# Efficiently Cooling Quantum Systems with Finite Resources: Insights from Thermodynamic Geometry

Philip Taranto,[1,2,*,†] Patryk Lipka-Bartosik,[3,4,*] Nayeli A. Rodríguez-Briones,[5,6] Martí Perarnau-Llobet,[7,3] Nicolai Friis,[5,8] Marcus Huber,[5,8,‡] and Pharnam Bakhshinezhad[5,§]

[1]*Department of Physics and Astronomy, University of Manchester, Manchester M13 9PL, United Kingdom*
[2]*Department of Physics, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo City, Tokyo 113-0033, Japan*
[3]*Department of Applied Physics, University of Geneva, 1211 Geneva, Switzerland*
[4]*Institute of Theoretical Physics, Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, 30-348 Kraków, Poland*
[5]*Atominstitut, Technische Universität Wien, Stadionallee 2, 1020 Vienna, Austria*
[6]*Miller Institute for Basic Research in Science, University of California, Berkeley, California 94720, USA*
[7]*Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, 08193 Bellatera (Barcelona), Spain*
[8]*Institute for Quantum Optics and Quantum Information—IQOQI Vienna, Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria*

(Received 16 April 2024; accepted 16 December 2024; published 18 February 2025)

Landauer's limit on heat dissipation during information erasure is critical as devices shrink, requiring optimal pure-state preparation to minimize errors. However, Nernst's third law states this demands infinite resources in energy, time, or control complexity. We address the challenge of cooling quantum systems with *finite* resources. Using Markovian collision models, we explore resource trade-offs and present efficient cooling protocols (that are optimal for qubits) for coherent and incoherent control. Leveraging thermodynamic length, we derive bounds on heat dissipation for swap-based strategies and discuss the limitations of preparing pure states efficiently.

DOI: 10.1103/PhysRevLett.134.070401

*Introduction*—One of the most essential tasks in quantum science is preparing pure quantum states, equivalent to cooling physical systems or erasing information. This is a critical prerequisite for quantum computation, where the output state from a calculation must be erased before it can be reused as an input for the next [1]. Failure to create sufficiently pure states leads to computational errors and reduces the accuracy of timekeeping [2,3] and measurement [4]. Without adequate purity, possibly due to limited resources or control, the frequency of gate and measurement errors increases, potentially relegating any "quantum advantage" to mere conjecture.

In this sense, thermodynamics links the degree of *control* over a system with one's *capacity* to perform useful tasks. Landauer established that a minimum amount of heat must be dissipated when erasing information encoded in *any* physical system, formalizing a connection between physics and information [5]. This limit applies to classical and quantum theory, gaining prominence as computing elements are miniaturized and become more susceptible to heat-induced errors. Efforts to saturate the Landauer bound involve engineering quasistatic interactions between information-bearing systems and controllable machines. However, determining the necessary conditions for Landauer-cost erasure has been impeded by inequivalent assumptions across experimental platforms.

A breakthrough by Reeb and Wolf reformulated the Landauer limit in the context of quantum information, providing platform-agnostic insights [6]. They demonstrated the need for an infinitely large energy gap in an infinite-dimensional machine to achieve perfect Landauer-cost erasure. Yet, infinite resources are not practically accessible, leading to the challenge of optimizing cooling with finite resources [7,8]. When resources are limited, factors such as the energy-level structure of the cooling machines and the complexity of their interactions with the target system influence the achievable final purity and associated energy cost, leading to a three-way trade-off among *energy*, *time*, and *control complexity* [9].

*These authors contributed equally to this work.
†Contact author: philip.taranto@manchester.ac.uk
‡Contact author: marcus.huber@tuwien.ac.at
§Contact author: pharnam.bakhshinezhad@tuwien.ac.at

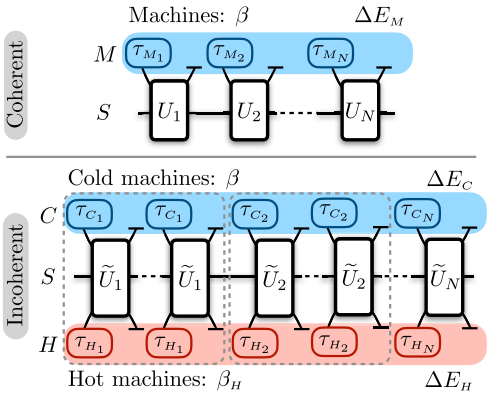


FIG. 1. Framework. A system $S$ is cooled via sequential interactions with a system $M$ comprising $N$ machines, using a Markovian collision model. In the coherent-control setting (top), arbitrary unitaries $U_i$ act between $S$ and initially thermal machines $\tau_{M_i}(\beta, H_{M_i})$. In the incoherent-control setting (bottom), $M$ splits into hot ($H$) and cold ($C$) systems at inverse temperatures $\beta_H$ and $\beta$, respectively. Here, the unitaries are restricted to be energy conserving, i.e., $[\tilde{U}_i, H_S + H_{C_i} + H_{H_i}] = 0$. Moreover, we allow repeated interactions between the target and copies of hot and cold machines, which we call a *stage* (gray outline). The energy cost is the change in energy of the appropriate machines, $\Delta E_X$ for $X \in \{M, C, H\}$.

Here, we explore the relationship between energy cost and time in the finite-resource regime. This general setting is difficult to analyze, as cooling performance depends on a complex interplay of microscopic details. To make progress, we focus on cooling procedures implemented via a Markovian collision model [10–15], where the target system is cooled through a sequence of unitary interactions with uncorrelated thermal machines. We connect these models to continuous trajectories in the state space and use the geometric technique of thermodynamic length [16–21] to bound the cooling performance. For a simple but insightful collision model based on swap operations, we derive the associated thermodynamic metric and optimal cooling protocols for the case of qubit systems. Our Letter, thus, contributes to the understanding of resource limitations for the important task of preparing pure states. This also represents a first step into connecting the framework of thermodynamic geometry with (Markovian) collision models, establishing methods that may prove valuable for analyzing cooling procedures in more complicated settings.

*Framework*—We consider cooling a target quantum system $S$ via unitary interactions with another system $M$ composed of $N$ subsystems $\{M_1, \ldots, M_N\}$ called *machines*. We describe the procedure by a Markovian *collision model* [10–15]. Here, the target unitarily interacts with a fresh machine at each time step, reflecting the property of memorylessness and rapid machine rethermalization between control operations (see Fig. 1).

All systems $X \in \{S, M_1, \ldots, M_N\}$ have an associated Hilbert space $\mathcal{H}_X$, on which states $\varrho_X$ are represented as positive semidefinite, unit-trace operators. Each system has a Hamiltonian whose spectral decomposition fixes its energy structure, $H_X := \sum_{i=0}^{d_X-1} E_X^{(i)} |i\rangle\langle i|$. We consider finite-dimensional systems $d_X := \dim(\mathcal{H}_X) < \infty$ and assume that $E_X^{(i+1)} \geq E_X^{(i)}$, with $E_X^{(0)} := 0$. With respect to any Hamiltonian $H$, the thermal state at inverse temperature $\beta := (k_B T)^{-1}$ is $\tau(\beta, H) := \mathcal{Z}^{-1}(\beta) \exp(-\beta H)$, where $\mathcal{Z}(\beta) := \mathrm{tr}[\exp(-\beta H)]$ is the partition function; when unambiguous, we will write $\tau(\beta)$. The thermal state uniquely maximizes entropy $S(\varrho) := -\mathrm{tr}[\varrho \log \varrho]$ for fixed average energy $E(\varrho) := \mathrm{tr}[H\varrho]$ providing a suitable initial machine state for *cooling schemes* (formalized below).

Boundary conditions: We investigate procedures that transition a target system from an initial state $\varrho_S$ to a final one $\varrho_S'$ while concurrently transforming the collection of machines from an initial thermal state $\tau_M(\beta)$ to a final state $\varrho_M'$. This transformation occurs via the global evolution described by

$$\varrho_{SM}' = U[\varrho_S \otimes \tau_M(\beta)]U^\dagger. \tag{1}$$

The goal of cooling is to manipulate the target so that the majority of its populations are transferred to the lowest energy eigenstates. Any meaningful notion of cooling can be captured by a majorization relation (see Supplemental Material, Sec. I [22]); for simplicity, we focus on processes that take the target from a Gibbs state characterized by an initial $\beta$ to a final one with $\beta_f := \lambda\beta$, where $\lambda > 1$. This fixes the boundary conditions.

Structural and control resources: Cooling performance is influenced by several factors, including the dimensions $d_X$ and Hamiltonians $H_X$ of all systems, the interaction range $k$ (the number of systems involved in each interaction), the number of machines $N$, and the dissipated heat $\Delta E_M := \mathrm{tr}[H_M(\varrho_M' - \varrho_M)]$, which establishes a lower bound for the energetic cost of any implementation. We distinguish *structural* resources—such as $d_X$ and $H_X$, which are fixed independently of the procedure—from *control* resources linked to the protocol, such as the interaction range $k$, total time duration (the number of time steps $N$ for fixed control complexity $k$), and the dissipated heat $\Delta E_M$.

Type of control: We consider two extremal control paradigms: *coherent* and *incoherent* (see Fig. 1) [9,28,29]. Coherent control allows a work source to implement any system-machine unitary, enabling arbitrary transformations as described in Eq. (1). In contrast, incoherent control employs energy-conserving unitaries between the target and machines at different temperatures to drive all heat and entropy flows. Coherent control represents the highest level of control in a thermodynamic setting, while incoherent control assumes less control, requiring only the switching on and off of interaction Hamiltonians. The settings of heat-bath algorithmic cooling [30–36] and of autonomous cooling [37–43] are contained within these paradigms, respectively.

Cooling schemes: We now define the concept of a cooling scheme, encompassing all aforementioned dependencies.

*Definition 1*—A *cooling scheme* is defined by the tuple $(\mathcal{B},\mathcal{S},\mathcal{C},\mathcal{T})$. Here, $\mathcal{B}$ denotes the *boundary conditions* of the problem, namely, the initial and final temperature of the target system. The *structural resources* $\mathcal{S}$ include $\beta$, $H_X$, and $d_X$, i.e., the initial temperature, Hamiltonians, and dimensions of all systems. The *control resources* $\mathcal{C}$ encompass the total number of machines $N$, the interaction range $k$, and the energy cost $\Delta E_M$. Finally, the *type* $\mathcal{T}$ indicates whether the procedure operates with the coherent or incoherent control.

Notably, Nernst's third law of thermodynamics and Landauer's bound exemplify instances where particular resource configurations preclude achievability. Nernst's law states that infinite resources are required to prepare a pure state [44–46]; in our context, this implies an infinitely large energy gap in the machine is necessary [6,9,28]. Similarly, Landauer's bound establishes that the entropy of the target cannot be reduced by $\tilde{\Delta}S \coloneqq S(\varrho_S) - S(\varrho_S')$ via interactions with a thermal machine without dissipating at least $\beta\Delta E_M$ [1,5,6,47]. Delineating the boundary of attainable cooling procedures for different resource configurations represents a significant open problem [8,9]. Here, we focus on achievable schemes, optimizing over finite resources to develop effective cooling procedures. Specifically, we seek the optimal energy-level structure of machines $\{H_{M_i}\}_i$ and interactions $\{U_i\}_i$ to minimize the energy cost of a cooling scheme characterized by fixed control complexity $k$ and finite duration, represented by the number of steps $N$. To achieve this objective, we leverage the geometric concept of thermodynamic length [16,18,19,21].

*Thermodynamic geometry*—The notion of thermodynamic length enables the characterization of path-dependent thermodynamic quantities, such as (dissipated) work, via a geometric approach [18,21,48]. To define this concept, consider a path in Hamiltonian space $H(t)$ parametrized by $t\in[0,1]$. The *thermodynamic length* associated to such a path is given by

$$\mathcal{L} \coloneqq \int_0^1 \sqrt{\mathrm{cov}_t\big(\beta\dot{H}(t),\beta\dot{H}(t)\big)}\,\mathrm{d}t, \tag{2}$$

where $\dot{H}(t) \coloneqq \partial_t H(t)$ and

$$\mathrm{cov}_t(A,B) \coloneqq \mathrm{tr}[\mathcal{J}_t(A)B] - \mathrm{tr}[\tau(t)A]\mathrm{tr}[\tau(t)B], \tag{3}$$

with $\mathcal{J}_t(A) \coloneqq \int_0^1 \tau(t)^{1-x}A\tau(t)^x\mathrm{d}x$ and $\tau(t) \coloneqq \tau[\beta,H(t)]$. The length squared $\mathcal{L}^2$ is related to the dissipated heat or excess work when slowly driving $H(t)$ while in contact with a bath at inverse temperature $\beta$ [17–20]. The minimal length connecting two end points $H(t_0)$ and $H(t_1)$ minimizes the dissipation along a path in Hamiltonian space and is found by solving the geodesic equations; for commuting Hamiltonians, an analytic solution to Eq. (2) is known (as shown in Ref. [49]).

The notion of thermodynamic length is typically employed in slowly driven systems [19–21,50–53] or discrete-time processes [48,54]. We will show that this useful concept can also be applied to the setting of Markovian collision models in the limit of a large number of machines $N$. The key observation here is that slowly changing the system's Hamiltonian can be approximated by a sequence of simple swap interactions in which the system interacts with thermal machine subsystems with carefully chosen Hamiltonians. This approach enables to characterize the backaction of the process in the machine or bath, in contrast to previous works, e.g., Refs. [19,48,54]. This connection allows us to develop new insights to our question of interest: Given the ability to apply $N$ unitary interactions (of fixed complexity $k$) between system and machines, what is the optimal machine energy-level distribution for cooling the target system?

*Coherent control*—In the coherent-control setting, given infinite resources, the Landauer bound sets the ultimate limit on cooling. We first focus on the role of finite structural complexity in said scenarios. We strive to identify the structural complexity that minimizes the energy cost when the system is cooled via a sequence of $N<\infty$ bipartite ($k=2$) interactions.

*Theorem 1*—In the coherent-control setting, given a qudit system with Hamiltonian $H_S$, the minimum energy cost of cooling from a thermal state $\tau_S(\beta,H_S)$ to $\tau_S(\lambda\beta,H_S)$ using swaps between the system and a fresh qudit machine with arbitrary Hamiltonian at each of the $N$ steps is given by

$$\beta\Delta E_M = \tilde{\Delta}S_S + \frac{1}{2N}(\mathcal{L}^*)^2 + \mathcal{O}(N^{-2}), \tag{4}$$

where $\mathcal{L}^*$ is the minimal thermodynamic length [21,49]:

$$\begin{aligned}\mathcal{L}^* &= 2\arccos\left(\mathrm{tr}\left[\sqrt{\tau_S(\beta,H_S)}\sqrt{\tau_S(\lambda\beta,H_S)}\right]\right)\\ &= 2\arccos\left(\frac{\mathcal{Z}\left[\beta(1+\lambda)/2\right]}{\sqrt{\mathcal{Z}(\beta)\mathcal{Z}(\lambda\beta)}}\right).\end{aligned} \tag{5}$$

Above, $\beta\Delta E_M$ is the (scaled) energy cost of cooling, $\tilde{\Delta}S_S$ is the Landauer bound (the minimum energy cost), and the remaining terms represent the additional finite-resource penalty. As $N$ increases, this additional cost decreases like $1/N$, approaching the Landauer bound. The thermodynamic length $\mathcal{L}^*$ quantifies how far apart the initial and final states are in a thermodynamic sense; the greater this distance, the more energy above Landauer cost is required for finite $N$.

*Sketch of proof*—The proof, fully detailed in the Supplemental Material, Sec. II [22], is based on the equality

form of Landauer's bound [6]:

$$\beta\Delta E_M = \tilde{\Delta}S_S + I(S:M)_{\varrho'_{\mathrm{SM}}} + D[\varrho'_M\|\tau_M(\beta)], \qquad (6)$$

which holds for any transformation described by Eq. (1) such that the entropy of the target changes from $S(\varrho_S)$ to $S(\varrho'_S) = S(\varrho_S) - \tilde{\Delta}S_S$. Here, $I(X:Y)_{\varrho_{XY}} := S(\varrho_X) + S(\varrho_Y) - S(\varrho_{XY})$ is the quantum mutual information, and $D(\varrho_X\|\varrho_Y) := \mathrm{tr}[\varrho_X(\log\varrho_X - \log\varrho_Y)]$ is the quantum relative entropy. This equality breaks down the total energy cost into three terms: the fundamental Landauer bound $\tilde{\Delta}S_S$, correlations built up between system and machine $I(S:M)_{\varrho'_{\mathrm{SM}}}$, and a term measuring how far the machine's final state deviates from thermal equilibrium $D[\varrho'_M|\tau_M(\beta)]$. The proof then proceeds in two steps: First, the bipartite interactions are chosen to be swaps between the qudit system and each of a sequence of qudit machines with increasing energy gaps, such that no correlations are built up between $S$ and $M$ as the system is cooled, i.e., $I(S:M)_{\varrho'_{\mathrm{SM}}} = 0$. Then, the relative-entropy term is minimized; for the sequence of swap operations considered, the relative entropy has the tight lower bound $\frac{1}{2N}(\mathcal{L}^*)^2 + \mathcal{O}(N^{-2})$, and we therefore assert the claim. ■

In summary, having fixed the parameters $d_X < \infty$, $N < \infty$, and $k = 2$, but not the structural complexity, i.e., the machine Hamiltonians $\{H_{M_i}\}_i$, of a coherent-control cooling scheme, we have optimized the remaining control-resource parameter, namely, the energy cost $\Delta E_M$. In the case of qubit target and machines, we show that such swaps constitute the optimal interaction (see Supplemental Material, Sec. II A [22]) and derive the Hamiltonian that saturates Eq. (4) (see Supplemental Material, Sec. II B [22]), thereby providing the optimal cooling scheme with respect to heat dissipation in the case of qubits. In Fig. 2, we compare this optimal protocol with other known coherent cooling schemes to demonstrate its effectiveness. Although the optimal energy structure in the case of swaps for higher dimensions is given by Eq. (5), determining the optimal operation in general remains an open problem. Indeed, when considering higher-dimensional machines (including those with interacting subsystems), further cooling advantages can be achieved [55,56] (see also Supplemental Material, Sec. II C [22]).

Some comments regarding optimality are in order. First, we are assuming that the cooling procedure is Markovian, i.e., that the machines are completely refreshed between steps of controlled evolution. In this setting, creating correlations costs energy [58–60], which implies that the optimal scheme must minimize the correlations built up between system and machines [61,62]. However, in a non-Markovian setting, correlations could potentially be used in later steps to lower the energy cost or improve performance [15,63,64]. Second, there is a nonzero energy cost for creating coherences [65]. Since we assume an initial state that is diagonal in the energy eigenbasis, this implies that the optimal cooling scheme must permute only populations of energy eigenstates (leading to a final system state that commutes with the initial one [21]), but for general initial states this need not be the case. Nonetheless, since we evaluate the cooling performance in terms of the heat dissipated by the machine and neglect any system-local energy cost (i.e., that associated to local basis changes), our results apply to arbitrary initial and final system states.

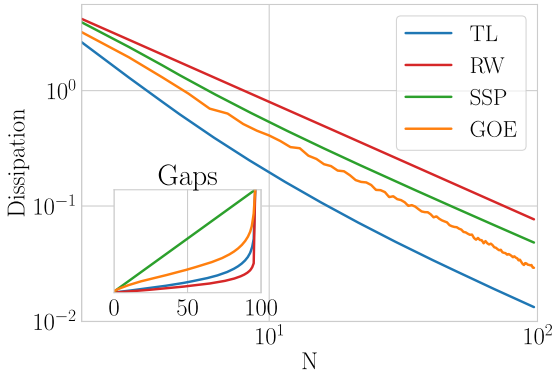


FIG. 2. Cooling with coherent control. Heat dissipation when coherently cooling a qubit target with energy gap $E_S = 1$ from $\beta = 10^{-8}$ to $\beta_f = 10$ (i.e., $\lambda = 10^9$) via sequential bipartite interactions with $N$ machine qubits. We compare four procedures that depend on the energy-level structure $\{H_{M_i}\}_i$, whose gaps are depicted in the inset for $N = 100$. TL is the optimal protocol deduced from the thermodynamic length (see Theorem 1); RW corresponds to the protocol from Ref. [6] where machine energy gaps change linearly; SSP corresponds to the protocol from Ref. [57] where the machine qubits' excited-state populations change linearly. GOE is a protocol where the machine gaps correspond to energy level spacings drawn from the Gaussian orthogonal ensemble (see Supplemental Material, Sec. III [22]); interestingly, this procedure also outperforms the RW and SSP protocols.

*Incoherent control*—We now consider the same question in the incoherent-control setting. In contrast to the coherent-control setting, here all heat and entropy flows occur solely by coupling the target system to a hot ($H$) and cold ($C$) machine, leading to an energy-conserving transformation overall. Beginning with $\varrho_{SCH} = \tau_S(\beta) \otimes \tau_C(\beta) \otimes \tau_H(\beta_H)$, where $\beta_H < \beta$, the considered evolution leads to the output state $\varrho'_{SCH} = \tilde{U}\varrho_{SCH}\tilde{U}^\dagger$, where $[\tilde{U}, H_S + H_C + H_H] = 0$ encodes energy conservation, i.e., that the total energy remains constant throughout. Note that, in the incoherent-control scenario, no energy-conserving unitary with another single thermal machine can lead to cooling the target [29]; hence, interactions between at least three systems must be considered.

In this setting, the Landauer bound is unattainable; instead, the ultimate limit is given by the Carnot-Landauer bound [9]

$$\Delta F_S^\beta + \eta\Delta E_H \le 0, \qquad (7)$$

which follows from the equality form

$$\Delta F_S^\beta + \eta \Delta E_H = -\beta^{-1}\{\Delta S_S + \Delta S_C + \Delta S_H + D[\varrho_C' \| \tau_C(\beta)] + D[\varrho_H' \| \tau_H(\beta_H)]\}. \quad (8)$$

Here, we have introduced the free energy $F_X^\beta(\varrho) \coloneqq \mathrm{tr}[H_X \varrho] - \beta^{-1} S(\varrho)$ and the Carnot efficiency $\eta \coloneqq 1 - \beta_H/\beta \in [0, 1]$.

In a similar vein to the coherent-control scenario, we wish to bound the right-hand side of Eq. (8) for any finite-resource implementation and, ideally, identify a protocol that saturates this bound. However, a number of problems immediately arise in the incoherent-control scenario, since one is restricted to the orbit of energy-conserving unitaries, i.e., $\tilde{U}$ such that $[\tilde{U}, H_S + H_C + H_H] = 0$. This constraint implies that the relative-entropy terms cannot be bounded simply by the thermodynamic length, which was possible in the coherent-control setting because the full swap led to a straightforward expression in terms of a sequence of relative-entropy terms applied to the chain of machines. Here, such a swap is prohibited by energy conservation. We now present an attainable energy bound for finite-resource cooling with incoherent control, which is generally optimal for qubits and optimal for qudits within the considered class of interactions, in analogy to Theorem 1 in the coherent-control setting.

*Theorem 2*—In the incoherent-control setting, given a qudit system with Hamiltonian $H_S$, the minimum energy cost for cooling from a thermal state $\tau_S(\beta, H_S)$ to $\tau_S(\lambda\beta, H_S)$ using energy-conserving interactions between the system and two fresh qudit machines at inverse temperature $\beta$ (cold) and $\beta_H \leq \beta$ (hot), with arbitrary Hamiltonians in the limit of infinite steps but with $N < \infty$ distinct energy gaps, is given by

$$\Delta F_S^\beta + \eta \Delta E_H = -\frac{1}{2N\beta}(\mathcal{L}^*)^2 + \mathcal{O}(N^{-2}). \quad (9)$$

This result shows that, like in the coherent case, the additional energy cost above the fundamental limit (here, the Carnot-Landauer bound) scales as $1/N$ in the finite-resource setting.

*Sketch of proof*—The proof, presented in the Supplemental Material, Sec. IV [22], is fundamentally different to its coherent-control counterpart. In the constructive direction, we propose a cooling scheme comprising interactions that exchange populations among levels $|i, i+1, i\rangle_{SCH} \leftrightarrow |i+1, i, i+1\rangle_{SCH}$. The energy-conserving nature allows us to calculate the energy cost per population transfer, which is related to the relative entropy between the initial and final states of the virtual-qubit subspaces of the hot-and-cold machine that permit cooling. We finally bound this quantity by the thermodynamic length. ∎

For qudits, it is not clear that the form of energy-conserving interactions considered here is optimal; nonetheless, within this family, we present a cooling scheme that attains the energy cost of Eq. (9) and saturates the Carnot-Landauer bound in the limit of infinitely many distinct energy gaps, i.e., diverging control complexity. In the special case of cooling a qubit target with (hot and cold) qubit machines, we show that the cycle $|010\rangle_{SCH} \leftrightarrow |101\rangle_{SCH}$ is indeed optimal. This is because, for any fixed set of energy gaps, the family of energy-conserving unitaries on three qubits that permit cooling without creating coherences or correlations must be of this form, and, thus, we can cover the entire orbit of unitaries in question. Such operations can be considered as a *virtual* swap between the target and the *virtual* qubit subspace of the machine spanned by $\{|01\rangle_{CH}, |10\rangle_{CH}\}$. In general, since such a subspace has norm strictly less than one, each such virtual swap will lead to the system qubit being at strictly higher temperature than the virtual qubit. However, in the limit of infinitely many repetitions within a single stage, the temperature of the system's qubit subspace of interest converges to the virtual temperature of the machine-qubit subspace [28,29]. As we are interested in finite resources, we assume that one performs a finite but sufficiently large number of virtual swaps so that the error is within specified tolerances. The relative entropy term that governs the finite-cooling behavior here (and which leads to the thermodynamic-length term) concerns the initial and final thermal states of the machine at the corresponding virtual temperature defined by the qubit subspace in question. Implementing the protocol that swaps the target successively with appropriately chosen virtual qubits of the machine in each stage minimizes the thermodynamic length and, therefore, provides the optimal incoherent cooling procedure.

*Role of correlations*—The constraint of energy conservation distinguishes the paradigms of coherent and incoherent control. In the latter setting, the virtual subspaces spanned by the hot-and-cold machines influence the performance of a cooling scheme, rather than the state of the machine *per se*. This suggests that *correlations* play a dominant role in the incoherent-control setting; we now formalize this intuition.

*Theorem 3*—For any incoherent-control cooling scheme, the sum of free-energy differences $\Delta F_X^\beta$ (with respect to inverse temperature $\beta$) is bounded by the sum of generated bipartite correlations $\Delta I_\alpha$:

$$\sum_{X \in \{S,C,H\}} \Delta F_X^\beta \leq -\frac{2}{3}\beta^{-1} \sum_\alpha \Delta I_\alpha, \quad (10)$$

where $\alpha \in \{SC, SH, CH\}$ and $I_{XY} \coloneqq I(X:Y)_{\varrho_{XY}}$ is the quantum mutual information.

A proof is given in Supplemental Material, Sec. V [22]. This bound has interesting implications. For instance, *a priori*, the only claim that one can deduce regarding the free-energy change of the hot machine is $\Delta F_H^\beta \leq 0$, which follows from both the system and cold machine

beginning in thermal states at inverse temperature $\beta$. However, using the relation $\beta\Delta F_X^\beta = D[\varrho_X'\|\tau_X(\beta)]$ for $X \in \{S, C\}$, we can derive the tighter bound

$$\beta\Delta F_H^\beta \leq -\frac{2}{3}\sum_\alpha \Delta I_\alpha - D[\varrho_S'\|\tau_S(\beta)] - D[\varrho_C'\|\tau_C(\beta)] \leq 0, \tag{11}$$

where the second inequality follows from the nonnegativity of both the mutual information and the relative entropy.

*Discussion*—Efficient cooling of quantum systems in practice necessitates the optimization of machines and interactions over complicated resource constraints. Here, we have made several contributions to this problem, focusing primarily on qubit systems while developing methods that may prove valuable in more general settings. First, we formalized the concept of a cooling scheme using a universal definition that captures all relevant dependencies, permitting fair comparison among different procedures. Second, for the case of fixed control complexity, we demonstrated simple protocols that asymptotically saturate the ultimate bounds and dissipate minimal heat in the regime of many (but finite) machines, establishing their optimality specifically for qubit systems under both coherent and incoherent control. Our main technical contribution bounds the heat dissipated by the machine in a cooling process in terms of the geometric concept of thermodynamic length. We make this connection by modeling cooling processes as Markovian collision models, linking prominent methods used in quantum thermodynamics and information theory. Finally, we analyzed the role of correlations in the incoherent-control setting, deriving a bound on free-energy differences in terms of correlations.

A key direction for future work is the extension of our analysis to higher-dimensional systems, where the optimality of our protocols remains an open question. While our framework provides a foundation for such investigations, the optimization problem becomes considerably more complex beyond the qubit setting, potentially requiring different mathematical techniques and approaches. Other important directions include extending our analysis beyond Markovian collision models and developing protocols that saturate the correlation bounds in the incoherent-control setting. Yet, like other higher-dimensional problems at the intersection of thermodynamics and information theory (e.g., regarding symmetrically thermalizing unitaries [60]), general solutions and optimality proofs may be difficult to obtain due to the large parameter spaces involved. In light of this observation, more pragmatic platform-specific approaches may be called for, and we hence envisage future attempts to address such questions to be tailored to more particular (experimental) setups.

*Acknowledgments*—We thank Fabien Clivaz, Morteza Rafiee, and Ralph Silva for insightful discussions. P. T. acknowledges funding from the Japan Society for the Promotion of Science (JSPS) Postdoctoral Fellowships for Research in Japan and the IBM-UTokyo Laboratory. P. L.-B. acknowledges the Swiss National Science Foundation (SNSF) through Project No. 192244 and the Polish National Science Centre through project Sonata 2023/37351/D/ST2/02309. M. P.-L. acknowledges the Swiss National Science Foundation for financial support through Ambizione grant PZ00P2-186067 and also acknowledges funding from the Spanish Agencia Estatal de Investigacion through the grant "Ramón y Cajal RYC2022-036958-I". N. A. R.-B. acknowledges the support from the Miller Institute for Basic Research in Science, at the University of California, Berkeley. N. F. acknowledges financial support from the Austrian Science Fund (FWF) through the stand-alone Project No. P 36478-N funded by the European Union—NextGenerationEU, as well as by the Austrian Federal Ministry of Education, Science and Research via the Austrian Research Promotion Agency (FFG) through the flagship Project No. FO999897481 funded by the European Union—NextGenerationEU. This publication was made possible through the support of Grant No. 62423 from the John Templeton Foundation. The opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of the John Templeton Foundation. M. H., N. A. R.-B., and P. B. acknowledge funding from the European Research Council (Consolidator grant "Cocoquest" 101043705). M. H. further acknowledges funding by FQXi (FQXi-IAF19-03-S2, within the project "Fueling quantum field machines with information") as well as the European flagship on quantum technologies ("ASPECTS" consortium 101080167).

[1] Charles H. Bennett, The thermodynamics of computation—a review, Int. J. Theor. Phys. **21**, 905 (1982).
[2] Jake Xuereb, Paul Erker, Florian Meier, Mark T. Mitchison, and Marcus Huber, Impact of imperfect timekeeping on quantum control, Phys. Rev. Lett. **131**, 160204 (2023).
[3] Emanuel Schwarzhans, Maximilian P. E. Lock, Paul Erker, Nicolai Friis, and Marcus Huber, Autonomous temporal probability concentration: Clockworks and the second law of thermodynamics, Phys. Rev. X **11**, 011046 (2021).
[4] Yelena Guryanova, Nicolai Friis, and Marcus Huber, Ideal projective measurements have infinite resource costs, Quantum **4**, 222 (2020).
[5] Rolf Landauer, Irreversibility and heat generation in the computing process, IBM J. Res. Dev. **5**, 183 (1961).
[6] David Reeb and Michael M. Wolf, An improved Landauer principle with finite-size corrections, New J. Phys. **16**, 103011 (2014).
[7] Florian Meier and Hayata Yamasaki, Energy-consumption advantage of quantum computation, arXiv:2305.11212.

[8] Anthony Munson, Naga Bhavya Teja Kothakonda, Jonas Haferkamp, Nicole Yunger Halpern, Jens Eisert, and Philippe Faist, Complexity-constrained quantum thermodynamics, arXiv:2403.04828.
[9] Philip Taranto, Faraj Bakhshinezhad, Andreas Bluhm, Ralph Silva, Nicolai Friis, Maximilian P. E. Lock, Giuseppe Vitagliano, Felix C. Binder, Tiago Debarba, Emanuel Schwarzhans, Fabien Clivaz, and Marcus Huber, Landauer versus Nernst: What is the true cost of cooling a quantum system? PRX Quantum **4**, 010332 (2023).
[10] Jayaseetha Rau, Relaxation phenomena in spin and harmonic oscillator systems, Phys. Rev. **129**, 1880 (1963).
[11] Mário Ziman, Peter Štelmachovič, Vladimir Bužek, Mark Hillery, Valerio Scarani, and Nicolas Gisin, Diluting quantum information: An analysis of information transfer in system-reservoir interactions, Phys. Rev. A **65**, 042105 (2002).
[12] Valerio Scarani, Mário Ziman, Peter Štelmachovič, Nicolas Gisin, and Vladimir Bužek, Thermalizing quantum machines: Dissipation and entanglement, Phys. Rev. Lett. **88**, 097905 (2002).
[13] Mário Ziman, Peter Štelmachovič, and Vladimir Bužek, Description of quantum dynamics of open systems based on collision-like models, Open Syst. Inf. Dyn. **12**, 81 (2005).
[14] Francesco Ciccarello, Collision models in quantum optics, Quantum Meas. Quantum Metrol. **4**, 53 (2017).
[15] Philip Taranto, Faraj Bakhshinezhad, Philipp Schüttelkopf, Fabien Clivaz, and Marcus Huber, Exponential improvement for quantum cooling through finite-memory effects, Phys. Rev. Appl. **14**, 054005 (2020).
[16] Frank Weinhold, Metric geometry of equilibrium thermodynamics, J. Chem. Phys. **63**, 2479 (1975).
[17] Peter Salamon and Richard Stephen Berry, Thermodynamic length and dissipated availability, Phys. Rev. Lett. **51**, 1127 (1983).
[18] Gavin E. Crooks, Measuring thermodynamic length, Phys. Rev. Lett. **99**, 100602 (2007).
[19] David A. Sivak and Gavin E. Crooks, Thermodynamic metrics and optimal paths, Phys. Rev. Lett. **108**, 190602 (2012).
[20] Matteo Scandi and Martí Perarnau-Llobet, Thermodynamic length in open quantum systems, Quantum **3**, 197 (2019).
[21] Paolo Abiuso, Harry J. D. Miller, Martí Perarnau-Llobet, and Matteo Scandi, Geometric optimisation of quantum thermodynamic processes, Entropy **22**, 1076 (2020).
[22] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.134.070401, which includes Refs. [23–27], for proofs and technical details.
[23] Fabien Clivaz, Optimal manipulation of correlations and temperature in quantum thermodynamics, Ph.D. thesis, University of Geneva, 2020, arXiv:2012.04321.
[24] Fumio Hiai and Dénes Petz, *Introduction to Matrix Analysis and Applications* (Springer Cham, Switzerland, 2014).
[25] Giacomo Livan, Marcel Novaes, and Pierpaolo Vivo, *Introduction to Random Matrices: Theory and Practice* (Springer, Cham, Switzerland, 2018).
[26] Hans A. Weidenmüller, Random-matrix model for thermalization, J. Phys. A **57**, 165002 (2024).
[27] Elliott H. Lieb and Mary Beth Ruskai, Proof of the strong subadditivity of quantum mechanical entropy, J. Math. Phys. (N.Y.) **14**, 1938 (1973).
[28] Fabien Clivaz, Ralph Silva, Géraldine Haack, Jonatan Bohr Brask, Nicolas Brunner, and Marcus Huber, Unifying paradigms of quantum refrigeration: A universal and attainable bound on cooling, Phys. Rev. Lett. **123**, 170605 (2019).
[29] Fabien Clivaz, Ralph Silva, Géraldine Haack, Jonatan Bohr Brask, Nicolas Brunner, and Marcus Huber, Unifying paradigms of quantum refrigeration: Fundamental limits of cooling and associated work costs, Phys. Rev. E **100**, 042130 (2019).
[30] Jonathan Baugh, Osama Moussa, Colm A. Ryan, Ashwin Nayak, and Raymond Laflamme, Experimental implementation of heat-bath algorithmic cooling using solid-state nuclear magnetic resonance, Nature (London) **438**, 470 (2005).
[31] Leonard J. Schulman, Tal Mor, and Yossi Weinstein, Physical limits of heat-bath algorithmic cooling, Phys. Rev. Lett. **94**, 120501 (2005).
[32] Sadegh Raeisi and Michele Mosca, Asymptotic bound for heat-bath algorithmic cooling, Phys. Rev. Lett. **114**, 100404 (2015).
[33] Nayeli Azucena Rodríguez-Briones and Raymond Laflamme, Achievable polarization for heat-bath algorithmic cooling, Phys. Rev. Lett. **116**, 170501 (2016).
[34] Daniel K. Park, Nayeli Azucena Rodriguez-Briones, Guanru Feng, Rahimi Rahimi, Jonathan Baugh, and Raymond Laflamme, Heat bath algorithmic cooling with spins: Review and prospects, in *Electron Spin Resonance (ESR) Based Quantum Computing*, edited by T. Takui, L. Berliner, and G. Hanson (Springer, New York, 2016), p. 227, 10.1007/978-1-4939-3658-8_8.
[35] Álvaro M. Alhambra, Matteo Lostaglio, and Christopher Perry, Heat-bath algorithmic cooling with optimal thermalization strategies, Quantum **3**, 188 (2019).
[36] Junan Lin, Nayeli A. Rodríguez-Briones, Eduardo Martín-Martínez, and Raymond Laflamme, Thermodynamic analysis of algorithmic cooling protocols: Efficiency metrics and improved designs, Phys. Rev. A **110**, 022215 (2024).
[37] Noah Linden, Sandu Popescu, and Paul Skrzypczyk, How small can thermal machines be? The smallest possible refrigerator, Phys. Rev. Lett. **105**, 130401 (2010).
[38] Amikam Levy and Ronnie Kosloff, Quantum absorption refrigerator, Phys. Rev. Lett. **108**, 070604 (2012).
[39] Nicolas Brunner, Noah Linden, Sandu Popescu, and Paul Skrzypczyk, Virtual qubits, virtual temperatures, and the foundations of thermodynamics, Phys. Rev. E **85**, 051117 (2012).
[40] Mark T. Mitchison, Marcus Huber, Javier Prior, Mischa P. Woods, and Martin B. Plenio, Realising a quantum absorption refrigerator with an atom-cavity system, Quantum Sci. Technol. **1**, 015001 (2016).
[41] Gleb Maslennikov, Shiqian Ding, Roland Hablützel, Jaren Gan, Alexandre Roulet, Stefan Nimmrichter, Jibo Dai, Valerio Scarani, and Dzmitry Matsukevich, Quantum absorption refrigerator with trapped ions, Nat. Commun. **10**, 202 (2019).
[42] Sreenath K. Manikandan, Étienne Jussiau, and Andrew N. Jordan, Autonomous quantum absorption refrigerators, Phys. Rev. B **102**, 235427 (2020).
[43] José Antonio Marín Guzmán, Paul Erker, Simone Gasparinetti, Marcus Huber, and Nicole Yunger Halpern,

Useful autonomous quantum machines, Rep. Prog. Phys. **87**, 12 (2024).
[44] Walther Nernst, Über die Beziehung zwischen Wärmeentwicklung und maximaler Arbeit bei kondensierten Systemen, in *Sitzungsberichte der Koniglich Preussischen Akademie der Wissenschaften* (Akademie-Verlag, Berlin, 1906), p. 933.
[45] Francesco Ticozzi and Lorenza Viola, Quantum resources for purification and cooling: Fundamental limits and opportunities, Sci. Rep. **4**, 5192 (2014).
[46] Lluis Masanes and Jonathan Oppenheim, A general derivation and quantification of the third law of thermodynamics, Nat. Commun. **8**, 14538 (2017).
[47] Massimiliano Esposito and Christian Van den Broeck, Second law and Landauer principle far from equilibrium, Europhys. Lett. **95**, 40004 (2011).
[48] James Nulton, Peter Salamon, Bjarne Andresen, and Qi Anmin, Quasistatic processes as step equilibrations, J. Chem. Phys. **83**, 334 (1985).
[49] Anna Jenčová, Geodesic distances on density matrices, J. Math. Phys. (N.Y.) **45**, 1787 (2004).
[50] Kay Brandner and Keiji Saito, Thermodynamic geometry of microscopic heat engines, Phys. Rev. Lett. **124**, 040602 (2020).
[51] Harry J. D. Miller and Mohammad Mehboudi, Geometry of work fluctuations versus efficiency in microscopic thermal machines, Phys. Rev. Lett. **125**, 260602 (2020).
[52] Geng Li, Jin-Fu Chen, C. P. Sun, and Hui Dong, Geodesic path for the minimal energy cost in shortcuts to isothermality, Phys. Rev. Lett. **128**, 230603 (2022).
[53] Adam G. Frim and Michael R. DeWeese, Geometric bound on the efficiency of irreversible thermodynamic cycles, Phys. Rev. Lett. **128**, 230601 (2022).
[54] Matteo Scandi, Harry J. D. Miller, Janet Anders, and Martí Perarnau-Llobet, Quantum work statistics close to equilibrium, Phys. Rev. Res. **2**, 023377 (2020).
[55] A. Rolandi, P. Abiuso, and M. Perarnau-Llobet, Collective advantages in finite-time thermodynamics, Phys. Rev. Lett. **131**, 210401 (2023).
[56] Patryk Lipka-Bartosik and Martí Perarnau-Llobet, Minimizing dissipation via interacting environments: Quadratic convergence to Landauer bound, arXiv:2411.00944.
[57] Paul Skrzypczyk, Anthony J. Short, and Sandu Popescu, Work extraction and thermodynamics for individual quantum systems, Nat. Commun. **5**, 4185 (2014).
[58] Marcus Huber, Martí Perarnau-Llobet, Karen V. Hovhannisyan, Paul Skrzypczyk, Claude Klöckl, Nicolas Brunner, and Antonio Acín, Thermodynamic cost of creating correlations, New J. Phys. **17**, 065008 (2015).
[59] David Edward Bruschi, Martí Perarnau-Llobet, Nicolai Friis, Karen V. Hovhannisyan, and Marcus Huber, The thermodynamics of creating correlations: Limitations and optimal protocols, Phys. Rev. E **91**, 032118 (2015).
[60] Faraj Bakhshinezhad, Fabien Clivaz, Giuseppe Vitagliano, Paul Erker, Ali T. Rezakhani, Marcus Huber, and Nicolai Friis, Thermodynamically optimal creation of correlations, J. Phys. A **52**, 465303 (2019).
[61] Maryam Afsary, Marzieh Bathaee, Faraj Bakhshinezhad, Ali T. Rezakhani, and Alireza Bahrampour, Binding energy of bipartite quantum systems: Interaction, correlations, and tunneling, Phys. Rev. A **101**, 013403 (2020).
[62] Patryk Lipka-Bartosik, Giovanni Francesco Diotallevi, and Pharnam Bakhshinezhad, Fundamental limits on anomalous energy flows in correlated quantum systems, Phys. Rev. Lett. **132**, 140402 (2024).
[63] Nayeli Azucena Rodríguez-Briones, Eduardo Martín-Martínez, Achim Kempf, and Raymond Laflamme, Correlation-enhanced algorithmic cooling, Phys. Rev. Lett. **119**, 050502 (2017).
[64] Nayeli Azucena Rodríguez-Briones, Jun Li, Xinhua Peng, Tal Mor, Yossi Weinstein, and Raymond Laflamme, Heat-bath algorithmic cooling with correlated qubit-environment interactions, New J. Phys. **19**, 113047 (2017).
[65] Avijit Misra, Uttam Singh, Samyadeb Bhattacharya, and Arun Kumar Pati, Energy cost of creating quantum coherence, Phys. Rev. A **93**, 052335 (2016).