Autonomous Temporal Probability Concentration: Clockworks and the Second Law of Thermodynamics

Emanuel Schwarzhans,[1,*] Maximilian P. E. Lock,[1] Paul Erker,[1] Nicolai Friis,[1] and Marcus Huber[1,2,†]

[1]*Institute for Quantum Optics and Quantum Information-IQOQI Vienna, Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria*

[2]*Vienna Center for Quantum Science and Technology, Atominstitut, TU Wien, 1020 Vienna, Austria*

(Received 13 July 2020; revised 26 November 2020; accepted 24 December 2020; published 8 March 2021)

According to thermodynamics, the inevitable increase of entropy allows the past to be distinguished from the future. From this perspective, any clock must incorporate an irreversible process that allows this flow of entropy to be tracked. In addition, an integral part of a clock is a clockwork, that is, a system whose purpose is to temporally concentrate the irreversible events that drive this entropic flow, thereby increasing the accuracy of the resulting clock ticks compared to counting purely random equilibration events. In this article, we formalize the task of *autonomous temporal probability concentration* as the inherent goal of any clockwork based on thermal gradients. Within this framework, we show that a perfect clockwork can be approximated arbitrarily well by increasing its complexity. Furthermore, we combine such an idealized clockwork model, comprised of many qubits, with an irreversible decay mechanism to showcase the ultimate thermodynamic limits to the measurement of time.

DOI: 10.1103/PhysRevX.11.011046 Subject Areas: Quantum Physics, Quantum Information

I. INTRODUCTION

Time plays a special role in quantum physics. While other physical quantities of interest are represented as Hermitian operators, there is no observable corresponding to time itself. That is, it is not possible to find an operator conjugate to the Hamiltonian (representing energy) that may serve as "time observable" in the same way as is done for position and momentum [1] (see, e.g., Ref. [2] for some caveats to this statement). Time thus plays the role of a parameter in the equations of motion. Consequently, the passage of time is estimated via the evolution of a reference system—a *clock*. By tracking the dynamical evolution of (observable quantities related to) such a clock system, it is possible to extract information about the flow of time; see, e.g., Refs. [3–10]. *But what makes a specific system useful as a clock?* To address this question, we consider time to be a continuously elapsing parameter t ("Schrödinger time") whose value is estimated by a clock in terms of discrete increments ("ticks"). According to quantum theory, the evolution of any closed system is time-reversal symmetric, and, therefore, any complete description of an instrument that measures time inevitably requires an irreversible part that breaks this symmetry. By definition, the equilibrium state of any system features no nontrivial evolution in time. Thus, the first necessary ingredient for building a clock is an out-of-equilibrium system, such that the clock can harness the irreversible transition to higher entropy to produce ticks.

Entropy-increasing processes are fundamentally stochastic. Consequently, individual events resulting from such a process provide little information about t and, thus, make for rather bad clocks. While one could, in principle, use any equilibrating system as a clock—such as a hot coffee mug cooling down on your desk—its ticks, e.g., the spontaneous emissions of thermal photons (which exhibit super-Poissonian statistics), come at highly irregular intervals with respect to Schrödinger time. Structuring this irregular entropy flow into a series of ticks to allow for a precise synchronization of events is exactly the purpose of a clock. In this article, we formalize the task of timekeeping by conceptually separating two stages:

(i) an irreversible process that follows the second law of thermodynamics, i.e., an out-of-equilibrium system moving toward equilibrium by means of discrete and stochastic events, and

(ii) an internal clockwork that temporally concentrates the probability of an irreversible event occurring, thereby mitigating the fluctuations of the intervals between the equilibration events.

As we see, the particular choice of (i) provides the context for evaluating clock performance, because it

*emanuel.schwarzhans@oeaw.ac.at

†marcus.huber@univie.ac.at

represents a basic form of clock itself, while at the same time limiting the performance of a clock for any given clockwork. Stage (ii) gives rise to a clearly defined mathematical task that we refer to as *autonomous temporal probability concentration* (ATPC).

Here, we consider clocks to be autonomous. That is, the Hamiltonian generating the evolution of the clockwork is energy conserving and time independent, and the irreversible process is memoryless and requires no external control, i.e., *no active measurement*. Although current quantum clocks are usually far from autonomous, as they require power input and are subject to losses, both of which are usually not fully accounted for in their analysis, we focus on autonomous clocks in order to provide a full analysis of the resources that are fundamentally required to operate a clock.

To describe the performance of a clock, we use two quantities: *accuracy* and *resolution*. The accuracy $N = (\bar{t}/\Delta t)^2$ is the average number of ticks until the clock is, on average, off by one tick with respect to Schrödinger time. The resolution $R = 1/\bar{t}$ is the average of the tick frequency with respect to Schrödinger time. Note that this choice for quantifying the resolution is not the only possibility. We choose the above definition since it represents a conservative figure of merit, in the sense that it prevents statistical outliers of ticks occurring at small times t from unduly inflating the estimated resolution, as we discuss in more detail in Sec. III.

That there is a trade-off relation between accuracy and resolution, and that there is a proportional relation between the entropy dissipated in the process and the clock performance, was first noticed in a model of an autonomous quantum clock as an open quantum system in Ref. [11] and recently corroborated in a mesoscopic experiment in Ref. [12].

Here, we combine these aspects and provide a detailed investigation of the trade-offs between accuracy, resolution, and entropy production for given energy and complexity within the framework of autonomous quantum clocks [13,14]. A central tenet for providing these trade-offs is the separation of timekeeping into two separate processes mentioned above: (i) the irreversible out-of-equilibrium transitions of the clockwork via interaction with an environment, resulting in distinguishable events registered as "ticks," which we model with a decay mechanism, and (ii) the internal closed-system (unitary) dynamics that provide a clockwork and temporally concentrate the population of states from which an irreversible transition can emerge. That is, the clockwork ensures that the circumstances that allow for a tick to happen (e.g., a specific energy level resonant with the out-of-equilibrium dynamics being highly populated) occur only within a very narrow time window.

We first find that a simple clockwork can concentrate probability only in a limited fashion, prompting the question of whether more complex designs could perform better. We answer this question by finding an analytical relation between ATPC and complexity of a specific clockwork model. More generally, we identify the important features of a clockwork that lead to this improvement and prove that, for cold environments, ATPC can be performed arbitrarily well. Then, we investigate whether perfect ATPC allows for perfect clocks and find that the answer is no. In fact, the irreversible process sets a limit to the clock quality, and, while increasing the complexity (and, thus, the concentration of probability) first improves the quality of the clock, after a certain point a further increase actually is detrimental. We thus illustrate the trade-offs between accuracy, resolution, entropy production, and clockwork complexity.

The specific clock model that we consider here consists of (i) external heat baths as out-of-equilibrium resources, (ii) a quantum system representing the "clockwork," and (iii) an external field that the clockwork can emit energy ("ticks," e.g., photons) into. In Sec. II, we first discuss the role and choice of the clockwork and formalize the task of ATPC. In Sec. III, we then discuss mechanisms for coupling the clockwork to an equilibrating process to produce ticks. In Sec. IV, we combine the two, to showcase the limitations set by the irreversible process and how the complexity of a clockwork can be utilized to reach the maximal potential of a clock. We continue in Sec. V with a discussion of the implications and the relation to other literature on clocks and end with a short conclusion in Sec. VI.

II. THERMAL MACHINES AND THE CLOCKWORK

Let us now consider a clockwork in the sense discussed above, that is, a device that contains a target subsystem, which is to be prepared for an out-of-equilibrium transition, thus resulting in a tick. From a thermodynamic perspective, such a preparation requires work to be performed on the target, which can be achieved by a quantum thermal machine. Operating such a machine, in turn, requires an out-of-equilibrium resource, which we here consider to be provided by thermal baths at different temperatures, i.e., a thermal gradient. More specifically, we assume that two independent baths are available, a hot bath and a cold bath, at temperatures T_H and T_C, respectively, where the latter represents the environment. This setup is depicted in Fig. 1.

This choice is motivated, first, by the general availability of heat baths; i.e., it is the most common out-of-equilibrium resource found in nature, such as, e.g., the Sun. Second, because systems are usually expected to thermalize (eventually) without detailed external control or timing, i.e., preparing such heat baths does not require any timing device or detailed control of the system's internal structure, just an increase in average energy. Consequently, heat baths allow for transparent bookkeeping of the relevant resources, i.e., of the average amount of entropy dissipated by the clockwork for each tick.

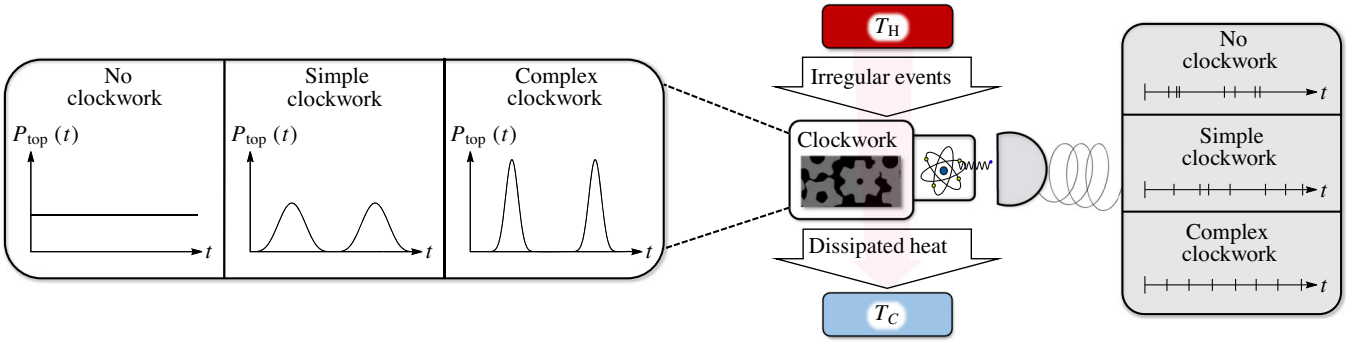

FIG. 1. Illustration of timekeeping at the level of individual irreversible events. The equilibration events that follow the second law of thermodynamics are inherently stochastic and irregular; in our example, we use radiative decays from an excited energy level of a quantum system (which we label "top"). By inserting an autonomously operating clockwork between the two out-of-equilibrium systems ("hot" decaying quantum systems and a "cold" environment), these decays are temporally structured by temporal variation of the population of the top level, a task that we refer to as autonomous temporal probability concentration. This task concentrates the probability of such a decay around the oscillatory peaks of excited population. The panel on the right showcases how greater clockwork complexity leads to a regularization of individual thermalization events, i.e., clock ticks, starting from a thermal population with randomly distributed ticks depicted above, continuing to a simple clockwork with limited population and still significant variance, resulting in ticks being more likely during peak populations and, thus, less frequent and more regular, and then finally a complex clockwork, increasing population while decreasing temporal variance, giving yet more regular ticks. The tick distribution on the right is exemplary and depicts random ticks, whose spacing approximates the cycle time of the clockwork as the temporal probability becomes more concentrated around sharp peaks.

A specific focus of the analysis performed here lies on the identification of trade-offs between different figures of merit for the clock performance for fixed energy input and clock complexity. In principle, the performance of a given clock also depends on the (difference between the) temperatures T_C and T_H. However, since we are primarily interested in upper bounds on the relevant figures of merit, we often concentrate on the case where the environment temperature is $T_C = 0$. For the sake of completeness, calculations for general T_C can be found in Appendixes B and C.

Our clockwork model then consists of two parts, a d-dimensional "*ladder*" target system (in the simplest case, a qubit, $d = 2$) and a machine, which itself has some substructure and couples to the ladder via unitary interaction. This interaction supplies work (which the machine draws from its coupling to the heat baths) to the ladder, driving it to its excited states. The ladder, in turn, couples irreversibly to an external field, and, thus, these excitations eventually result in ticks (i.e., energy emitted into the field). Here, we consider a model where only a nonzero population $P_{\text{top}}(t)$ of the "top level"—the most highly excited state of the ladder—can lead to a tick. Barring some improbable combination of selection rules, such a single sharp energy transition can in practice, of course, only be approximated. However, as becomes clear once we introduce our model, allowing the possibility of clock ticks occurring due to other transitions serves to spread the temporal profile of the ticks, decreasing probability concentration. In the spirit of deriving idealized but fundamental bounds, we therefore focus on decays resulting from only one particular transition. As a consequence, the quality of the clockwork depends on the properties of the particular probability distribution $P_{\text{top}}(t)$ as a function of Schrödinger time t. In particular, an ideal clockwork should be capable of producing

$$P_{\text{top}}(t) = \begin{cases} 1, & \text{if } t = t_0, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

While one would expect a perfect clockwork to be capable of producing this distribution, it is also clear that it is not always desirable in conjunction with an irreversible mechanism. If the probability is arbitrarily temporally concentrated, i.e., it is close to one for only a short period of time, but the coupling of the ladder to the external field is of finite strength, then the emission of the ladder energy into the field has a chance not to occur during the peak, thus skipping this tick and worsening the clock performance. Nonetheless, an ideal clockwork should be capable of approximating this ideal distribution to the desired precision set by the irreversible mechanism. Arguably, it seems implausible that a heat engine itself, which intrinsically also harnesses the stochastic flow of energy from a hot to a cold bath, should be able to produce such a perfect signal. However, it may be reasonable to expect that a sufficiently complex clockwork, itself driven by a heat engine, could approximate the ideal ATPC of Eq. (1). In the following, we therefore investigate the role of the complexity of the internal structure of the machine in approximating the ideal ATPC. In order to do so, we decompose the machine into a set of elementary few-qubit machines, each realizing an effective virtual qubit [15]. This set allows the number of (elementary) machines to be used as a proxy for the

complexity of the clockwork's microscopic structure. In terms of these quantifiers, i.e., the dimension d of the target system and the number $M(d-1)$ of virtual-qubit machines (we consider each of the $d-1$ transitions between neighboring energy levels of the ladder to be coupled to M virtual-qubit machines), a central result on autonomous probability concentration that we derive in this paper can be phrased as follows.

Result 1: Autonomous temporal probability concentration of qubit machines.—Driving a d-dimensional target system at temperature $T_C = 0$, with M virtual-qubit machines per transition between neighboring levels, autonomously allows a top-level probability of

$$P_{\text{top}}(t) = \left\{1 - \left[1 - \left(\frac{\mathcal{Z}_H - 1}{\mathcal{Z}_H}\right)^{d-1}\right]^M\right\} \sin^{2(d-1)}(gt) \tag{2}$$

to be reached. Here, $\mathcal{Z}_H$ is the partition function of a qubit coupled to the hot bath and can, thus, take values between 1 and 2.

In other words, we show in the following that the behavior of an ideal clockwork [i.e., Eq. (1)] can be approximated arbitrarily well by increasing the complexity of the clockwork, that is, by increasing M and d.

A. Two-qubit machine

We begin by considering the simplest possible heat-engine-driven clockwork: a two-dimensional ladder coupled to a cold bath (the environment) and to a two-qubit machine, i.e., $d = 2$ and $M = 1$. In terms of Hilbert space dimension, this thermal machine is the smallest possible [15], consisting of a cold qubit and a hot qubit, in contact with the cold environment and a hot bath, respectively, as illustrated in Fig. 2.

Before the machine is activated, the qubits interact only with their respective baths. Under the assumption of weak coupling between the qubits and baths, each qubit thermalizes to the corresponding bath temperature. Denoting the energy gaps of the hot, cold, and ladder qubits as E_H, E_C, and E_L, respectively, the reduced states of the qubits can be represented by the thermal states

$$\rho_i = \frac{e^{-\beta_i H_i}}{\mathcal{Z}_i}, \tag{3}$$

with $i = H, C, L$ and where $\mathcal{Z}_i = 1 + e^{-\beta_i E_i}$ are the respective partition functions and H_i the corresponding free Hamiltonians with eigenstates $|0_i\rangle$ and $|1_i\rangle$. The total initial state of the clockwork—the machine and the ladder—thus takes the form $\rho_0 = \rho_H \otimes \rho_C \otimes \rho_L$.

We further assume that the timescale of the interaction between the machine and target qubits is much shorter than that of their thermalization with the respective baths. Consequently, the relevant dynamics of the clockwork are well described by energy-conserving unitary processes on the clockwork Hilbert space $\mathcal{H} = \mathcal{H}_H \otimes \mathcal{H}_C \otimes \mathcal{H}_L$. This description corresponds to assuming that the energy scales of the clockwork are much greater than that of its coupling to the environment. It is suited to our purpose of obtaining fundamental limits to the task of ATPC, as relaxing this assumption results in the regularity of the clockwork being impaired by the randomness of the bath. Now, since the purpose of the machine is to transfer energy to the target system, we are interested in designing the internal structure of the clockwork, namely, the energy levels of the free Hamiltonian $H_0 = H_H + H_C + H_L$ and an interaction Hamiltonian H_{int} such that $[H_0, H_{\text{int}}] = 0$ and $[H_L, H_{\text{int}}] \neq 0$. This structure can be achieved by choosing the energy gaps to satisfy $E_H \geq E_C$ and $E_L = E_H - E_C$, which results in two degenerate energy levels of H_0: $|0_C 1_H 0_L\rangle$ and $|1_C 0_H 1_L\rangle$. This result, in turn, allows us to define an interaction Hamiltonian that acts nontrivially only within the degenerate subspace, given by

$$H_{\text{int}} = g(|1_C 0_H 1_L\rangle\langle 0_C 1_H 0_L| + |0_C 1_H 0_L\rangle\langle 1_C 0_H 1_L|), \tag{4}$$

where $g \in \mathbb{R}$ is a coupling constant. The unitary dynamics generated by the total Hamiltonian $H = H_0 + H_{\text{int}}$ hence conserves the total energy of the clockwork, since $[H_0, H_{\text{int}}] = 0$. However, since $[H_L, H_{\text{int}}] \neq 0$, the interaction, once activated, can perform work on the ladder.

The resulting dynamics leads to an increase of the population of the top energy level $|1_L\rangle$ of the ladder, which (in units where $\hbar = 1$) is given by

$$P_{\text{top}}(t) = \text{Tr}(|1_L\rangle\langle 1_L| e^{-iHt} \rho_0 e^{iHt}). \tag{5}$$

The maximally reachable population depends on the temperatures of the baths, as well as the energy gaps of the machine qubits [15]. The top-level probability in Eq. (5) evaluates to (see Appendix A)

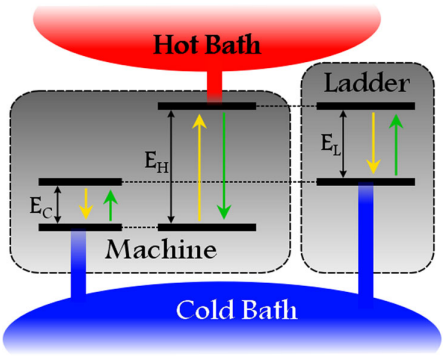

FIG. 2. Energy-level structure of the minimal thermal clockwork. The transitions induced by H_{int} are indicated by arrows. The green arrows indicate the transition where the ladder gets exited. The yellow arrows show the reverse transition. Coupling the qubit with the biggest energy gap to the hot bath introduces a bias toward the transition that is indicated by the green arrows.

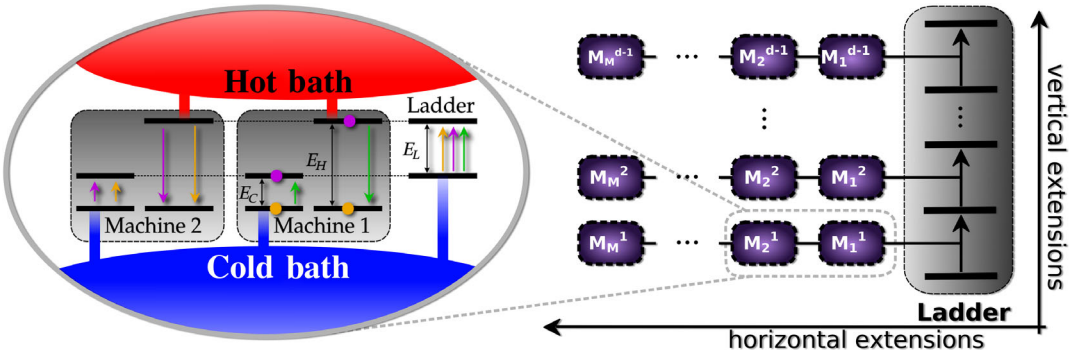

FIG. 3. On the right-hand side, the notion of "horizontal" and "vertical" extension is schematically illustrated. The ladder dimension d determines the number of vertical extensions ("rows"). Each pair of neighboring energy levels of the ladder couples to M copies of the two-qubit machine, where M determines the number of horizontal extensions ("columns"). On the left-hand side, the energy-level structure of a horizontally extended machine for a two-machine clockwork coupling to a two-level ladder ($d = 2$) is illustrated. This case is a special case of the generalized clockwork shown on the right-hand side. The different energy-conserving transitions induced by H_{int} are indicated by the differently colored arrows and dots, where the dots represent a conditioning of the transitions of machine 2 either on the ground state (orange) or on the exited state (purple) of machine 1. The reverse transitions are not depicted here for the sake of clarity.

$$P_{\mathrm{top}}(t) = \left(\frac{\mathcal{Z}_H - 1}{\mathcal{Z}_C \mathcal{Z}_H \mathcal{Z}_L}\right) \sin^2(gt) + \left(\frac{(\mathcal{Z}_C - 1)(\mathcal{Z}_L - 1)}{\mathcal{Z}_C \mathcal{Z}_H \mathcal{Z}_L}\right) \cos^2(gt) + \frac{\mathcal{Z}_L - 1}{\mathcal{Z}_L} - \frac{(\mathcal{Z}_C - 1)(\mathcal{Z}_L - 1)}{\mathcal{Z}_C \mathcal{Z}_H \mathcal{Z}_L}. \tag{6}$$

For $T_C = 0$, this probability simplifies to

$$P_{\mathrm{top}}(t) = \left(1 - \frac{1}{\mathcal{Z}_H}\right) \sin^2(gt). \tag{7}$$

Thus, even when $T_C = 0$, this function is far away from the ideal shape in Eq. (1), in terms of both its maximal value and the width of the distribution around its peak. Even in the limit $T_H \to \infty$, the maximal value reached at $t = (\pi/2g)$ is only $\frac{1}{2}$. Moreover, this top-level population could also be achieved by directly coupling the ladder to the hot bath. Thus, the two-qubit machine does not provide the desired ATPC by itself. However, in the following, we present a generalization of this framework which allows arbitrarily precise ATPC and, hence, an ideal clockwork to be approximated to within any given error.

B. Generalized machines

In the following, we present a generalized clockwork model that allows both the "sharpness" and the amplitude of $P_{\mathrm{top}}(t)$ to be controlled, while keeping track of all the relevant resources. This model can be achieved by two qualitatively different but compatible extensions that we refer to as "horizontal" and "vertical" extensions, as illustrated in Fig. 3. The horizontal extension allows the amplitude of $P_{\mathrm{top}}(t)$ to be increased, while the vertical extension allows the width of the peak of $P_{\mathrm{top}}(t)$ to be decreased, thus increasing its sharpness. Specifically, we add more levels to the target ladder and with it more two-qubit machines, interacting with each successive transition (vertical extension); to a given transition, we add more machines (horizontal extension). We start by collecting all interactions along a vertical column (see Fig. 3) of machines interacting with the ladder into a term H_1. This collection vertically extends the interaction of a single two-qubit machine [Eq. (4)] along all ladder states, i.e.,

$$H_1 = g \sum_{n=1}^{d-1} (|1_C 0_H\rangle\langle 0_C 1_H|_{M_1^n} \otimes |n+1_L\rangle\langle n_L| + \mathrm{H.c.}), \tag{8}$$

for the first vertical column, where M_i^j denotes the Hilbert space of the jth two-qubit machine acting on the ith ladder transition. We then add another term H_2, which does the same for the second vertical column, albeit with an additional projector onto the subspace orthogonal to the one on which H_1 acts nontrivially to ensure commutativity of H_1 and H_2. This process continues for M vertical columns, always projecting onto the orthogonal subspace of all previously added machines. Using $\mathrm{M}_{(i)}$ to denote the Hilbert space of the vertical group of the ith machine, i.e., $M_{(i)} := \bigotimes_{j=1}^{d-1} M_i^j$, we can then write our generalization of the interaction Hamiltonian from the previous section in a compact notation as

$$H_{\mathrm{int}} = \sum_{k=1}^{M} \bigotimes_{i=1}^{k-1} \mathbb{1}_{M_{(i)}} \otimes J_{M_{(k)}L} \bigotimes_{i=k+1}^{M} \Pi_{M_{(i)}} = \sum_{k=1}^{M} H_k. \tag{9}$$

Here, we define the projectors

$$\Pi_{M_{(i)}} := \mathbb{1}_{M_{(i)}} - \sum_{n=0}^{d-1} |n_{M_{(i)}}\rangle\langle n_{M_{(i)}}| \tag{10}$$

and the operator

$$J_{M_{(k)}L} := ig\sum_{n=1}^{d-1}\sqrt{n(d-n)}(|n_{M_{(k)}}, n_L\rangle \times \langle n-1_{M_{(k)}}, n-1_L| - \text{H.c.}), \tag{11}$$

and the states $|n_{M_{(k)}}\rangle$ are defined as

$$|n_{M_{(k)}}\rangle := \bigotimes_{j=1}^{n}|1_C0_H\rangle_{M_k^j}\bigotimes_{l=n+1}^{d-1}|0_C1_H\rangle_{M_k^l}. \tag{12}$$

That is, the state $|n_{M_{(k)}}\rangle$ can be considered to be the nth excited state of the kth vertical group $\mathrm{M}_{(k)}$ in the sense that the first n machines M_k^j for $j=1,\ldots,n$ are in the "used" state $|1_C0_H\rangle_{M_k^j}$, whereas the remaining $d-n+1$ machines M_k^l, with $l=n+1,\ldots,d-1$, are in the "unused" state $|0_C1_H\rangle_{M_k^l}$. The relative normalization factor $\sqrt{n(d-n)}$ of the different interaction terms in the Hamiltonian is precisely chosen such that the different subspace rotations are in phase to single out a $\sin^{2(d-1)}(gt)$ scaling of $P_{\text{top}}(t)$ as opposed to a mixture of different powers of trigonometric functions. For further details, see Appendix C.

In the following, we briefly discuss the horizontal and vertical extensions separately to outline their physical impact.

1. Horizontal extension

As shown in Appendix B, for $T_C=0$, the interaction Hamiltonian for $d=2$ in Eq. (9) then modifies the top-level probability from Eq. (7) to

$$P_{\text{top}}(t)=\left(1-\frac{1}{\mathcal{Z}_H^M}\right)\sin^2(gt). \tag{13}$$

For finite T_C, the weight of this sinusoidal term changes, and there are additional constant and cosine terms, whose relative weight increases with increasing T_C (see Appendix B).

From Eq. (13), we see that the maximal value of $P_{\text{top}}(t)$ increases with increasing M, and total population inversion can be achieved in the limit $M\to\infty$. However, in order to achieve ATPC, only increasing the magnitude of $P_{\text{top}}(t)$ is not sufficient, since this increase neglects the temporal concentration. In the next section, we therefore introduce the vertical extension, which allows us to temporally concentrate $P_{\text{top}}(t)$, leading to sharper peaks.

2. Vertical extension

For the vertical extension, we generalize the ladder to a nondegenerate system with d evenly spaced energy eigenstates, with the gap between neighboring states equal to E_L. To each of the $d-1$ pairs of neighboring energy levels of the vertically extended ladder, a two-qubit machine can be coupled in the way described in the previous section. In total, the vertically extended clockwork thus consists of a d-dimensional ladder and $d-1$ two-qubit machines, as illustrated in Fig. 3. The resulting top-level probability for $T_C=0$ becomes

$$P_{\text{top}}(t)=\left(\frac{\mathcal{Z}_H-1}{\mathcal{Z}_H}\right)^{(d-1)}\sin^{2(d-1)}(gt). \tag{14}$$

We thus see that just vertically extending the machine makes the temporal distribution sharper, but it also decreases the achievable top-level population.

3. General extended clockwork

Finally, by combining the horizontal and vertical extensions, we can combine the advantages of both, i.e., simultaneously increase the top-level population and the sharpness of the temporal distribution. Straightforward calculation of the top-level probability for $T_C=0$ (shown in detail in Appendix C) yields

$$P_{\text{top}}(t)=\left\{1-\left[1-\left(\frac{\mathcal{Z}_H-1}{\mathcal{Z}_H}\right)^{d-1}\right]^M\right\}\sin^{2(d-1)}(gt). \tag{15}$$

For T_C only slightly greater than zero, Eq. (15) smoothly approximates the above top-level probability (see Appendix C).

A direct consequence of the particular form of the top-level probability in Eq. (15) is that the amplitude and temporal variance (sharpness) of $P_{\text{top}}(t)$ can be optimized to within any desired error by controlling the number of machines (M) per neighboring pair of energy levels in the horizontal extension and the dimension of the ladder (d) in the vertical extension, respectively.

Since we restrict the clockwork to consist of qubit machines that have up to $M(d-1)$-body interactions, and the Hilbert space of the machine is $4^{M(d-1)}$ dimensional, the most reasonable quantifier of complexity should be related simply to M and d. We, therefore, focus on elucidating the role of M and d separately. What we can now see is that, in order to decrease the temporal variance (increasing d) while increasing the amplitude (increasing M), the complexity necessarily has to increase. Thus, for a fixed complexity, there exists a trade-off between temporal variance and probability amplitude.

In the following sections, we include the irreversible decay mechanism to numerically analyze how accuracy and resolution of clocks are influenced by changes in M and d.

III. IRREVERSIBILITY AND CLOCK TICKS

Any autonomous quantum clock (or any clock for that matter) inevitably produces entropy in order to tick [11], as it needs to be subject to an irreversible evolution. While the

internal clockwork produces a temporally well-concentrated and repeating distribution, there needs to be an irreversible process that turns this entropy into a measurable signal. For this process to happen, there needs to be a system that is driven out of equilibrium in order to relax back to equilibrium while producing a tick. In our case, the system that is driven out of equilibrium is the ladder. The system with respect to which the ladder is driven out of equilibrium is assumed to couple to the ladder such that the top level is unstable and decays, emitting energy into that system. As an example, one can take this system to be a photon field at the environment temperature T_C that couples to the ladder, such that when the top-level population decays to the ground state it emits a photon of energy $E_\gamma = (d-1)E_L$. However, note that the assumption that only this particular ladder transition couples to the field is an idealization for the purpose of deriving fundamental bounds (as discussed in Sec. II). Since any irreversible process can be viewed as a reversible process on a larger Hilbert space, the presence of such a decay channel must, in principle, also allow for the reverse process of exciting the ladder while absorbing energy, e.g., in the form of a photon in the example above. However, the probability for this process to happen can be made arbitrarily small by demanding that the background temperature of the field satisfies $(E_\gamma/k_B T_C) \gg 1$.

The number of possible decay processes is vast. However, since our aim is to capture all resources that are necessary to operate a clock, allowing for decay processes that require memory misses the purpose, since the required resources are not clearly defined for them. We therefore require the photon field to be memoryless, i.e., that correlations with the ladder are diluted very quickly and are, thus, negligible. The resulting dynamics are governed by the law of exponential decay and, thus, constitute an ideal case, giving an effective upper bound to the clock performance and allowing us to keep track of the resources that are invested. In particular, the probability density for a tick occurring at time t is given by (see Appendix D)

$$P_{\text{tick}}(t) = cP_{\text{top}}(t)e^{-c\int P_{\text{top}}(t')dt'}, \tag{16}$$

where c is the coupling strength of the photon field with the top level of the ladder.

Let us now consider the energetic resources required to run the clock. We first note that, taking $T_C = 0$, as the clockwork state evolves, each branch of its superposition where the ladder's top level is excited corresponds to a transition of $d-1$ machines: $|0_{M_{(k)}}\rangle \to |d-1_{M_{(k)}}\rangle$, where the value of k differs between branches, and we recall that the $\{|n_{M_{(k)}}\rangle\}$ are defined in Eq. (12). Thus, regardless of which branch is realized, if the ladder's top level is excited, then the heat flow from the hot bath into the total system is given by $Q_{\text{in}} = (d-1)E_H$. This heat flow does the work of driving the ladder from $|0\rangle_L$ to $|d-1\rangle_L$, i.e., $W = (d-1)E_L$. After the clock ticks and the cold qubits of the machines rethermalize, $Q_{\text{out}} = (d-1)E_C$ of heat is dissipated into the cold bath. Since $E_H = E_L + E_C$, we, thus, have the usual relation for a thermal machine, i.e., $Q_{\text{in}} = W + Q_{\text{out}}$, and the thermal efficiency of the process is $\eta_{\text{th}} := (W/Q_{\text{in}}) = (E_L/E_L + E_C)$. From this result, one can see that, as E_C/E_L decreases, we approach the Carnot efficiency bound $\eta_{\text{th}} \leq 1$.

Curiously, for $T_C = 0$ and $M \to \infty$, the top-level population is just $P_{\text{top}}(t) = \sin^{2(d-1)}(gt)$. If interpreted as a heat engine whose purpose is to charge a battery (the ladder), then one can indeed reach an efficiency of $\eta_{\text{th}} \approx 1$ and still charge the battery in finite time $\tau = (\pi/2g)$. Even the task of ATPC can be achieved to arbitrary precision at perfect efficiency. One can interpret this result as sufficient clockwork complexity permitting perfect efficiency at finite power.

In any case, the efficacy of ATPC and the resulting clock dynamics are essentially determined by the ladder dimension d and the number of driving machines M, which together correspond to a simple notion of *clockwork complexity*. In order to investigate how these affect the quality of the clock, we quantify this quality using two notions introduced in Ref. [11]. These are the accuracy, which is the average number of ticks until the next tick is off by the average time between two ticks, i.e.,

$$N = \left(\frac{\bar{t}}{\Delta t}\right)^2, \tag{17}$$

and the resolution, which is the inverse average time between two ticks, i.e.,

$$R = \frac{1}{\bar{t}}. \tag{18}$$

Here, $\bar{t} = \int_0^\infty tP_{\text{tick}}(t)dt$ and $(\Delta t)^2 = \overline{t^2} - (\bar{t})^2$ with $\overline{t^2} = \int_0^\infty t^2 P_{\text{tick}}(t)dt$.

Let us remark here that our choice of quantifier for the resolution R is not the only option. For instance, another candidate to quantify the resolution would be $\overline{1/t} = \int_0^\infty \frac{1}{t} P_{\text{tick}}(t)dt$. However, for this choice, small times would contribute much more strongly to the average than for the inverse of the average times. For example, already a single outlier at a very small value of t would result in a very large value of $\overline{1/t}$, and one would conclude that the resolution was very high, even if most of the events were observed at larger values of t. Conversely, choosing $1/\bar{t}$ as a quantifier means that even a few outliers at large values t would result in a low resolution. Therefore, our choice $R = 1/\bar{t}$ represents the more conservative of these choices, ensuring that our description results in upper bounds on the resolution.

In the following section, we present numerical calculations of the accuracy as a function of the resolution, the clockwork complexity, and the energy dissipated per tick.

For comparison, let us take as a baseline an example where there are no qubit machines employed, and the ladder simply begins in equilibrium with the hot bath and emits this energy into the cold bath via the irreversible process; i.e., there is no ATPC. In that case, the top-level probability is constant, i.e., $P_{\mathrm{top}}(t) = \exp[-\beta_H(d-1)E_L]/\mathcal{Z}_L$, which results in $R = (1/\bar{t}) = (1/\Delta t) = c\exp[-\beta_H(d-1)E_L]/\mathcal{Z}_L$, and, thus, the resolution is essentially determined by the decay rate c and the population of the decaying level. The accuracy is simply $N = 1$. This result highlights the main purpose of a clockwork: An individual event resulting from pure thermalization results in an accuracy of 1 and comes at a work cost of $(d-1)E_L$, whereas the clockwork can increase the accuracy while keeping the work cost of one tick constant.

IV. NUMERICAL RESULTS

Since we are interested in upper bounds on the clock quality, for the following results we assume the temperature of the hot bath to be infinite, i.e., $T_H \to \infty$, as well as $T_C = 0$. The curves in the following figures are generated numerically by varying three free parameters, namely, M, d, and c (varying g has the inverse effect of varying c—see Appendix F). In particular, each curve corresponds to fixed values of M and c while d varies; we display the accuracy N on the vertical axes and the ladder dimension d (Fig. 4), the resolution R (Fig. 5), and the energy dissipation rate ε (Fig. 6), respectively, on the horizontal axes. First, we analyze the relation between the sharpness of the peak of $P_{\mathrm{top}}(t)$ and the clock accuracy. Recalling the discussion in Sec. II B 2, we note that the sharpness of $P_{\mathrm{top}}(t)$ increases with increasing d, and the latter may therefore be used as a measure of the sharpness of $P_{\mathrm{top}}(t)$. In Fig. 4, the behavior of the accuracy as d increases is depicted for different ladder-bath coupling strengths c and different values of M. For small d, we see that the accuracy increases linearly. However, increasing the sharpness beyond a certain point leads to a decrease in accuracy (this behavior is discussed in detail in Sec. V). The value of c therefore puts a bound on the maximally achievable accuracy for all potential clocks. The same limiting behavior is apparent if we fix c and vary g instead (see Appendix F), for reasons that we discuss below.

In order to analyze both accuracy and resolution with respect to the clockwork complexity in Fig. 5, we compare those two quantities for different fixed values of M while varying d. Increasing M allows us to reach a higher maximal accuracy, while increasing d (which increases from right to left in Fig. 5) allows us to trade resolution for accuracy up to the optimal point, after which the accuracy reduces again. We further observe that all clocks with the same c and g lie under a curve defined by the case of $M \to \infty$. Increasing c allows for clocks of higher quality, i.e., that have higher accuracy and resolution. Furthermore, the position of the maximum depends on the value of g. Here, g is chosen to be equal to E_C. Increasing this value shifts the peak to the right, i.e., to higher resolutions (see Appendix F).

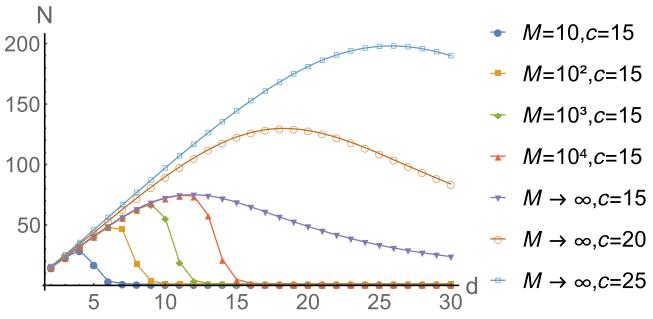

FIG. 4. The effect of the ladder dimension d on the clock accuracy N for different M. The top three curves show this effect in the limit $M \to \infty$ for different c, where $[c] = \mathrm{s}^{-1}$.

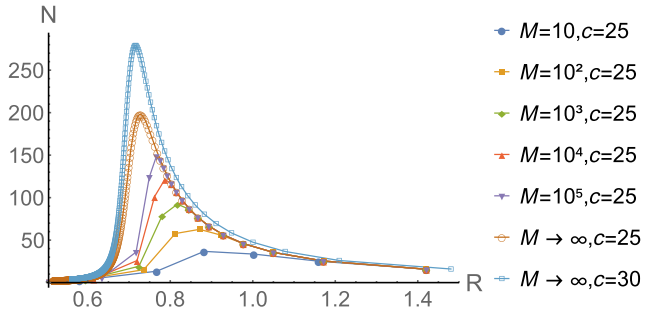

FIG. 5. The trade-off between clock accuracy N and resolution R for clockworks of various complexities, where R is increased by decreasing d. For finite horizontal extensions M, we show this behavior for a fixed photon field coupling $c = 25$, where $[c] = \mathrm{s}^{-1}$. Increasing M allows for higher maximal accuracy, and, thus, the orange line represents an upper bound of the accuracy for a given resolution for all clocks with $c = 25$. For $M \to \infty$, we see that increasing c increases the potentially achievable combinations accuracy and resolution. We choose $g = E_C$ in all cases.

Finally, in order to analyze the effect of the energy dissipation rate $\varepsilon = Q_{\mathrm{out}}/\bar{t}$ on the clock accuracy with respect to different values of M and d, which are simply related to the complexity of the clockwork (see Sec. II B 3), in Fig. 6 we plot the accuracy over the energy dissipation rate ε for clockworks of different complexity. In particular, we compare different values of M while varying d. What we can see in Fig. 6 is that, for fixed M (at fixed c and g), increasing the energy dissipation rate (which is achieved by increasing d) increases the accuracy at a certain slope until a maximum is reached. Increasing d further decreases the accuracy. Furthermore, for a given c, increasing M leads to

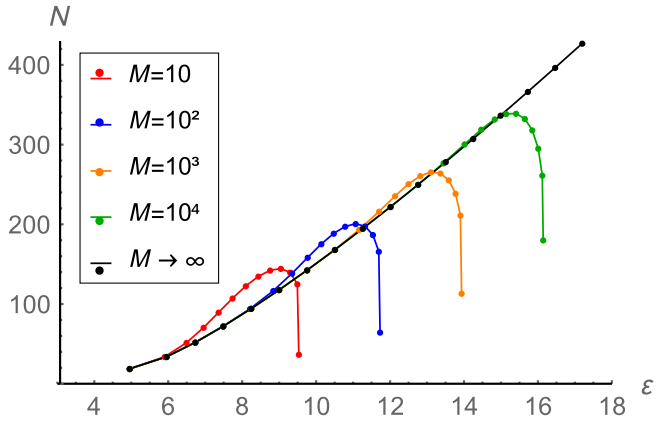

FIG. 6. The accuracy N as a function of the energy dissipation rate $\varepsilon = Q_{\text{out}}/\bar{t}$ (with $[\varepsilon] = E_C\ \text{s}^{-1}$) for different numbers of horizontal extensions M, with $c = 10^5\ \text{s}^{-1}$. As M increases, the maxima of the curves obtained by varying d are shifted toward higher values of N and ε, i.e., up and to the right, reaching finite values (not shown) in the limit $M \to \infty$ (black curve). The individual peaks obtained for a given coupling strength c and number M of horizontal extensions correspond to the clocks that achieve maximal accuracy under these constraints. We exclude suboptimal cases, i.e., cases where increasing d reduces resolution and accuracy.

a lower slope (approaching the slope of $M \to \infty$) while allowing for higher maximal accuracy, which suggests that a greater maximal accuracy can be achieved at the cost of a greater energy dissipation rate. We also note that increasing c increases the maximally achievable accuracy N, as can also be seen in Figs. 4 and 5. At the same time, increasing c increases the resolution (which can be seen in Fig. 6). Since Q_{out} does not depend on c, this increase in resolution leads to a higher energy-dissipation rate for all values of d and M, thus increasing the slope of the N-ε curve (in the linear regime). For reasons of visual clarity, this increase is not depicted in Fig. 6. One should note, however, that increasing c indefinitely would eventually break the assumptions inherent in our analysis and force us to explore deviations from a Markovian exponential decay toward memory effects in the irreversible dynamics.

V. DISCUSSION

Our results have two general implications. The first concerns the task of autonomous probability concentration. We show that, in principle, sufficiently increasing the clockwork complexity alone is enough to concentrate the temporal probability arbitrary well. In particular, this task can again be split into two conceptually different subtasks: maximizing the achievable population and improving the temporal sharpness with which this (maximal) population can be reached. By splitting the clockwork into a target ladder and virtual-qubit machines coupled to the different ladder transitions, we are able to analyze how more complex clockworks can help to achieve the two respective subtasks. While we work with equally spaced ladder systems, the same result (qualitatively) also holds for arbitrarily spaced target Hamiltonians, simply by redefining the respective coupling strengths (the g's) of the interaction Hamiltonians. We equip our clockwork with a particular tensor product structure, the division into two-qubit machines, for the sake of keeping track of its complexity. Our machine operates optimally within the framework set by this division, but whether more general machines could also achieve the same performance with a smaller overall size remains an open question.

In our analysis, we optimize the internal structure of the clock, i.e., the clockwork, to concentrate the probability in a fashion that most closely resembles the temporal distribution of an ideal clockwork. For given c, this optimization amplifies the clock quality only up to a limit, which we showcase in Fig. 4. This limit can intuitively be understood by considering the two key timescales of the clock, namely, that of the clockwork's dynamics and that of the irreversible decay. Increasing d while keeping c fixed, one eventually reaches a point where $P_{\text{top}}(t)$ is so well concentrated temporally that the comparative slowness of the decay mechanism reduces the probability that the clock will tick. In other words, it becomes more likely that the decay mechanism will skip that peak. We see the inverse of this behavior if, instead of c, we consider curves of fixed g (see Appendix F), as increasing g speeds up the clockwork, effectively making the limit imposed by c more restrictive.

This result brings us to the second implication of our work. The irreversible mechanism, in our case characterized by the parameter c, puts an absolute upper bound on the achievable combinations of resolution and accuracy, i.e., the clock quality, and thus determines the potential for how well a particular physical process can be used as the basis for a clock. The question of how well this upper bound can be approximated brings us to the role of our two extensions. First of all, the horizontal extension, i.e., the coupling of multiple elementary machines to a single transition between neighboring ladder levels, primarily serves the purpose of increasing the possible population inversion and with it the achievable top-level population. As we see in Eq. (16), c always appears multiplied by the prefactor of the sine in Eq. (15), resulting in an *effective coupling*:

$$C_M = c\left\{1 - \left[1 - \left(\frac{\mathcal{Z}_H - 1}{\mathcal{Z}_H}\right)^{d-1}\right]^M\right\}. \tag{19}$$

From this result, we can see that increasing the horizontal extension M is physically equivalent to increasing the coupling c, and this equivalence is why they play the same role in Fig. 4 (though we note that C_M is bounded with respect to M but not with respect to c). One cannot make a similar statement to relate c with the vertical extension d,

since the exponent of the sine in Eq. (15) varies as d does. As noted above, d sharpens the temporal distribution, thus increasing the accuracy of the clock as long as the limit set by c and M [via Eq. (19)] is not surpassed, as demonstrated in Fig. 4.

In the regime where the accuracy grows linearly with the sharpness (Fig. 4), which is determined by d (see Appendix F), there exists a trade-off relation between accuracy and resolution. To see this relation, note that the resolution decreases monotonically with d [see Fig. 7(b)]. For fixed c and g, the case of $M \to \infty$ represents an upper bound on the clockwork quality, i.e., on the achievable combinations of accuracy and resolution, which is illustrated in Fig. 5.

In our computations, we focus on the limit $T_C \to 0$. State-of-the-art atomic clocks operate at optical frequencies [16] and at even higher frequencies in novel proposals [17]. The vacuum state of any optical-frequency mode of the electromagnetic field has a population of approximately 1 at room temperature, and this situation is thus virtually indistinguishable from temperature 0. For clocks operating at much lower frequencies, such as those based on microwave transitions, cooling the environmental degrees of freedom into which the irreversible mechanism dissipates heat would be necessary in order to approach the fundamental limits we derive here.

It is nonetheless important to stress that we are interested in fundamental limits of timekeeping and the associated complexity and cost, which one of the reasons why we consider autonomous clocks in contrast with atomic clocks, for example, which require external control. For the practical purpose of building clocks for everyday use, atomic clocks can require as little as 30 mW [18], which is many orders of magnitude above the scale defined by the system energy and the timescale of the relevant processes but still insignificant for global energy use. The majority of that cost, however, comes from the fact that at some point that single event needs to be amplified and registered by a measurement apparatus, whose inherent irreversible nature is also thermodynamic and comes with its own costs and limitations [19]. Conversely, the inevitable imperfections of clocks and the associated costs also limit the achievable quality of measurements and, consequently, of all estimation procedures, e.g., of work itself [20].

As far as current and future prospects are concerned, the limits derived in this paper have more fundamental relevance for the autonomous control of quantum systems by a quantum clock [21,22]. Here, a small quantum system is envisioned to be controlled by an autonomous quantum clock. This control is important for any type of unitary process requiring precise timing, from small machines operating in cycles [23] to general repeating unitary processes such as circuit-based cooling models [24–28].

Coming back to the actual energy cost, there are a number of interesting observations that follow from our clock model. First of all, the horizontal extension always comes at a finite energy cost and dissipation for each tick for any clock. The vertical extension, on the other hand, linearly increases the energy cost and dissipation and, as long as the limit imposed by the c and M is not exceeded, also increases the accuracy as d increases. This limit can be qualitatively understood as a point after which further concentrating probability is actually detrimental to clock performance. We can observe that the increase in accuracy follows a linear behavior in d, before switching to a sublinear behavior close to the peak, after which it actually decreases accuracy again. Thus, in addition to recovering the observation that $N \propto \Delta S$, i.e., that a clock's accuracy is essentially determined by the entropy it dissipates (which seems to be a prevalent feature in all classical and quantum clocks [12]), we pinpoint which combination of the resources M, d, and c allow us to maintain this linear regime. The proportionality factor itself can be identified numerically from the plots in Fig. 6.

Finally, let us put our clockwork model in context with recent literature on quantum clocks. In Refs. [29–32], the relationship between achievable clock accuracy as a function of "clock dimension" is studied by means of repeated applications of maps from a clock system to a register. These works provide fundamental bounds for clock accuracy for fixed system dimension d, showing that the accuracy of classical (incoherent) clocks can at best scale linearly in d, whereas a quantum clock's accuracy (with states featuring coherence) may scale as d^2. The clock system considered in these works is exactly what we here refer to as the ladder system. Meanwhile, the map that Refs. [29–33] refer to as being responsible for creating a tick event in the register subsumes the interactions between the ladder, the qubit machines, and the heat baths and also includes the irreversible mechanism and the subsequent readout. In other words, our work provides a concrete physical realization of the maps that effect the transfer of ticks to the register. In Fig. 4, we see that, in the regime of the clockwork not exceeding the clock potential dictated by the irreversible mechanism, the accuracy scales linearly with the ladder dimension d, which is already the optimal achievable scaling [32].

VI. CONCLUSION

In this article, we have put forward a framework for studying fundamental limits of timekeeping. The conceptual split of any such task into a *clockwork*, which creates a temporally concentrated probability distribution and a mechanism for irreversibility, allowed us to derive an analytic formula for the achievable temporal probability concentration of the clockwork. The irreversible mechanism provides a context for the operation of the clock by allowing the passage of time to be tracked in the first place. Meanwhile, the chosen irreversible mechanism sets the reference timescale that ultimately constrains the

potential of any clockwork that harnesses this mechanism to form a clock. But it is the clockwork that needs to be appropriately tuned to achieve maximal performance given these constraints. By composing the clockwork of the smallest possible thermal machines, we were further able to conceptually split the task of autonomous probability concentration into two subtasks. First, by having more machines interact with a single transition, we can increase the maximum top-level probability and, with it, the effective coupling to the irreversible mechanism. Second, by concatenating multiple transitions of this kind, we are able to sharpen the temporal distribution. This sharpening reveals the intricate ways in which the complexity of the clockwork determines its performance.

In the future, it will be interesting to study more exotic irreversible mechanisms beyond exponential decay and whether they could be harnessed to further improve the clock quality. Moreover, one might study increasing clockwork complexity in a manner other than the addition of qubits. Furthermore, beyond simply asking if optimal ATPC is achievable in some limit, it would be of interest to ask which temporal probability profile [i.e., $P_{\text{tick}}(t)$] optimizes clock performance under some constraints and the extent to which such a profile is achievable with a quantum machine. Another interesting avenue would be to make a stronger connection to experimental implementations of clocks and how the notion of irreversibility versus temporal probability concentration can be more formally made on larger scales. One experiment in that direction is performed in Ref. [12], using a nanomechanical membrane. Photoisomerization could also provide a potential platform for implementing such a clock at molecular scales [34]. While there are many open paths to explore and questions to answer, our results consolidate the fact that perfect clocks are practically impossible when derived from first principles and that significant thermodynamic resources have to be invested to reach the potential of any physical system to act as a clock.

ACKNOWLEDGMENTS

We acknowledge support from the Austrian Science Fund (FWF) through the START Project No. Y879-N2, the Zukunftskolleg ZK03, and Project No. P 31339-N27. We further acknowledge funding from the Erwin Schrödinger Center for Quantum Science & Technology (ESQ) Discovery grant "Emergence of physical laws: From mathematical foundations to applications in many body physics."

APPENDIXES

In these Appendixes, we provide detailed derivations and background information for the results presented in the main text. In Appendix A, we first present a derivation of the top-level probability for the two-qubit machine of Sec. II A. In Appendix B, we then present the derivation of the top-level probability for the horizontal extension of the clockwork for arbitrary temperatures T_H and T_C. In Appendix C, we then again focus on the case $T_C = 0$, for which we derive the top-level probability in the full horizontal and vertical extension. Appendix D contains the derivation of the tick probability. Appendix E presents the details for the numerical computation of accuracy and resolution. Appendix F discusses the behavior of clocks with changing ladder dimension d as well as changing coupling constant g.

APPENDIX A: TOP-LEVEL PROBABILITY OF A TWO-QUBIT CLOCKWORK

Here, we present a derivation of the top-level probability $P_{\text{top}}(t)$ from Eq. (7). That is, we consider the minimal clockwork discussed in Sec. II A, which consists of a hot qubit (coupling to the hot bath at temperature T_H, as well as a cold qubit and a ladder (both coupling to a cold bath at temperature T_C). The derivation presented here is a special case ($M = 1$ and $T_C = 0$) of the more general derivation of the top-level probability within the horizontal extension that we present in Appendix B (where $M > 1$ and both T_H and T_C can take on arbitrary values). Nevertheless, we first go through the much simpler derivation for $M = 1$ and $T_C = 0$ here, which serves as a guiding example for the much more involved general calculation that is to follow.

Assuming that the systems have thermalized with their respective baths, the initial state of the clockwork is given by

$$\rho_0 = |0\rangle\langle 0|_C \otimes \tau_H \otimes |0\rangle\langle 0|_L = |0\rangle \times \langle 0|_C \otimes \frac{1}{\mathcal{Z}_H}(|0\rangle\langle 0|_H + e^{-\beta_H E_H}|1\rangle\langle 1|_H) \otimes |0\rangle\langle 0|_L, \tag{A1}$$

where $\mathcal{Z}_H = 1 + e^{-\beta_H E_H}$ is the partition function of the hot qubit. Since the interaction term H_{int} in the total Hamiltonian $H = H_0 + H_{\text{int}}$ is chosen such that the free energy (that is, with respect to the free Hamiltonian H_0) is conserved, $[H_0, H_{\text{int}}] = 0$, and because the initial state ρ_0 is diagonal in the eigenbasis of H_0, we can write the top-level probability from Eq. (5) as

$$\begin{aligned} P_{\text{top}}(t) &= \text{Tr}(|1\rangle\langle 1|_L e^{-iHt}\rho_0 e^{iHt}) \\ &= \text{Tr}(|1\rangle\langle 1|_L e^{-iH_{\text{int}}t}\rho_0 e^{iH_{\text{int}}t}) \\ &= \text{Tr}(T_0|0\rangle\langle 0|_C \otimes \tau_H T_0^\dagger), \end{aligned} \tag{A2}$$

where T_0 is a matrix encoding the transition amplitude between the ground state and excited state of the ladder, given by

$$T_0 =_L \langle 1|e^{-iH_{\rm int}t}|0\rangle_L = \sum_{k=0}^{\infty} \frac{(-it)^k}{k!}\, {}_L\langle 1|H_{\rm int}^k|0\rangle_L. \tag{A3}$$

Now, because $H_{\rm int}^2$ is proportional to the identity on its support, that is,

$$\begin{aligned} H_{\rm int}^2 &= g^2(|1\rangle\langle 0|_C \otimes |0\rangle\langle 1|_H \otimes |1\rangle\langle 0|_L + \text{H.c.})^2 \\ &= g^2(|1\rangle\langle 1|_C \otimes |0\rangle\langle 0|_H \otimes |1\rangle\langle 1|_L + |0\rangle\langle 0|_C \otimes |1\rangle\langle 1|_H \otimes |0\rangle\langle 0|_L), \end{aligned} \tag{A4}$$

even powers of $H_{\rm int}$ do not contribute to T_0. However, for odd powers, we have $H_{\rm int}^{2k+1} = g^{2k}H_{\rm int}$, such that we find ${}_L\langle 1|H_{\rm int}^{2k+1}|0\rangle_L = g^{2k+1}|1\rangle\langle 0|_C \otimes |0\rangle\langle 1|_H$. With this result, we can evaluate the transition matrix T_0, i.e.,

$$\begin{aligned} T_0 &= \sum_{k=0}^{\infty} \frac{(-it)^{2k+1}}{(2k+1)!}\, {}_L\langle 1|H_{\rm int}^{2k+1}|0\rangle_L \\ &= \sum_{k=0}^{\infty} \frac{(-igt)^{2k+1}}{(2k+1)!} |1\rangle\langle 0|_C \otimes |0\rangle\langle 1|_H = \sin(gt)|1\rangle\langle 0|_C \otimes |0\rangle\langle 1|_H. \end{aligned} \tag{A5}$$

Inserting the result into Eq. (A2), we finally arrive at the top-level probability

$$\begin{aligned} P_{\rm top}(t) &= \mathrm{Tr}(T_0|0\rangle\langle 0|_C \otimes \tau_H T_0^\dagger) \\ &= \sin^2(gt)\, {}_L\langle 1|\tau_H|0\rangle_L \\ &= \sin^2(gt)\frac{e^{-\beta_H E_H}}{1+e^{-\beta_H E_H}} = \sin^2(gt)\left(1-\frac{1}{\mathcal{Z}_H}\right). \end{aligned} \tag{A6}$$

APPENDIX B: THE HORIZONTAL EXTENSION

In this Appendix, we present more technical details of the horizontal extension of the autonomous clockwork from a single ($M = 1$) to multiple ($M > 1$) two-qubit machines interacting with the same two-level ($d = 2$) transition of the ladder system. In particular, we derive the top-level probability for the horizontal extension for arbitrary temperatures T_C and T_H.

Following a similar approach as in Eq. (A2), we define transition operators

$$\begin{aligned} T_n &:= {}_L\langle 1|e^{-iH_{\rm int}t}|n\rangle_L = {}_L\langle 1|\sum_{j=0}^{\infty}\frac{(-it)^j}{j!}H_{\rm int}^j|n\rangle_L \\ &= {}_L\langle 1|\sum_{j=0}^{\infty}\frac{(-it)^j}{j!}\sum_{k=1}^{M}H_k^j|n\rangle_L, \end{aligned} \tag{B1}$$

for $n = 0, 1$, where the last equality follows from the fact that the interaction terms H_k given by the terms in Eq. (9) have mutually disjoint support, i.e., $H_k H_{k'} = 0$ for $k \neq k'$. Before we calculate these transition operators, we note that H_k satisfies the cyclic property

$$H_k^{2q} = g^{2q}\bigotimes_{i=1}^{k-1}\mathbb{1}_{M_i} \otimes (|0_C1_H\rangle\langle 0_C1_H|_{M_k} \otimes |0\rangle\langle 0|_L + |1_C0_H\rangle\langle 1_C0_H|_{M_k} \otimes |1\rangle\langle 1|_L)\bigotimes_{j=k+1}^{M}\Pi_{M_j} \quad \text{for } q \in \mathbb{N}_{>0}, \tag{B2}$$

$$H_k^{2q+1} = g^{2q+1}\bigotimes_{i=1}^{k-1}\mathbb{1}_{M_i} \otimes (\sigma_{M_k}^- \otimes \sigma_L^+ + \sigma_{M_k}^+ \otimes \sigma_L^-)\bigotimes_{j=k+1}^{M}\Pi_{M_j} = g^{2q+1}H_k \quad \text{for } q \in \mathbb{N}, \tag{B3}$$

where $M_i = M_i^1 = C_i^1 \otimes H_i^1$ denotes the Hilbert space of the ith horizontal extension, $\sigma_L^+ := |1_L\rangle\langle 0_L| = (\sigma_L^-)^\dagger$, $\sigma_{M_k}^+ := |0_C1_H\rangle\langle 1_C0_H|_{M_k} = (\sigma_{M_k}^+)^\dagger$, and $|l_Cm_H\rangle\langle p_Cq_H|_{M_k} := |l\rangle\langle p|_{C_k} \otimes |m\rangle\langle q|_{H_k}$, with $l, m, p, q = 0, 1$.

Now, for the transition operator T_0, we note that only odd powers of H_k can map $|0\rangle_L$ to $|1\rangle_L$, and therefore

$$T_0 = \sum_{q=0}^{\infty}\sum_{k=1}^{M}\frac{(-it)^{2q+1}}{(2q+1)!}\, {}_L\langle 1|H_k^{2q+1}|0\rangle_L = \sin(gt)\left(\sum_{k=1}^{M}\bigotimes_{j=1}^{k-1}\mathbb{1}_{M_j} \otimes \sigma_{M_k}^- \bigotimes_{l=k+1}^{M}\Pi_{M_l}\right). \tag{B4}$$

We can calculate T_1 similarly, noting that only even powers of H_k contain the factor $|1\rangle\langle 1|_L$:

$$T_1 = \sum_{q=0}^{\infty}\sum_{k=1}^{M}\frac{(-it)^{2q}}{(2q)!}\, {}_L\langle 1|H_k^{2q}|1\rangle_L = \cos(gt)\left(\sum_{k=1}^{M}\bigotimes_{j=1}^{k-1}\mathbb{1}_{M_j} \otimes |1_C0_H\rangle\langle 1_C0_H|_{M_k} \bigotimes_{l=k+1}^{M}\Pi_{M_l}\right) + \tilde{\Pi}, \tag{B5}$$

where $\tilde{\Pi}$ is a projection defined by

$$\tilde{\Pi} \coloneqq \mathbb{1}_{\mathcal{H}\setminus L} - \sum_{k=1}^{M} \bigotimes_{j=1}^{k-1} \mathbb{1}_{M_j} \otimes |1_C 0_H\rangle\langle 1_C 0_H|_{M_k} \bigotimes_{l=k+1}^{M} \Pi_{M_l} = \bigotimes_{k=1}^{M} \Pi_{M_k} + \sum_{k=1}^{M} \bigotimes_{j=1}^{k-1} \mathbb{1}_{M_j} \otimes |0_C 1_H\rangle\langle 0_C 1_H|_{M_k} \bigotimes_{j'=k+1}^{M} \Pi_{M_{j'}}. \tag{B6}$$

In order to evaluate the top-level probability, let us briefly inspect the initial state ρ_0 in this situation, which is given by

$$\begin{aligned}\rho_0 &= \bigotimes_{i=1}^{M} \tau_{C_i} \otimes \tau_{H_i} \otimes \tau_L \\ &= \underbrace{\bigotimes_{i=1}^{M} \left(\frac{1}{\mathcal{Z}_C}|0\rangle\langle 0|_{C_i} + \frac{\mathcal{Z}_C - 1}{\mathcal{Z}_C}|1\rangle\langle 1|_{C_i}\right) \otimes \left(\frac{1}{\mathcal{Z}_H}|0\rangle\langle 0|_{H_i} + \frac{\mathcal{Z}_H - 1}{\mathcal{Z}_H}|1\rangle\langle 1|_{H_i}\right)}_{\mathrm{Tr}_L(\rho_0)} \otimes \left(\frac{1}{\mathcal{Z}_L}|0\rangle\langle 0|_L + \frac{\mathcal{Z}_L - 1}{\mathcal{Z}_L}|1\rangle\langle 1|_L\right).\end{aligned} \tag{B7}$$

We split the initial state into two parts: one where the ladder is initially exited and one where it is not, i.e.,

$$\rho_0 = \frac{1}{\mathcal{Z}_L}\mathrm{Tr}_L(\rho_0) \otimes |0\rangle\langle 0|_L + \frac{\mathcal{Z}_L - 1}{\mathcal{Z}_L}\mathrm{Tr}_L(\rho_0) \otimes |1\rangle\langle 1|_L. \tag{B8}$$

The top-level probability can then be seen to split into two separate contributions, corresponding to the two terms in Eq. (B8), that is,

$$P_{\mathrm{top}}(t) = \frac{1}{\mathcal{Z}_L}\mathrm{Tr}[T_0 \mathrm{Tr}_L(\rho_0) T_0^\dagger] + \frac{\mathcal{Z}_L - 1}{\mathcal{Z}_L}\mathrm{Tr}[T_1 \mathrm{Tr}_L(\rho_0) T_1^\dagger]. \tag{B9}$$

We first consider the part of ρ_0 where the ladder is initially in the ground state. Considering $\bigotimes_{j=1}^{k-1} \mathbb{1}_{M_j} \otimes \sigma^-_{M_k} \bigotimes_{l=k+1}^{M} \Pi_{M_l}$ in Eq. (B4), we see that for each k there are $k-1$ machines that are acted upon only by identities, meaning the partial trace over each of these machines simply contributes a factor 1. There are $M-k$ machines that are acted upon by an operator Π_{M_j}, each leading to a factor of

$$\begin{aligned}&\mathrm{Tr}\left[\frac{1}{\mathcal{Z}_H \mathcal{Z}_C}\Pi_{M_j}[|0\rangle\langle 0|_{C_j} + (\mathcal{Z}_C - 1)|1\rangle\langle 1|_{C_j}] \otimes [|0\rangle\langle 0|_{H_j} + (\mathcal{Z}_H - 1)|1\rangle\langle 1|_{H_j}]\Pi^\dagger_{M_j}\right] \\ &\quad = \mathrm{Tr}\left[\frac{1}{\mathcal{Z}_H \mathcal{Z}_C}[|0_C 0_H\rangle\langle 0_C 0_H|_{M_j} + (\mathcal{Z}_H - 1)(\mathcal{Z}_C - 1)|1_C 1_H\rangle\langle 1_C 1_H|_{M_j}]\right] = \frac{1 + (\mathcal{Z}_C - 1)(\mathcal{Z}_H - 1)}{\mathcal{Z}_H \mathcal{Z}_C}\end{aligned} \tag{B10}$$

in $\mathrm{Tr}[T_0 \mathrm{Tr}_L(\rho_0) T_0^\dagger]$. In addition, there is always exactly one machine which is acted upon by $\sigma^-_{M_k}$, contributing a factor of

$$\begin{aligned}&\mathrm{Tr}\left[\frac{1}{\mathcal{Z}_H \mathcal{Z}_C}\sigma^-_{M_k}[|0\rangle\langle 0|_{C_k} + (\mathcal{Z}_C - 1)|1\rangle\langle 1|_{C_k}] \otimes [|0\rangle\langle 0|_{H_k} + (\mathcal{Z}_H - 1)|1\rangle\langle 1|_{H_k}]\sigma^+_{M_k}\right] \\ &\quad = \mathrm{Tr}\left[\frac{1}{\mathcal{Z}_H \mathcal{Z}_C}(\mathcal{Z}_H - 1)|1_C 0_H\rangle\langle 1_C 0_H|\right] = \frac{\mathcal{Z}_H - 1}{\mathcal{Z}_H \mathcal{Z}_C}\end{aligned} \tag{B11}$$

in $\mathrm{Tr}[T_0 \mathrm{Tr}_L(\rho_0) T_0^\dagger]$. The first term of $P_{\mathrm{top}}(t)$ is thus given by the sum over all $k \in 1, 2, \ldots, M$ [see Eq. (B4)], multiplied by the initial population of the ladder ground state, resulting in

$$\frac{1}{\mathcal{Z}_L}\mathrm{Tr}[T_0 \mathrm{Tr}_L(\rho_0) T_0^\dagger] = \frac{1}{\mathcal{Z}_L}\frac{\mathcal{Z}_H - 1}{\mathcal{Z}_H \mathcal{Z}_C}\sum_{k=1}^{M}\left(\frac{1 + (\mathcal{Z}_C - 1)(\mathcal{Z}_H - 1)}{\mathcal{Z}_H \mathcal{Z}_C}\right)^{k-1} \sin^2(gt). \tag{B12}$$

The second part of $P_{\mathrm{top}}(t)$ can be calculated in the same way. Comparing Eq. (B5) with Eq. (B4), one sees that (aside from the oscillating scalar factors) the first term of T_1 differs from T_0 by replacing $\sigma^-_{M_k}$ in the latter with $|1_C 0_H\rangle\langle 1_C 0_H|_{M_k}$, and the corresponding factor $\mathcal{Z}_H - 1/\mathcal{Z}_H \mathcal{Z}_C$ is thus replaced by

$$\mathrm{Tr}\left[\frac{1}{\mathcal{Z}_H\mathcal{Z}_C}|1_C0_H\rangle\langle 1_C0_H|_{M_k}[|0\rangle\langle 0|_{C_k}+(\mathcal{Z}_C-1)|1\rangle\langle 1|_{C_k}]\otimes[|0\rangle\langle 0|_{H_k}+(\mathcal{Z}_H-1)|1\rangle\langle 1|_{H_k}]|1_C0_H\rangle\langle 1_C0_H|_{M_k}\right]$$
$$=\mathrm{Tr}\left[\frac{1}{\mathcal{Z}_H\mathcal{Z}_C}(\mathcal{Z}_C-1)|1_C0_H\rangle\langle 1_C0_H|_{M_k}\right]=\frac{\mathcal{Z}_C-1}{\mathcal{Z}_H\mathcal{Z}_C}. \tag{B13}$$

The transition operator T_1 additionally contains the static term $\tilde{\Pi}$, which means that $P_{\mathrm{top}}(t)$ contains the additional term

$$\mathrm{Tr}\left[\frac{1}{\mathcal{Z}_H^M\mathcal{Z}_C^M}\tilde{\Pi}\bigotimes_{k=1}^{M}[|0\rangle\langle 0|_{C_k}+(\mathcal{Z}_C-1)|1\rangle\langle 1|_{C_k}]\otimes[|0\rangle\langle 0|_{H_k}+(\mathcal{Z}_H-1)|1\rangle\langle 1|_{H_k}]\tilde{\Pi}\right]$$
$$=\frac{[1+(\mathcal{Z}_C-1)(\mathcal{Z}_H-1)]^M}{\mathcal{Z}_H^M\mathcal{Z}_C^M}+\sum_{k=1}^{M}\frac{[1+(\mathcal{Z}_C-1)(\mathcal{Z}_H-1)]^{k-1}(\mathcal{Z}_H-1)}{\mathcal{Z}_H^{k-1}\mathcal{Z}_C^{k-1}}. \tag{B14}$$

Thus, the second term of $P_{\mathrm{top}}(t)$ becomes

$$\frac{\mathcal{Z}_L-1}{\mathcal{Z}_L}\mathrm{Tr}[T_1\mathrm{Tr}_L(\rho_0)T_1^\dagger]=\frac{\mathcal{Z}_L-1}{\mathcal{Z}_L}\frac{\mathcal{Z}_C-1}{\mathcal{Z}_H\mathcal{Z}_C}\sum_{k=1}^{M}\left(\frac{1+(\mathcal{Z}_C-1)(\mathcal{Z}_H-1)}{\mathcal{Z}_H\mathcal{Z}_C}\right)^{k-1}\cos^2(gt)$$
$$+\frac{\mathcal{Z}_L-1}{\mathcal{Z}_L}\left(\frac{[1+(\mathcal{Z}_C-1)(\mathcal{Z}_H-1)]^M}{\mathcal{Z}_H^M\mathcal{Z}_C^M}+\sum_{k=1}^{M}\frac{[1+(\mathcal{Z}_C-1)(\mathcal{Z}_H-1)]^{k-1}(\mathcal{Z}_H-1)}{\mathcal{Z}_H^{k-1}\mathcal{Z}_C^{k-1}}\right), \tag{B15}$$

and the total top-level probability of the horizontal extension is given by

$$P_{\mathrm{top}}(t)=\sum_{k=1}^{M}\left(\frac{1+(\mathcal{Z}_C-1)(\mathcal{Z}_H-1)}{\mathcal{Z}_H\mathcal{Z}_C}\right)^{k-1}\left(\frac{1}{\mathcal{Z}_L}\frac{\mathcal{Z}_H-1}{\mathcal{Z}_H\mathcal{Z}_C}\sin^2(gt)+\frac{\mathcal{Z}_L-1}{\mathcal{Z}_L}\frac{\mathcal{Z}_C-1}{\mathcal{Z}_H\mathcal{Z}_C}\cos^2(gt)\right)$$
$$+\frac{\mathcal{Z}_L-1}{\mathcal{Z}_L}\left(\frac{[1+(\mathcal{Z}_C-1)(\mathcal{Z}_H-1)]^M}{\mathcal{Z}_H^M\mathcal{Z}_C^M}+\sum_{k=1}^{M}\frac{[1+(\mathcal{Z}_C-1)(\mathcal{Z}_H-1)]^{k-1}(\mathcal{Z}_H-1)}{\mathcal{Z}_H^{k-1}\mathcal{Z}_C^{k-1}}\right). \tag{B16}$$

Taking the limit $T_C\to 0$, the top-level probability becomes

$$P_{\mathrm{top}}(t)=\left(1-\frac{1}{\mathcal{Z}_H^M}\right)\sin^2(gt). \tag{B17}$$

APPENDIX C: DETAILS ON THE VERTICAL EXTENSION

Here, we present a detailed derivation of the top-level probability $P_{\mathrm{top}}(t)$ for the vertical extension of the clockwork. That is, we consider $M(d-1)$ two-qubit machines coupled to the d-level ladder, such that M machines nontrivially couple to each of the $d-1$ ladder transitions. For the purpose of this derivation, we consider the general case that both the cold and hot bath have finite temperatures, in particular, $T_C\geq 0$ and $T_H<\infty$, but we assume that $T_C<T_H$. To label the different machines, we denote the Hilbert space of the jth machine coupling to the ith ladder transition (i.e., the transition between the ladder levels $|i-1\rangle_L$ and $|i\rangle_L$) by M_j^i, where $i\in\{0,1,\ldots,d-1\}$ and $j\in\{1,2,\ldots,M\}$. Moreover, we denote the Hilbert space of the collection of all machines within the same column (see, e.g., the illustration in Fig. 3), i.e., the collection of jth machines for all ladder transitions, as $M_{(j)}:=\bigotimes_{i=1}^{d-1}M_j^i$. Following these conventions, we define the fully (horizontally and vertically) extended interaction Hamiltonian as

$$H_{\mathrm{int}}=\sum_{k=1}^{M}\bigotimes_{i=1}^{k-1}\mathbb{1}_{M_{(i)}}\otimes J_{M_{(k)}L}\bigotimes_{j=k+1}^{M}\Pi_{M_{(j)}}=\sum_{k=1}^{M}H_k. \tag{C1}$$

Here, the operator H_k acts nontrivially on the joint Hilbert space $M_{(k)}$ of the kth column and the ladder via the operator

$$J_{M_{(k)}L}:=ig\sum_{n=1}^{d-1}\sqrt{n(d-n)}(|n_{M_{(k)}},n_L\rangle$$
$$\times\langle n-1_{M_{(k)}},n-1_L|-\mathrm{H.c.}). \tag{C2}$$

However, the action of $J_{M_{(k)}L}$ is conditioned on the states of the machines corresponding to the columns $M_{(k+1)}$, $M_{(k+2)}$, …, $M_{(M)}$ through the projectors

$$\Pi_{M_{(k)}}:=\mathbb{1}_{M_{(k)}}-\sum_{n=0}^{N-1}|n_{M_{(k)}}\rangle\langle n_{M_{(k)}}|=\mathbb{1}_{M_{(k)}}-\bar{\Pi}_{M_{(k)}}, \tag{C3}$$

where $|0_{M_{(k)}}\rangle := \bigotimes_{j=1}^{d-1} |0_C 1_H\rangle_{M_k^j}$ and the states $|n_{M_{(k)}}\rangle$ for $n = 1, \ldots, d-1$ are defined as

$$|n_{M_{(k)}}\rangle := \bigotimes_{j=1}^{n} |1_C 0_H\rangle_{M_k^j} \bigotimes_{l=n+1}^{d-1} |0_C 1_H\rangle_{M_k^l}. \tag{C4}$$

The state $|n_{M_{(k)}}\rangle$ can be considered to be the nth excited state of the kth vertical group $\mathrm{M}_{(k)}$ in the sense that the first n machines M_k^j for $j = 1, \ldots, n$ are in the used state $|1_C 0_H\rangle_{M_k^j}$, whereas the remaining $d-n+1$ machines M_k^l for $l = n+1, \ldots, d-1$ are in the unused state $|0_C 1_H\rangle_{M_k^l}$. Similarly, the state $|0_{M_{(k)}}\rangle$ represents the corresponding "ground state."

Further note that $J_{M_{(k)}L}$ acts as an effective generator of rotations on the states $|n_L, n_{M_{(k)}}\rangle$ for $n = 0, \ldots, d-1$. To be more precise, $J_{M_{(k)}L}$ can be considered to be a spin-j representation [for $j = (d-1/2)$] of the generator of rotations around the y axis on the subspace $\mathcal{W}_{(k)} := \mathrm{span}(\{|n_{M_{(k)}}\rangle\}_{n=0,\ldots,d-1}) \subset M_{(k)}$ of the Hilbert space $M_{(k)}$ of the kth vertical group of machines. Let us further denote the orthogonal complement of $\mathcal{W}_{(k)}$ with respect to $M_{(k)}$ by $\mathcal{W}_{(k)}^{\perp}$, such that $M_{(k)} = \mathcal{W}_{(k)} \oplus \mathcal{W}_{(k)}^{\perp}$. We then observe that $\ker(J_{M_{(k)}L}) = \mathcal{W}_{(k)}^{\perp} \otimes \mathcal{H}_L$. Since $\Pi_{M_{(k)}}$ projects onto $\mathcal{W}_{(k)}^{\perp}$, we see that H_k has support only on the subspace $\mathrm{supp}(H_k) = \bigotimes_{i=1}^{k-1} M_{(i)} \otimes \mathcal{W}_{(k)} \otimes_{j=k+1}^{m} \mathcal{W}_{(j)}^{\perp} \otimes \mathcal{H}_L$ of the total Hilbert space of the ladder and all machines. Moreover, these subspaces are orthogonal for different values of k, i.e.,

$$H_k H_{k'} = 0 \quad \forall \; k \neq k'. \tag{C5}$$

As a consequence, we have $H_{\mathrm{int}}^q = (\sum_{k=1}^{M} H_k)^q = \sum_{k=1}^{M} H_k^q$, which we can use in the power expansion of $e^{-iH_{\mathrm{int}}t}$, that is,

$$\begin{aligned} e^{-iH_{\mathrm{int}}t} &= \sum_{q=0}^{\infty} \frac{(-it)^q}{q!} H_{\mathrm{int}}^q = \mathbb{1} + \sum_{q=1}^{\infty} \frac{(-it)^q}{q!} H_{\mathrm{int}}^q = \mathbb{1} + \sum_{k=1}^{M} \sum_{q=1}^{\infty} \frac{(-it)^q}{q!} H_k^q \\ &= \mathbb{1} + \sum_{k=1}^{M} \bigotimes_{i=1}^{k-1} \mathbb{1}_{M_{(i)}} \otimes \left(\sum_{q=1}^{\infty} \frac{(-it)^q}{q!} J_{M_{(k)}L}^q \right) \bigotimes_{j=k+1}^{M} \Pi_{M_{(j)}}, \end{aligned} \tag{C6}$$

where we isolate the leading-order term ($q = 0$) in the expansion and use the fact that the $\mathbb{1}_{M_{(i)}}$ and $\Pi_{M_{(j)}}$ are idempotent. Next, we observe that by definition $J_{M_{(k)}L}^q \bar{\Pi}_{M_{(k)}} = J_{M_{(k)}L}^q$ for all $q \geq 1$. We then define the operator $U_{M_{(k)}L}(t) := e^{-iJ_{M_{(k)}L}t}$ and write

$$U_{M_{(k)}L} \bar{\Pi}_{M_{(k)}} = e^{-iJ_{M_{(k)}L}t} \bar{\Pi}_{M_{(k)}} = \left(\mathbb{1}_{M_{(k)}L} + \sum_{q=1}^{\infty} \frac{(-it)^q}{q!} J_{M_{(k)}L}^q \right) \bar{\Pi}_{M_{(k)}} = \mathbb{1}_L \otimes \bar{\Pi}_{M_{(k)}} + \sum_{q=1}^{\infty} \frac{(-it)^q}{q!} J_{M_{(k)}L}^q. \tag{C7}$$

Inserting Eq. (C7) into Eq. (C6), we obtain

$$\begin{aligned} e^{-iH_{\mathrm{int}}t} &= \mathbb{1} - \mathbb{1}_L \otimes \sum_{k=1}^{M} \bigotimes_{i=1}^{k-1} \mathbb{1}_{M(i)} \overline{\Pi}_{M(k)} \otimes \sum_{j=k+1}^{M} \Pi_{M(j)} + \sum_{k=1}^{M} \bigotimes_{i=1}^{k-1} \mathbb{1}_{M(i)} \otimes U_{M(k)^L} \overline{\Pi}_{M(k)} \otimes \sum_{j=k+1}^{M} \Pi_{M(j)} \\ &= \mathbb{1} - \mathbb{1}_L \otimes \sum_{k=1}^{M} \tilde{\Pi}_{[k]} + \sum_{k=1}^{M} \tilde{U}_{[k]} = \overline{\mathbb{1}} + \sum_{k=1}^{M} \tilde{U}_{[k]}, \end{aligned} \tag{C8}$$

where we define $\overline{\mathbb{1}} := \mathbb{1} - \mathbb{1}_L \otimes \sum_{k=1}^{M} \tilde{\Pi}_{[k]}$. The projectors $\tilde{\Pi}_{[k]}$ and operators $\tilde{U}_{[k]}$ are defined, respectively, as

$$\tilde{\Pi}_{[k]} := \bigotimes_{i=1}^{k-1} \mathbb{1}_{M_{(i)}} \otimes \bar{\Pi}_{M_{(k)}} \bigotimes_{j=k+1}^{M} \Pi_{M_{(j)}}, \tag{C9}$$

$$\tilde{U}_{[k]} := \bigotimes_{i=1}^{k-1} \mathbb{1}_{M_{(i)}} \otimes U_{M_{(k)}L} \bar{\Pi}_{M_{(k)}} \bigotimes_{j=k+1}^{M} \Pi_{M_{(j)}}, \tag{C10}$$

such that $\tilde{U}_{[k]} \tilde{U}_{[k']} = 0$ and $\tilde{\Pi}_{[k]} \tilde{\Pi}_{[k']} = 0 \; \forall \; k \neq k'$, while $(\mathbb{1}_L \otimes \tilde{\Pi}_{[k']}) \tilde{U}_{[k]} = \tilde{U}_{[k]} (\mathbb{1}_L \otimes \tilde{\Pi}_{[k']}) = \delta_{kk'} \tilde{U}_{[k]}$ and $\overline{\mathbb{1}} \tilde{U}_{[k]} = \tilde{U}_{[k]} \overline{\mathbb{1}} = 0$.

With this result, we are now in a position to provide a compact expression of the top-level probability $P_{\mathrm{top}}(t)$, which takes the form

$$P_{\text{top}}(t) = \text{Tr}[|d-1\rangle\langle d-1|_L \rho] = \text{Tr}[{}_L\langle d-1|e^{-iH_{\text{int}}t}\rho_0 e^{iH_{\text{int}}t}|d-1\rangle_L]$$
$$= \text{Tr}\left[{}_L\langle d-1|\left(\bar{\mathbb{1}} + \sum_{k=1}^{M}\tilde{U}_{[k]}\right)\rho_0\left(\bar{\mathbb{1}} + \sum_{k=1}^{M}\tilde{U}_{[k]}^{\dagger}\right)|d-1\rangle_L\right]$$
$$= \text{Tr}[{}_L\langle d-1|\bar{\mathbb{1}}\rho_0\bar{\mathbb{1}}|d-1\rangle_L] + \sum_{k=1}^{M}\text{Tr}[{}_L\langle d-1|\tilde{U}_{[k]}\rho_0\tilde{U}_{[k]}^{\dagger}|d-1\rangle_L], \tag{C11}$$

where we use the assumption that the initial state ρ_0 is diagonal with respect to the joint eigenbasis of the orthogonal projectors $\tilde{\Pi}_{[k]}$, which is the case here because the ladder and all machines qubits are initially thermal with respect to either the cold or hot bath.

For the first term in $P_{\text{top}}(t)$, we then have

$$\text{Tr}[{}_L\langle d-1|\bar{\mathbb{1}}\rho_0\bar{\mathbb{1}}|d-1\rangle_L] = \text{Tr}[{}_L\langle d-1|\bar{\mathbb{1}}\rho_0|d-1\rangle_L] = \text{Tr}\left[{}_L\langle d-1|\left(\mathbb{1} - \mathbb{1}_L \otimes \sum_{k=1}^{M}\tilde{\Pi}_{[k]}\right)\rho_0|d-1\rangle_L\right]$$
$$= {}_L\langle d-1|\tau_L(\beta_C)|d-1\rangle_L\left(1 - \sum_{k=1}^{M}\text{Tr}[\tilde{\Pi}_{[k]}\text{Tr}_L(\rho_0)]\right). \tag{C12}$$

Here, we further have

$$\text{Tr}[\tilde{\Pi}_{[k]}\text{Tr}_L(\rho_0)] = \text{Tr}\left[\left(\bigotimes_{i=1}^{k-1}\mathbb{1}_{M_{(i)}} \otimes \bar{\Pi}_{M_{(k)}} \bigotimes_{j=k+1}^{M}\Pi_{M_{(j)}}\right)\bigotimes_{l=1}^{M}\tau_{M_{(l)}}\right] = \text{Tr}[\bar{\Pi}_{M_{(k)}}\tau_{M_{(k)}}]\prod_{j=k+1}^{M}\text{Tr}[\Pi_{M_{(j)}}\tau_{M_{(j)}}]$$
$$= \text{Tr}[\bar{\Pi}_{M_{(k)}}\tau_{M_{(k)}}]\prod_{j=k+1}^{M}(1 - \text{Tr}[\bar{\Pi}_{M_{(j)}}\tau_{M_{(j)}}]), \tag{C13}$$

where we can use Eq. (C3) to calculate

$$\text{Tr}[\bar{\Pi}_{M_{(j)}}\tau_{M_{(j)}}] = \sum_{n=0}^{N-1}\langle n_{M_{(j)}}|\tau_{M_{(j)}}|n_{M_{(j)}}\rangle = \sum_{n=0}^{N-1}\prod_{i=1}^{n}\langle 1_C 0_H|\tau_{M_j^i}|1_C 0_H\rangle\prod_{l=n+1}^{N-1}\langle 0_C 1_H|\tau_{M_j^l}|0_C 1_H\rangle$$
$$= \sum_{n=0}^{N-1}\langle 1|\tau_C|1\rangle^n\langle 0|\tau_H|0\rangle^n\langle 0|\tau_C|0\rangle^{d-n-1}\langle 1|\tau_H|1\rangle^{d-n-1} = \frac{1}{\mathcal{Z}_C^{d-1}\mathcal{Z}_H^{d-1}}\sum_{n=0}^{d-1}(\mathcal{Z}_C-1)^n(\mathcal{Z}_H-1)^{d-n-1}$$
$$= \frac{(\mathcal{Z}_H-1)^d - (\mathcal{Z}_C-1)^d}{\mathcal{Z}_H^{d-1}\mathcal{Z}_C^{d-1}(\mathcal{Z}_H - \mathcal{Z}_C)}. \tag{C14}$$

Inserting Eqs. (C14) and (C13) into Eq. (C12) and evaluating the sum over k, we obtain

$$\text{Tr}[{}_L\langle d-1|\bar{\mathbb{1}}\rho_0\bar{\mathbb{1}}|d-1\rangle_L] = {}_L\langle d-1|\tau_L(\beta_C)|d-1\rangle_L\left(1 - \frac{(\mathcal{Z}_H-1)^d - (\mathcal{Z}_C-1)^d}{\mathcal{Z}_H^{d-1}\mathcal{Z}_C^{d-1}(\mathcal{Z}_H - \mathcal{Z}_C)}\right)^M. \tag{C15}$$

Turning to the second term of $P_{\text{top}}(t)$ in Eq. (C11), we express the individual terms in the sum over k as

$$\text{Tr}[{}_L\langle d-1|\tilde{U}_{[k]}\rho_0\tilde{U}_{[k]}^{\dagger}|d-1\rangle_L] = \text{Tr}[{}_L\langle d-1|\bar{\Pi}_{M_{(k)}}U_{M_{(k)}L}\tau_L \otimes \tau_{M_{(k)}}U_{M_{(k)}L}^{\dagger}\bar{\Pi}_{M_{(k)}}|d-1\rangle_L]\prod_{j=k+1}^{M}\text{Tr}[\Pi_{M_{(j)}}\tau_{M_{(j)}}]. \tag{C16}$$

Then, we note that we can write

$$\bar{\Pi}_{M_{(k)}}U_{M_{(k)}L} = \bar{\Pi}_{M_{(k)}} \otimes \mathbb{1}_L + \sum_{q=1}^{\infty}\frac{(-it)^q}{q!}J_{M_{(k)}L}^q = \sum_{m,n=0}^{N-1}|m_{M_{(k)}}, n_L\rangle\langle m_{M_{(k)}}, n_L| + \sum_{q=1}^{\infty}\frac{(-it)^q}{q!}J_{M_{(k)}L}^q$$
$$= \sum_{\substack{m,n=0\\m\neq n}}^{d-1}|m_{M_{(k)}}, n_L\rangle\langle m_{M_{(k)}}, n_L| + \sum_{n=0}^{d-1}|n_{M_{(k)}}, n_L\rangle\langle n_{M_{(k)}}, n_L| + \sum_{q=1}^{\infty}\frac{(-it)^q}{q!}J_{M_{(k)}L}^q$$
$$= \sum_{\substack{m,n=0\\m\neq n}}^{N-1}|m_{M_{(k)}}, n_L\rangle\langle m_{M_{(k)}}, n_L| + \sum_{n=0}^{d-1}|n_{M_{(k)}}, n_L\rangle\langle n_{M_{(k)}}, n_L|e^{-iJ_{M_{(k)}L}t}, \tag{C17}$$

where we separate the terms corresponding to projectors onto the kernel and support of the operator $J_{M_{(k)}L}$ in the second step. With this result, we can simplify the first factor appearing on the right-hand side of Eq. (C16) to

$$\begin{aligned}
&\mathrm{Tr}[{}_L\langle d-1|\bar{\Pi}_{M_{(k)}}U_{M_{(k)}L}\tau_L\otimes\tau_{M_{(k)}}U^{\dagger}_{M_{(k)}L}\bar{\Pi}_{M_{(k)}}|d-1\rangle_L]\\
&=\sum_{n=0}^{d-2}\langle n_{M_{(k)}},d-1_L|\tau_L\otimes\tau_{M_{(k)}}|n_{M_{(k)}},d-1_L\rangle+\mathrm{Tr}\Bigg[{}_L\langle d-1|\Bigg(\sum_{n=0}^{d-1}|n_{M_{(k)}},n_L\rangle\langle n_{M_{(k)}},n_L|e^{-iJ_{M_{(k)}L}t}\Bigg)(\tau_L\otimes\tau_{M_{(k)}})\\
&\quad\times\Bigg(e^{iJ_{M_{(k)}L}t}\sum_{n'=0}^{d-1}|n'_{M_{(k)}},n'_L\rangle\langle n'_{M_{(k)}},n'_L|\Bigg)|d-1\rangle_L\Bigg]\\
&={}_L\langle d-1|\tau_L(\beta_C)|d-1\rangle_L\sum_{n=0}^{d-2}\langle n_{M_{(k)}}|\tau_{M_{(k)}}|n_{M_{(k)}}\rangle+\sum_{n=0}^{d-1}\langle n_{M_{(k)}},n_L|\tau_{M_{(k)}}\otimes\tau_L|n_{M_{(k)}},n_L\rangle|\langle N-1_{M_{(k)}},d-1_L|e^{-iJ_{M_{(k)}L}t}|n_{M_{(k)}},n_L\rangle|^2\\
&={}_L\langle d-1|\tau_L(\beta_C)|d-1\rangle_L\sum_{n=0}^{d-2}\frac{(\mathcal{Z}_C-1)^n(\mathcal{Z}_H-1)^{d-n-1}}{\mathcal{Z}_C^{d-1}\mathcal{Z}_H^{d-1}}\\
&\quad+\sum_{n=0}^{d-1}{}_L\langle n|\tau_L(\beta_C)|n\rangle_L\frac{(\mathcal{Z}_C-1)^n(\mathcal{Z}_H-1)^{d-n-1}}{\mathcal{Z}_C^{d-1}\mathcal{Z}_H^{d-1}}|\langle d-1_{M_{(k)}},d-1_L|e^{-iJ_{M_{(k)}L}t}|n_{M_{(k)}},n_L\rangle|^2.
\end{aligned}\tag{C18}$$

Reinserting the first term appearing in the last step of Eq. (C18) back into Eq. (C16) and evaluating the sum over k in Eq. (C11), we obtain another [i.e., in addition to that in Eq. (C15)] time-independent contribution to the top-level probability, given by

$$\begin{aligned}
&{}_L\langle d-1|\tau_L(\beta_C)|d-1\rangle_L\sum_{k=1}^{M}\Bigg(1-\frac{(\mathcal{Z}_H-1)^d-(\mathcal{Z}_C-1)^d}{\mathcal{Z}_H^{d-1}\mathcal{Z}_C^{d-1}(\mathcal{Z}_H-\mathcal{Z}_C)}\Bigg)^{M-k}\sum_{n=0}^{d-2}\frac{(\mathcal{Z}_C-1)^n(\mathcal{Z}_H-1)^{d-n-1}}{\mathcal{Z}_C^{d-1}\mathcal{Z}_H^{d-1}}\\
&\quad={}_L\langle d-1|\tau_L(\beta_C)|d-1\rangle_L\Bigg(1-\frac{(\mathcal{Z}_C-1)^{d-1}(\mathcal{Z}_H-\mathcal{Z}_C)}{(\mathcal{Z}_H-1)^d-(\mathcal{Z}_C-1)^d}\Bigg)\Bigg[1-\Bigg(1-\frac{(\mathcal{Z}_H-1)^d-(\mathcal{Z}_C-1)^d}{\mathcal{Z}_H^{d-1}\mathcal{Z}_C^{d-1}(\mathcal{Z}_H-\mathcal{Z}_C)}\Bigg)^M\Bigg].
\end{aligned}\tag{C19}$$

For the second term appearing in the last step of Eq. (C18), we note that, since $J_{M_{(k)}L}$ corresponds to the spin-j representation (with $j=d-1/2$) of the generator of rotations around the y axis on the subspace spanned by the vectors $|n_{M_{(k)}},n_L\rangle$ for $n=0,1,\ldots,d-1$, the matrix elements $\langle d-1_{M_{(k)}},d-1_L|e^{-iJ_{M_{(k)}L}t}|n_{M_{(k)}},n_L\rangle$ coincide with the elements of the Wigner (small) d matrix $d^j_{\mu,m}(\beta):=\langle j,\mu|e^{-i\beta J_y}|j,m\rangle$ for $\mu=j$, $m=n-j$, and $\beta=2gt$; see, e.g., Ref. [35] or [36]. In particular, Eq. (B7) in Ref. [35] lets us write

$$|\langle d-1_{M_{(k)}},d-1_L|e^{-iJ_{M_{(k)}L}t}|n_{M_{(k)}},n_L\rangle|^2=\binom{d-1}{n}\cos^{2n}(gt)\sin^{2(d-n-1)}(gt).\tag{C20}$$

The prefactors of these sinusoidal contributions are then obtained by combining the second term in the last step of Eq. (C18) with Eq. (C16) and evaluating the sum over k in Eq. (C11), which yields

$$\begin{aligned}
&{}_L\langle n|\tau_L(\beta_C)|n\rangle_L\frac{(\mathcal{Z}_C-1)^n(\mathcal{Z}_H-1)^{d-n-1}}{\mathcal{Z}_C^{d-1}\mathcal{Z}_H^{d-1}}\sum_{k=1}^{M}\Bigg(1-\frac{(\mathcal{Z}_H-1)^d-(\mathcal{Z}_C-1)^d}{\mathcal{Z}_H^{d-1}\mathcal{Z}_C^{d-1}(\mathcal{Z}_H-\mathcal{Z}_C)}\Bigg)^{M-k}\\
&\quad={}_L\langle n|\tau_L(\beta_C)|n\rangle_L\frac{(\mathcal{Z}_C-1)^n(\mathcal{Z}_H-1)^{d-n-1}(\mathcal{Z}_H-\mathcal{Z}_C)}{(\mathcal{Z}_H-1)^d-(\mathcal{Z}_C-1)^d}\Bigg[1-\Bigg(1-\frac{(\mathcal{Z}_H-1)^d-(\mathcal{Z}_C-1)^d}{\mathcal{Z}_H^{d-1}\mathcal{Z}_C^{d-1}(\mathcal{Z}_H-\mathcal{Z}_C)}\Bigg)^M\Bigg].
\end{aligned}\tag{C21}$$

Finally, we can collect Eqs. (C20) and (C21) and combine them with the time-independent terms in $P_{\mathrm{top}}(t)$ to arrive at

$$P_{\mathrm{top}}(t) = \sum_{n=0}^{d-2} {}_L\langle n|\tau_L(\beta_C)|n\rangle_L (\mathcal{Z}_C-1)^n (\mathcal{Z}_H-1)^{d-n-1} f(M,d,\beta_C,\beta_H) \binom{d-1}{n} \cos^{2n}(gt)\sin^{2(d-n-1)}(gt)$$
$$+ {}_L\langle d-1|\tau_L(\beta_C)|d-1\rangle_L \left\{1-[1-\cos^{2(d-1)}(gt)](\mathcal{Z}_C-1)^{d-1} f(M,d,\beta_C,\beta_H)\right\}, \tag{C22}$$

where the coefficient $f(M,d,\beta_C,\beta_H)$ is given by

$$f(M,d,\beta_C,\beta_H) = \frac{\mathcal{Z}_H-\mathcal{Z}_C}{(\mathcal{Z}_H-1)^d-(\mathcal{Z}_C-1)^d}\left[1-\left(1-\frac{(\mathcal{Z}_H-1)^d-(\mathcal{Z}_C-1)^d}{\mathcal{Z}_H^{d-1}\mathcal{Z}_C^{d-1}(\mathcal{Z}_H-\mathcal{Z}_C)}\right)^M\right]. \tag{C23}$$

The expression in Eq. (C22) holds for arbitrary temperatures T_C and $T_H > T_C$ and includes the desired term proportional to $\sin^{2(d-1)}(gt)$ in the sum for $n=0$. In particular, this term is the only term in $P_{\mathrm{top}}(t)$ that remains when taking the limit $T_C \to 0$, in which case $\mathcal{Z}_C \to 1$, ${}_L\langle d-1|\tau_L(\beta_C)|d-1\rangle_L \to 0$, and ${}_L\langle 0|\tau_L(\beta_C)|0\rangle_L = 1$, and we have

$$\lim_{T_C\to 0} P_{\mathrm{top}}(t) = \left\{1-\left[1-\left(\frac{\mathcal{Z}_H-1}{\mathcal{Z}_H}\right)^{d-1}\right]^M\right\}\sin^{2(d-1)}(gt), \tag{C24}$$

as stated in Eq. (15) of the main text.

To see that small deviations from the ideal case where $T_C = 0$ still allow for $P_{\mathrm{top}}(t)$ to be close to the corresponding value of the ideal case, i.e., to show the stability of our approach to ATPC, we analyze the behavior of $P_{\mathrm{top}}(t)$ in the limits $M \to \infty$ and $d \to \infty$ at finite temperatures. To this end, we first inspect Eq. (C23) and note that the term that is potentiated by M is smaller than 1. To see this result, we first write

$$\frac{(\mathcal{Z}_H-1)^d-(\mathcal{Z}_C-1)^d}{\mathcal{Z}_H^{d-1}\mathcal{Z}_C^{d-1}(\mathcal{Z}_H-\mathcal{Z}_C)} = \frac{(\mathcal{Z}_H-1)^d-(\mathcal{Z}_C-1)^d}{\mathcal{Z}_H^{d}\mathcal{Z}_C^{d}}\frac{\mathcal{Z}_H\mathcal{Z}_C}{\mathcal{Z}_H-\mathcal{Z}_C} = \frac{x^d-y^d}{x-y}, \tag{C25}$$

where we define $x := (\mathcal{Z}_H - 1/\mathcal{Z}_H\mathcal{Z}_C)$ and $y := (\mathcal{Z}_C - 1/\mathcal{Z}_H\mathcal{Z}_C)$ with the property $0 \le y < x \le \frac{1}{2}$. The expression on the right-hand side of Eq. (C25) is smaller than or equal to 1 if $x - x^d \ge y - y^d$, which is the case if $x - x^d$ is monotonically increasing on the interval $[0, \frac{1}{2}]$. Inspecting the derivative, we have $(\partial/\partial x)(x-x^d) = 1 - dx^{d-1} \ge 0$ since $dx^{d-1} \le d/2^{d-1} \le 1$ for $d \ge 2$. Consequently, we have $(x^d - y^d/x - y) \le 1$ and $\lim_{M\to\infty}[1-(x^d-y^d/x-y)]^M = 0$. Therefore, we see that

$$\lim_{M\to\infty} f(M,d,\beta_C,\beta_H) = \frac{\mathcal{Z}_H-\mathcal{Z}_C}{(\mathcal{Z}_H-1)^d-(\mathcal{Z}_C-1)^d}. \tag{C26}$$

Since we know that $P_{\mathrm{top}}(t)$ must lie in [0, 1], showing that the first term of Eq. (C22) ($n=0$) remains close to 1 when M and N go to infinity is sufficient to show that our approach is stable with respect to deviations from $T_H \to \infty$ and $T_C \to 0$, i.e.,

$$\lim_{M\to\infty}\lim_{d\to\infty} {}_L\langle 0|\tau_L(\beta_C)|0\rangle_L (\mathcal{Z}_H-1)^{d-1} f(M,d,\beta_C,\beta_H)\binom{d-1}{0}\sin^{2(d-1)}(gt)$$
$$= \lim_{d\to\infty}\frac{1}{\mathcal{Z}_L}(\mathcal{Z}_H-1)^{d-1}\frac{\mathcal{Z}_H-\mathcal{Z}_C}{(\mathcal{Z}_H-1)^d-(\mathcal{Z}_C-1)^d}\sin^{2(d-1)}(gt) = \lim_{d\to\infty}\frac{1}{\mathcal{Z}_L}\frac{\mathcal{Z}_H-\mathcal{Z}_C}{\mathcal{Z}_H-1}\frac{1}{1-(\frac{\mathcal{Z}_C-1}{\mathcal{Z}_H-1})^d}\sin^{2(d-1)}(gt)$$
$$= \begin{cases} \frac{\mathcal{Z}_H-\mathcal{Z}_C}{\mathcal{Z}_L(\mathcal{Z}_H-1)}, & \text{if } t=\frac{\pi}{2g}, \\ 0, & \text{otherwise.}\end{cases} \tag{C27}$$

The value of the expression in Eq. (C27) for $t = (\pi/2g)$ can further be written as

$$\frac{\mathcal{Z}_H-\mathcal{Z}_C}{\mathcal{Z}_L(\mathcal{Z}_H-1)} = \frac{1+e^{-\beta_H E_H}-(1+e^{-\beta_C E_C})}{e^{-\beta_H E_H}\sum_{n=0}^{d-1} e^{-n\beta_C(E_H-E_C)}} = \frac{1-e^{\beta_H E_H-\beta_C E_C}}{1+\sum_{n=1}^{\infty} e^{-n\beta_C(E_H-E_C)}}$$
$$= 1 - e^{-\beta_C(E_H-E_C)} - e^{-(\beta_C E_C-\beta_H E_H)} + e^{-E_H(\beta_C-\beta_H)}. \tag{C28}$$

Recalling that $E_H > E_C$, we can see that the quantity in Eq. (C28) remains close to 1 for finite temperatures when $\beta_C \gg \beta_H$ such that $\beta_H E_H < \beta_C E_C$ and when $k_B T_C \ll (E_H - E_C)$, which are both in keeping with the assumptions made in the main text.

APPENDIX D: TICK PROBABILITY DENSITY

In this Appendix, we show how our derivation of the tick probability results in an exponential decay. The derivation should not be understood as a new result but rather as a reminder and clarification. Treating the decay of the top level as a random event, we can approximate its probability of occurring in the time interval Δt by

$$\Delta P = \Gamma \Delta t, \tag{D1}$$

where Γ is given in terms of probability per unit time. In our case here, we have that Γ corresponds to the top-level population times the constant c, i.e., $\Gamma(t) = P_{\text{top}}(t)c$. Let us denote the cumulative probability that no decay occurs until time t as $P(0,t)$. We can then approximate the probability that no event occurs until $t + \Delta t$ as the probability of no event happening until t times the probability that no event happens in the interval Δt, i.e.,

$$P(0, t + \Delta t) = P(0,t)[1 - \Gamma(t)\Delta t], \tag{D2}$$

which leads to

$$\frac{P(0, t + \Delta t) - P(0,t)}{\Delta t} = -\Gamma(t)P(0,t). \tag{D3}$$

If we further let $\Delta t \to dt$, we get that

$$\frac{dP(0,t)}{dt} = -\Gamma(t)P(0,t) = -cP_{\text{top}}(t)P(0,t) \tag{D4}$$

and, consequently,

$$P(0,t) = e^{-c\int_0^t P_{\text{top}}(t')dt'}. \tag{D5}$$

Given this expression for the cumulative probability that no event occurs until time t, we can proceed to calculate the probability density of a decay event occurring between time t and $t + dt$. To do so, we differentiate the cumulative probability of having had a decay at time t with respect to t, i.e., $d[1 - P(0,t)]/dt$, which results in

$$P_{\text{tick}}(t) = cP_{\text{top}}(t)e^{-c\int P_{\text{top}}(t')dt'}. \tag{D6}$$

APPENDIX E: NUMERICAL CALCULATION OF ACCURACY AND RESOLUTION

In order to execute the numerical calculations of the resolution and the accuracy efficiently, we need to simplify the necessary integrals [defined in Eqs. (18) and (17)]. In this Appendix, we present details on our approach to this problem. For simplicity, we showcase the calculations for $T_C \to 0$ and $T_H \to \infty$.

Assessing resolution and accuracy for a given set of parameters d, M, c, and g breaks down to calculating the first and second moments of the tick distribution, i.e.,

$$\bar{t} = \int_0^\infty tP_{\text{tick}}(t)dt, \qquad \overline{t^2} = \int_0^\infty t^2 P_{\text{tick}}(t)dt, \tag{E1}$$

where

$$P_{\text{tick}}(t) = c\left\{1 - \left[1 - \left(\frac{\mathcal{Z}_H - 1}{\mathcal{Z}_H}\right)^{d-1}\right]^M\right\}\sin^{2(d-1)}(gt)\exp\left[-c\left\{1 - \left[1 - \left(\frac{\mathcal{Z}_H - 1}{\mathcal{Z}_H}\right)^{d-1}\right]^M\right\}\int_0^t dt'\sin^{2(d-1)}(gt')\right]. \tag{E2}$$

In order to simplify the cumbersome expressions, we use the following notation. We denote the kth moment as

$$I_k = c\left\{1 - \left[1 - \left(\frac{\mathcal{Z}_H - 1}{\mathcal{Z}_H}\right)^{d-1}\right]^M\right\}\int_0^\infty dt t^k \sin^{2(d-1)}(gt)\exp\left[-c\left\{1 - \left[1 - \left(\frac{\mathcal{Z}_H - 1}{\mathcal{Z}_H}\right)^{d-1}\right]^M\right\}\int_0^t dt'\sin^{2(d-1)}(gt')\right]. \tag{E3}$$

The *effective coupling* is defined as

$$C_M := c\left\{1 - \left[1 - \left(\frac{\mathcal{Z}_H - 1}{\mathcal{Z}_H}\right)^{d-1}\right]^M\right\}. \tag{E4}$$

Furthermore, let

$$f(t) = C_M \int_0^\infty dt' \sin^{2(d-1)}(gt'). \tag{E5}$$

This equation leads to a much simpler form for the different moments:

$$I_k = C_M \int_0^\infty dt t^k \sin^{2d}(gt)e^{-f(t)}. \tag{E6}$$

We can solve the integral in the term $f(t)$ with a solution introduced by Wiener [37]:

$$\int_0^x dx' \sin^{2(d-1)}(x') = \frac{1}{4^{d-1}}\left[\binom{2(d-1)}{(d-1)}x + \sum_{p=1}^{d-1}\frac{(-1)^p}{p}\binom{2(d-1)}{d-1-p}\sin(2px)\right], \tag{E7}$$

where $\binom{n}{m} = \frac{n!}{m!(n-m)!}$ is the binomial coefficient. Employing the solution, we get

$$f(t) = \frac{C_M}{4^{d-1}}\left\{\binom{2(d-1)}{d-1}t + \frac{1}{g}\sum_{p=1}^{d-1}\binom{2(d-1)}{d-1-p}\sin(2pgt)\right\} \tag{E8}$$

and, thus,

$$I_k = C_M\int_0^\infty dt t^k \sin^{2(d-1)}(gt)\exp\left[-\frac{C_M}{4^{d-1}}\binom{2(d-1)}{d-1}t\right] \times \exp\left[-\frac{C_M}{4^{d-1}g}\sum_{p=1}^{d}\frac{(-1)^p}{p}\binom{2(d-1)}{d-1-p}\sin(2pgt)\right]. \tag{E9}$$

By introducing a "cycle" counting variable $q = \lfloor gt/\pi \rfloor \in_0$ and its residue $\Theta_q = qt - q\pi$, $\Theta_q \epsilon [0,\pi)$, i.e., substituting with $t = (q\pi + \Theta_q/g)$, we arrive at

$$I_k = \frac{C_M}{g}\sum_{q=0}^{\infty}\int_0^\pi d\Theta_q\left(\frac{q\pi+\Theta_q}{g}\right)^k \sin^{2(d-1)}(\Theta_q) \times \exp\left[-\frac{C_M}{4^{d-1}g}\binom{2(d-1)}{d-1}(q\pi+\Theta_q)\right] \tag{E10}$$

$$\times \exp\left[-\frac{C_M}{4^{d-1}g}\sum_{p=1}^{d-1}\frac{(-1)^p}{p}\binom{2(d-1)}{d-1-p}\sin(2p\Theta_q)\right], \tag{E11}$$

where we use that $\sin^{2(d-1)}(x \pm n\pi) = \sin^{2(d-1)}(x)$ for $n\epsilon\mathbb{Z}$ as well as $\sin[2(x \pm n\pi)] = \sin(2x)$.

We are interested only in explicitly calculating the cases $k = 1$ and $k = 2$. Note that k appears only in the term $(q\pi + \Theta_q/g)^k$ and that $(q\pi + \Theta_q/g)^2 = (1/g^2)[(q\pi)^2 + 2q\pi\Theta_q + \Theta_q^2]$, which allows us to define the function

$$E(\Theta_q) = \exp\left[-\frac{C_M}{4^{d-1}g}\binom{2(d-1)}{d-1}\Theta_q\right] \times \exp\left[-\frac{C_M}{4^{d-1}g}\sum_{p=1}^{d-1}\frac{(-1)^p}{p}\binom{2(d-1)}{d-1-p}\sin(2p\Theta_q)\right], \tag{E12}$$

such that

$$I_k = \frac{C_M}{g}\sum_{q=0}^{\infty}\exp\left[-\frac{C_M}{4^{d-1}g}\binom{2(d-1)}{d-1}q\pi\right] \times \int_0^\pi d\Theta_q\left(\frac{q\pi+\Theta_q}{g}\right)^k \sin^{2(d-1)}(\Theta_q)E(\Theta_q). \tag{E13}$$

As a last step, we observe that $E(\Theta_q)$ does not depend on q directly but only through Θ_q, so the only direct dependence on q in the integral comes from the term $(q\pi + \Theta_q/g)^k$. This result leads us to define

$$\tilde{I}_j \coloneqq \int_0^\pi d\Theta_q \Theta_q^j \sin^{2(d-1)}(\Theta_q)E(\Theta_q), \tag{E14}$$

such that we can write the desired first and second moments of the tick distribution as

$$I_1 = \frac{C_M}{g^2}\sum_{q=0}^{\infty}\exp\left[-\frac{C_M}{4^{d-1}g}\binom{2(d-1)}{d-1}q\pi\right]\{q\pi\tilde{I}_0 + \tilde{I}_1\} \tag{E15}$$

and

$$I_2 = \frac{C_M}{g^3}\sum_{q=0}^{\infty}\exp\left[-\frac{C_M}{4^{d-1}g}\binom{2(d-1)}{d-1}q\pi\right]\{(q\pi)^2\tilde{I}_0 + 2q\pi\tilde{I}_1 + \tilde{I}_2\}, \tag{E16}$$

respectively. In this way, only $\tilde{I}_j$ needs to be calculated numerically for $j = \{0,1,2\}$, which decreases the effective computational costs enormously.

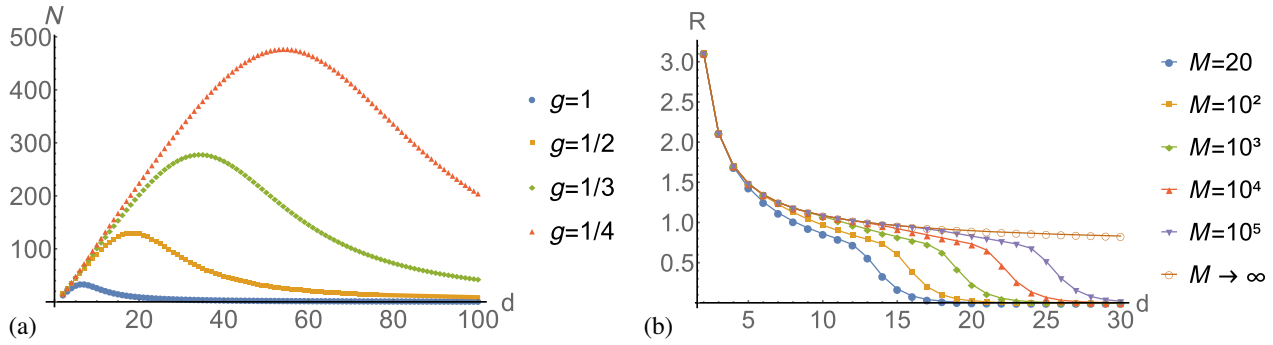

FIG. 7. (a) The accuracy is shown as a function of d for different values of g (and fixed $c = 10\ \mathrm{s}^{-1}$), where $[g] = E_C$. The maximally achievable accuracy increases with decreasing g. (b) The resolution is shown as a function d (at fixed $g = 1E_C$ and $c = 1000\ \mathrm{s}^{-1}$). The lines with solid dots show cases of finite M. The line with orange circles illustrates the behavior for the case $M \to \infty$, which provides an upper bound to the cases with finite M. For finite M, increasing d reduces the top-level population, eventually becoming so small that the decay event is considerably more likely to skip the first peak of $P_{\mathrm{top}}(t)$. This result leads to an additional reduction in resolution, initiating a drop of the resolution eventually approaching 0 (with d).

APPENDIX F: HOW THE SHARPNESS OF $P_{\mathrm{top}}(t)$ INFLUENCES ACCURACY AND RESOLUTION

The aim of this Appendix is to give further insight about the behavior of clocks with changing ladder dimension d as well as changing coupling constant g, in particular, with respect to Figs. 4–6 in Sec. IV.

First, let us discuss the relationship between accuracy and sharpness of $P_{\mathrm{top}}(t)$. The intuition is that clockworks that are capable of producing a very sharp temporal probability distribution should have the potential to give rise to highly accurate clocks, given a suitable irreversible process for the tick production. Since the maximal amplitude of $P_{\mathrm{top}}(t)$ is given by $1 - [1 - (\mathcal{Z}_H - 1/\mathcal{Z}_H)^{d-1}]^M$ (for $T_C = 0$), increasing M leads to an amplitude of $P_{\mathrm{top}}(t)$ that approaches 1 very quickly. Assuming that M is chosen large enough so that the maximal amplitude is within a desired distance to the value 1, the only parameter left that influences the sharpness of the probability distribution is the ladder dimension d. We can therefore use d as a proxy for sharpness. We then proceed by numerically calculating the accuracy in this situation for given values of c and g. The results are shown in Fig. 7(a) and indicate that the accuracy grows linearly with d. In this regime, the sharpness therefore determines the accuracy up to a constant factor. However, one should note that this linear relationship holds only in a regime where the decay process happens fast enough, i.e., assuming a sufficiently large value of c (or small enough value of g). If $P_{\mathrm{top}}(t)$ is too sharp compared to the timescale of the decay process, increasing the ladder dimension leads to a reduction of the accuracy [as seen in Figs. 7(a) and 4]. This result implies that for a given combination of c and g there are certain choices of d that lead to suboptimal clocks. Considering the resolution as a function of d [Fig. 7(b)] in the limit $M \to \infty$, we do not observe an optimal configuration. The resolution simply decreases with increasing d, indicating a trade-off relation between accuracy and resolution in the regime where the accuracy increases linearly with d. Thus, plotting accuracy over resolution reveals the trade-off relation depicted in Fig. 5. However, considering finite M, the resolution reaches a point at which it starts dropping to zero quickly. The reason for this drop can again be found in the amplitude of $P_{\mathrm{top}}(t)$, which goes to zero for large enough d and fixed M. Thus, not only the accuracy (see Fig. 4) but also the resolution is bounded from above by the corresponding resolution obtained for $M \to \infty$. There, c and g determine this upper bound.

Furthermore, considering only the cases where $M \to \infty$, Fig. 8 shows that increasing g at fixed c shifts the point of maximal accuracy to the right, i.e., toward higher resolutions. Thus, the value of g determines the lowest resolution at which (optimal) clocks can operate, i.e., the *maximal cycle time*. However, this increase in resolution comes at the cost of accuracy, as increasing g reduces the maximally reachable accuracy.

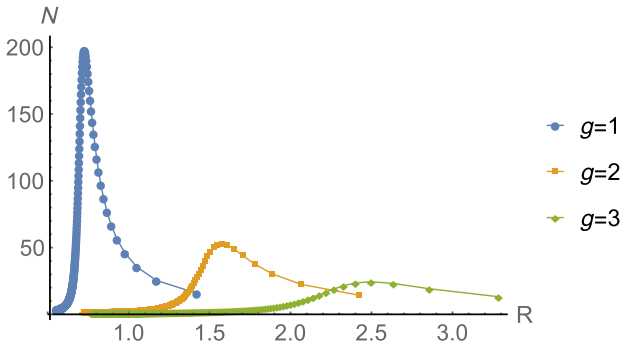

FIG. 8. The trade-off between clock accuracy N and resolution R for clockworks of various coupling constants g, where $[g] = E_C$ and R is increased by decreasing d. Here, we consider only cases of $M \to \infty$ and $c = 25\ \mathrm{s}^{-1}$. Increasing g shifts the peak toward the right, i.e., to higher resolutions, while decreasing the maximum accuracy.

[1] W. Pauli, *Die Allgemeinen Prinzipien der Wellenmechanik* (Springer, Berlin, 1958), https://doi.org/10.1007/978-3-642-61287-9.

[2] J. C. Garrison and J. Wong, *Canonically Conjugate Pairs, Uncertainty Relations, and Phase Operators*, J. Math. Phys. (N.Y.) **11**, 2242 (1970).

[3] S. L. Braunstein, C. M. Caves, and G. J. Milburn, *Generalized Uncertainty Relations: Theory, Examples, and Lorentz Invariance*, Ann. Phys. (N.Y.) **247**, 135 (1996).

[4] G. J. Milburn and T. J. Milburn, *A Quantum Optomechanical Mach Clock*, arXiv:1708.02369.

[5] P. Erker, *The Quantum Hourglass—Approaching Time Measurement with Quantum Information Theory*, Master's Thesis, ETH Zurich, 2014, https://doi.org/10.3929/ethz-a-010514644.

[6] A. C. Barato and U. Seifert, *Cost and Precision of Brownian Clocks*, Phys. Rev. X **6**, 041053 (2016).

[7] S. Khandelwal, M. P. E. Lock, and M. P. Woods, *Universal Quantum Modifications to General Relativistic Time Dilation in Delocalised Clocks*, Quantum **4**, 309 (2020).

[8] A. R. H. Smith and M. Ahmadi, *Quantum Clocks Observe Classical and Quantum Time Dilation*, Nat. Commun. **11**, 5360 (2020).

[9] A. R. H. Smith and M. Ahmadi, *Quantizing Time: Interacting Clocks and Systems*, Quantum **3**, 160 (2019).

[10] G. J. Milburn, *The Thermodynamics of Clocks*, Contemp. Phys. **60**, 69 (2020).

[11] P. Erker, M. T. Mitchison, R. Silva, M. P. Woods, N. Brunner, and M. Huber, *Autonomous Quantum Clocks: Does Thermodynamics Limit Our Ability to Measure Time?*, Phys. Rev. X **7**, 031022 (2017).

[12] A. N. Pearson, Y. Guryanova, P. Erker, E. A. Laird, G. A. D. Briggs, M. Huber, and N. Ares, *Measuring the Thermodynamic Cost of Timekeeping*, arXiv:2006.08670.

[13] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, *The Role of Quantum Information in Thermodynamics—A Topical Review*, J. Phys. A **49**, 143001 (2016).

[14] M. T. Mitchison, *Quantum Thermal Absorption Machines: Refrigerators, Engines and Clocks*, Contemp. Phys. **60**, 164 (2019).

[15] N. Brunner, N. Linden, S. Popescu, and P. Skrzypczyk, *Virtual Qubits, Virtual Temperatures, and the Foundations of Thermodynamics*, Phys. Rev. E **85**, 051117 (2012).

[16] A. D. Ludlow, M. M. Boyd, J. Ye, E. Peik, and P. O. Schmidt, *Optical Atomic Clocks*, Rev. Mod. Phys. **87**, 637 (2015).

[17] B. Seiferle, L. von der Wense, P. V. Bilous, I. Amersdorffer, C. Lemell, F. Libisch, S. Stellmer, T. Schumm, C. E. Düllmann, A. Pálffy, and P. G. Thirolf, *Energy of the* ^{229}Th *Nuclear Clock Transition*, Nature (London) **573**, 243 (2019).

[18] R. Lutwak, J. Deng, W. Riley, M. Varghese, J. Leblanc, G. Tepolt, M. Mescher, D. K. Serkland, K. M. Geib, and G. M. Peake, *The Chip-Scale Atomic Clock-Low-Power Physics Package, technical report, Symmetricom—Technology Realization Center, Beverly, MA*, 2004.

[19] Y. Guryanova, N. Friis, and M. Huber, *Ideal Projective Measurements Have Infinite Resource Costs*, Quantum **4**, 222 (2020).

[20] T. Debarba, G. Manzano, Y. Guryanova, M. Huber, and N. Friis, *Work Estimation and Work Fluctuations in the Presence of Non-ideal Measurements*, New J. Phys. **21**, 113002 (2019).

[21] A. S. L. Malabarba, A. J. Short, and P. Kammerlander, *Clock-Driven Quantum Thermal Engines*, New J. Phys. **17**, 045027 (2015).

[22] H. Ball, W. D. Oliver, and M. J. Biercuk, *The Role of Master Clock Stability in Scalable Quantum Information Processing*, npj Quantum Inf. **2**, 16033 (2016).

[23] T. Feldmann and J. P. Palao, *Performance of Quantum Thermodynamic Cycles*, 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, Switzerland, 2019), Chap. 3, pp. 67–85, https://doi.org/10.1007/978-3-319-99046-0_3.

[24] S. Raeisi, M. Kieferová, and M. Mosca, *Novel Technique for Robust Optimal Algorithmic Cooling*, Phys. Rev. Lett. **122**, 220501 (2019).

[25] Á. M. Alhambra, M. Lostaglio, and C. Perry, *Heat-Bath Algorithmic Cooling with Optimal Thermalization Strategies*, Quantum **3**, 188 (2019).

[26] F. Clivaz, R. Silva, G. Haack, J. B. Brask, N. Brunner, and M. Huber, *Unifying Paradigms of Quantum Refrigeration: A Universal and Attainable Bound on Cooling*, Phys. Rev. Lett. **123**, 170605 (2019).

[27] F. Clivaz, R. Silva, G. Haack, J. B. Brask, N. Brunner, and M. Huber, *Unifying Paradigms of Quantum Refrigeration: Fundamental Limits of Cooling and Associated Work Costs*, Phys. Rev. E **100**, 042130 (2019).

[28] N. A. Rodríguez-Briones and R. Laflamme, *Achievable Polarization for Heat-Bath Algorithmic Cooling*, Phys. Rev. Lett. **116**, 170501 (2016).

[29] S. Ranković, Y.-C. Liang, and R. Renner, *Quantum Clocks and Their Synchronisation—The Alternate Ticks Game*, arXiv:1506.01373.

[30] S. Stupar, C. Klumpp, R. Renner, and N. Gisin, *Performance of Stochastic Clocks in the Alternate Ticks Game*, arXiv:1806.08812.

[31] M. P. Woods, R. Silva, G. Pütz, S. Stupar, and R. Renner, *Quantum Clocks Are More Accurate than Classical Ones*, arXiv:1806.00491.

[32] Y. Yang and R. Renner, *Ultimate Limit on Time Signal Generation*, arXiv:2004.07857.

[33] M. P. Woods, *Autonomous Ticking Clocks from Axiomatic Principles*, Quantum **5**, 381 (2021).

[34] N. Y. Halpern and D. T. Limmer, *Fundamental Limitations on Photoisomerization from Thermodynamic Resource Theories*, Phys. Rev. A **101**, 042116 (2020).

[35] M. Morrison and G. Parker, *A Guide to Rotations in Quantum Mechanics*, Aust. J. Phys. **40**, 465 (1987).

[36] E. P. Wigner and J. J. Griffin, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra*, Pure and Applied Physics (Academic, New York, 1959).

[37] J. Wiener, *Integrals of* $\cos^{2n}(x)$ *and* $\sin^{2n}(x)$, Coll. Math. J. **31**, 60 (2000).