$ during the controlled phase.

V. CONCLUSION

We have proposed a mathematically explicit method for distinguishing heat from work dependent on the sensor and control capabilities at hand, based on a control-theoretic observability decomposition due to D. D'Alessandro. We also derived a dissipation operator, and proposed an accompanying entropy function which for fully controllable and observable systems reduces to the ordinary von-Neumann entropy, but which can also

yield non-zero values for pure states if the observables under consideration are informationally incomplete (i.e. if the state estimation problem is not solvable). And perhaps more importantly, this entropy can change also for an isolated system.

Thermodynamics is fundamentally the science of the difference between heat and work, and the possibilities of their interconversion. This is the key distinction that the whole edifice is built around, so a thorough understanding of it is of paramount importance for the development of the entire field. A proper understanding of the heat-work decomposition will allow us to extend thermodynamics further into the quantum realm, improve our understanding of dissipation and thermalization, and shed light on some of the largest conceptual problems in physics.

The nature of the heat-work decomposition is a major open problem in quantum thermodynamics whose solution would have fundamental ramifications for both pure and applied science. We believe that our proposal provides the correct answer to the conundrum.

ACKNOWLEDGMENTS

We acknowledge funding from the Swedish Research Council (VR) grant 2019-03901, and the European Research Council (ERC) through the Synergy Grant 854843 - FASTCORR. We thank Yishao Zhou (Stockholm University) for being an excellent teacher and for encouraging the lead author to study quantum control and pursue this project. Erik Sjöqvist is acknowledged for fruitful discussions.

Appendix A

1. Notation & conventions

- The inner-product is always of Hilbert-Schmidt type: $\langle A, B \rangle = \text{Tr}(A^\dagger B)$.
- We work in Hartree atomic units where $\hbar = e = m_e = k_e = 1$.
- The anti-commutator is denoted $[A, B]_+ = AB + BA$.
- The Lie algebra $\mathfrak{u}(n)$ is the set of skew-Hermitian matrices of dimension n , and $\mathfrak{su}(n)$ is the subset of those which are traceless. The corresponding Lie groups are $U(n)$, the set of unitary matrices of the same dimension $U(n)$, and the subset of those with determinant unity, $SU(n)$.
- \hat{I}_2 is the two-dimensional identity matrix.
- $\text{ad}_A(B)$ is the set of all commutators $[a, b]$ where $a \in A$ and $b \in B$. If there is a superscript ad^i it

represents a set formed by taking i repeated commutators of elements in B with elements in A . For $i = 2$ the elements are of the form $[a_2, [a_1, b]]$.

2. Algorithm for computing \mathcal{L} and \mathcal{V}

For a bilinear Hamiltonian of the form

$$\hat{H}(t) = \hat{H}_0 + \sum_{j=1}^k u_j(t) \hat{H}_{c,j}, \quad (\text{A1})$$

the dynamical Lie algebra \mathcal{L} is defined as the Lie algebra generated by the set $\{i\hat{H}_1, i\hat{H}_{c,1}, \dots, i\hat{H}_{c,k}\}$. We denote the generating set by G . Denote the set of basis elements of our would be algebra $\mathcal{B}_{\mathcal{L}}$. The first step in the algorithm is to

- **Step 0:** Add all linearly independent elements of G to $\mathcal{B}_{\mathcal{L}}$. These are the elements \mathcal{L} of depth 0.
- **Step k :** Compute the set of all commutators of $B_{\mathcal{L}}$ with G . Each result which is linearly independent of the rest of $\mathcal{B}_{\mathcal{L}}$, add it to $\mathcal{B}_{\mathcal{L}}$. These are the basis elements of depth k . If no new linearly independent elements are obtained, the algorithm terminates and $\mathcal{B}_{\mathcal{L}}$ is a basis of \mathcal{L} . Otherwise, proceed to step $k + 1$.

The algorithm to compute a basis $\mathcal{B}_{\mathcal{V}}$ of \mathcal{V} is identical to the above, except that the generating set is \mathcal{L} and the depth 0 element(s) the measured observable(s). In other words, add your measured observable to $\mathcal{B}_{\mathcal{V}}$. Then compute all commutators of depth 1 of \mathcal{L} with $\mathcal{B}_{\mathcal{V}}$, and add the linearly independent results to $\mathcal{B}_{\mathcal{V}}$. Then just repeat the procedure until no new linearly independent elements are generated.

3. Proof of Proposition 1

Assume that $\dim(\mathcal{V}) = r$ and let ρ be written in the form

$$\hat{\rho}(\theta) = \hat{\rho}_u + \sum_{j=1}^r \theta_j \hat{B}_j,$$

where we've introduced the notation \hat{B}_j for the basis elements of $i\mathcal{V}$, and $\theta_j \triangleq \text{Tr}\{\hat{B}_j \hat{\rho}\}$. The r -parameter estimation problem is now formulated as a problem of finding the projections of $\hat{\rho}$ along every basis operator of $i\mathcal{V}$. Taking the partial derivative of $\hat{\rho}$ with respect to one of the parameters we obtain

$$\frac{\partial}{\partial \theta_j} \hat{\rho}(\theta) = \hat{B}_j = \frac{1}{2} [\hat{L}_{\theta_j}, \hat{\rho}]_+ = \frac{1}{2} (\hat{L}_{\theta_j} \hat{\rho} + \hat{\rho} \hat{L}_{\theta_j}).$$

Since the basis operators \hat{B}_j are traceless, we see that the expectation values of the SLD operators vanish, i.e.

$$\langle \hat{L}_j \rangle = \text{Tr}\{\hat{L}_{\theta_j} \hat{\rho}\} = \text{Tr}\{\hat{B}_j\} = 0.$$

By employing the cyclic property of the trace, a straightforward calculation shows that

$$\text{Tr}\{[\hat{B}_i, \hat{B}_j]_+\} = \text{Tr}\{2\hat{L}_{\theta_i} \hat{\rho} \hat{L}_{\theta_j} \hat{\rho} + \hat{\rho}^2 \hat{L}_{\theta_i} \hat{L}_{\theta_j} + \hat{\rho}^2 \hat{L}_{\theta_j} \hat{L}_{\theta_i}\}.$$

If the quantum state is pure, then the density matrix is a projector with $\hat{\rho}^2 = \hat{\rho} = |\psi\rangle\langle\psi|$, and

$$\begin{aligned} \text{Tr}\{2\hat{L}_{\theta_i} \hat{\rho} \hat{L}_{\theta_j} \hat{\rho}\} &= 2\text{Tr}\{\hat{L}_{\theta_i} |\psi\rangle\langle\psi| \underbrace{|\psi\rangle\langle\psi| \hat{L}_{\theta_j} |\psi\rangle\langle\psi|}_{\text{constant}}\} = \\ &= 2\langle\psi|\hat{L}_{\theta_j}|\psi\rangle \text{Tr}\{\hat{L}_{\theta_i} |\psi\rangle\langle\psi|\} = 2\langle\hat{L}_{\theta_j}\rangle \langle\hat{L}_{\theta_i}\rangle = 0. \end{aligned}$$

Writing out the Fisher information explicitly

$$F_{i,j} = \frac{1}{2} \text{Tr}\{\hat{\rho} \hat{L}_{\theta_i} \hat{L}_{\theta_j} + \hat{\rho} \hat{L}_{\theta_j} \hat{L}_{\theta_i}\},$$

we now see that under the assumption of purity,

$$F_{\theta_i, \theta_j} = \frac{1}{2} \text{Tr}\{[\hat{B}_i, \hat{B}_j]_+\}.$$

By the cyclic property of the trace, and the hermiticity of the matrices $\{\hat{B}_i\}$, this is equal to $\langle \hat{B}_i, \hat{B}_j \rangle$. \square

-
- [1] R. Alicki and R. Kosloff, *Introduction to Quantum Thermodynamics: History and Prospects*, in *Thermodynamics in the Quantum Regime : Fundamental Aspects and New Directions*, edited by F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso (Springer, Cham, 2018) pp. 1–33.
- [2] F. Binder, L. A. Correa, C. Gogolin, J. Anders, and G. Adesso, eds., *Thermodynamics in the Quantum Regime*, Vol. 195 (Springer International Publishing, Cham, 2018).
- [3] G. Mahler, *Quantum Thermodynamic Processes* (Jenny Stanford Publishing, New York, 2014).
- [4] R. Alicki, *The quantum open system as a model of the heat engine*, *Journal of Physics A: Mathematical and General* **12**, L103 (1979).
- [5] B. Ahmadi, S. Salimi, and A. S. Khorashad, *On the contribution of work or heat in exchanged energy via interaction in open bipartite quantum systems*, *Scientific Reports*

- 13**, 160 (2023).
- [6] A. Vallejo, A. Romanelli, and R. Donangelo, *Qubit thermodynamics far from equilibrium: Two perspectives about the nature of heat and work in the quantum regime*, *Physical Review E* **103**, 042105 (2021).
- [7] H. Hossein-Nejad, E. J. O'Reilly, and A. Olaya-Castro, *Work, heat and entropy production in bipartite quantum systems*, *New Journal of Physics* **17**, 075014 (2015).
- [8] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, *Quantum thermalization through entanglement in an isolated many-body system*, *Science* **353**, 794 (2016).
- [9] T. Langen, S. Erne, R. Geiger, B. Rauer, T. Schweigler, M. Kuhnert, W. Rohringer, I. E. Mazets, T. Gasenzer, and J. Schmiedmayer, *Experimental observation of a generalized Gibbs ensemble*, *Science* **348**, 207 (2015).
- [10] M. Gring, M. Kuhnert, T. Langen, T. Kitagawa, B. Rauer, M. Schreitl, I. Mazets, D. A. Smith, E. Demler, and J. Schmiedmayer, *Relaxation and Prethermalization in an Isolated Quantum System*, *Science* **337**, 1318 (2012).
- [11] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, *Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas*, *Nature Physics* **8**, 325 (2012).
- [12] J. M. Deutsch, *Eigenstate thermalization hypothesis*, *Reports on Progress in Physics* **81**, 082001 (2018).
- [13] E. D. Sontag, *Mathematical Control Theory*, Vol. 6 (Springer New York, New York, NY, 1998).
- [14] J. C. Maxwell, *Theory of Heat* (Cambridge University Press, Cambridge, 2011).
- [15] C. E. Shannon, *A Mathematical Theory of Communication*, *Bell System Technical Journal* **27**, 379 (1948).
- [16] R. Landauer, *The physical nature of information*, *Physics Letters A* **217**, 188 (1996).
- [17] D. D. Alessandro, *On quantum state observability and measurement*, *Journal of Physics A: Mathematical and General* **36**, 9721 (2003).
- [18] H. Park, J. Lee, S. Han, S. Oh, and H. Seo, *Decoherence of nitrogen-vacancy spin ensembles in a nitrogen electron-nuclear spin bath in diamond*, *npj Quantum Information* **8**, 95 (2022).
- [19] A. Bechtold, D. Rauch, F. Li, T. Simmet, P.-L. Audebert, A. Regler, K. Müller, N. A. Sinitsyn, and J. J. Finley, *Three-stage decoherence dynamics of an electron spin qubit in an optically active quantum dot*, *Nature Physics* **11**, 1005 (2015).
- [20] M. Arfaoui and S. Jaziri, *Decoherence of a spin-valley qubit in a MoS2 quantum dot*, *Journal of Physics Communications* **6**, 115004 (2022).
- [21] D. D'Alessandro, *Introduction to Quantum Control and Dynamics* (Chapman and Hall/CRC, Boca Raton, 2021).
- [22] F. Albertini and D. D'Alessandro, *Notions of controllability for bilinear multilevel quantum systems*, *IEEE Transactions on Automatic Control* **48**, 1399 (2003).
- [23] G. M. Huang, T. J. Tarn, and J. W. Clark, *On the controllability of quantum-mechanical systems*, *Journal of Mathematical Physics* **24**, 2608 (1983).
- [24] D. D'Alessandro and R. Romano, *Further Results on the Observability of Quantum Systems Under General Measurement*, *Quantum Information Processing* **5**, 139 (2006).
- [25] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, *Quantum Fisher information matrix and multiparameter estimation*, *Journal of Physics A: Mathematical and Theoretical* **53**, 023001 (2020).
- [26] M. Yu, Y. Liu, P. Yang, M. Gong, Q. Cao, S. Zhang, H. Liu, M. Heyl, T. Ozawa, N. Goldman, and J. Cai, *Quantum Fisher information measurement and verification of the quantum Cramér-Rao bound in a solid-state qubit*, *npj Quantum Information* **8**, 56 (2022).
- [27] J. Liu, J. Chen, X.-X. Jing, and X. Wang, *Quantum Fisher information and symmetric logarithmic derivative via anti-commutators*, *Journal of Physics A: Mathematical and Theoretical* **49**, 275302 (2016).
- [28] J. M. Horowitz and T. R. Gingrich, *Thermodynamic uncertainty relations constrain non-equilibrium fluctuations*, *Nature Physics* **16**, 15 (2020).
- [29] C. Arenz, G. Gualdi, and D. Burgarth, *Control of open quantum systems: case study of the central spin model*, *New Journal of Physics* **16**, 065023 (2014).
- [30] J. Johansson, P. Nation, and F. Nori, *QuTiP 2: A Python framework for the dynamics of open quantum systems*, *Computer Physics Communications* **184**, 1234 (2013).
- [31] J. Johansson, P. Nation, and F. Nori, *QuTiP: An open-source Python framework for the dynamics of open quantum systems*, *Computer Physics Communications* **183**, 1760 (2012).