Letter

Optimal control of quantum thermal machines using machine learning

Ilia Khait,[1] Juan Carrasquilla,[2,3] and Dvira Segal[1,4,*]

[1]*Department of Physics, University of Toronto, 60 Saint George St., Toronto, Ontario M5S 1A7, Canada*
[2]*Vector Institute, MaRS Centre, Toronto, Ontario M5G 1M1, Canada*
[3]*Department of Physics & Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[4]*Department of Chemistry and Centre for Quantum Information and Quantum Control, University of Toronto, 80 Saint George St., Toronto, Ontario M5S 3H6, Canada*

(Received 16 August 2021; revised 29 November 2021; accepted 25 January 2022; published 11 March 2022)

We develop a deep learning (DL) framework assisted by differentiable programming for discovery of optimal quantum control protocols under hard constraints. To that end, we use neural network representations to our protocols, whose learning process is done with exact gradients. We find high-quality solutions to the optimization problem of finite-time thermodynamical process in a quantum thermal machine. Using this DL algorithm, we show that a previously employed, intuitive energetic cost of the thermal machine driving suffers from a fundamental flaw, which we resolve with an alternative construction for the cost function. Our DL-quantum control framework can be utilized to solve other quantum dynamics and thermodynamics problems.

DOI: 10.1103/PhysRevResearch.4.L012029

I. INTRODUCTION

Many problems in physics are formulated as optimization tasks by identifying a cost function that must be minimized. Prime examples are Hamilton's principle of least action in Lagrangian mechanics [1], Fermat's law of least time in classical optics [2], and more recently, variational algorithms in quantum computing [3]. Similarly, since its inception, thermodynamics has been concerned with performance optimization by identifying constrains and bounds on energy conversion processes. The ideal Carnot engine is designed to reach the maximal efficiency. However, this upper bound is theoretically obtained for arbitrarily slow, quasistatic processes; thus, the extracted power reduces to zero. Quasistatic processes are described using the framework of equilibrium thermodynamics. In contrast, real thermal devices operate on finite-time cycles, and they are naturally described in terms of finite-time thermodynamics [4,5]. This theory is concerned with, e.g., how the efficiency of thermal machines erodes when heat-to-work conversion processes take place in finite-time cycles [6,7].

In this paper, we propose a general framework for quantum control of driven quantum systems. This approach differs from traditional methods that rely on specific type of control inputs, e.g., bang-bang control [8–10], numeric [11], or analytical approximations [12]. Instead, a field of machine learning (ML)-assisted quantum control is rising [13–18]. Specifically, reinforcement learning (RL) is used to study optimal control in the quantum domain and shows promising results [16,19,20]. Here, we harness state-of-the-art ML techniques and use a general variational ansatz that relies on a neural network (NN), which we advance to treat hard-constrained problems. Building on exact gradients [21] and the full representability of NNs [22], our approach allows for unbiased and exact discovery of protocols in driven quantum systems. Our ML-quantum control approach is general and can be used to design and optimize a large variety of problems that require modifying the time-dependent, spatial, or interaction form of potentials [16,23–29]. Bringing ML and quantum control techniques to quantum thermal machines opens the door for various avenues such as device characterization, protocol discovery, and process optimizations.

Quantum thermal machines, in which, e.g., quantum coherences, correlations, and quantum statistics play a decisive role, cater fundamental understanding of thermodynamics at the nano and atomistic scale [30,31]. Beyond fundamental interest, quantum thermal machines promise compact, fast, and efficient work extraction and refrigeration schemes for quantum devices. It remains, however, a challenge to harness such effects and achieve a quantum advantage in thermal machines [32–36].

Optimizing the performance of *nanoscale, quantum* thermal machines is a central problem in the rapidly emerging field of quantum thermodynamics. Techniques such as shortcut to adiabaticity (STA) allow the design of finite-time protocols, which reproduce the same final state of an adiabatic time evolution yet at a price of supplemental work on the system [37–42]. Much theoretical and experimental effort [38,43–50] has been made to realize and characterize these systems. Here, we focus on a specific class of STA protocols—local counterdiabatic driving (LCD)—which is advantageous to the realization of quantum engines since it only requires the application of local time-dependent potentials.

Advances in diverse research topics, such as image recognition and natural language processing have led physicists

*dvira.segal@utoronto.ca

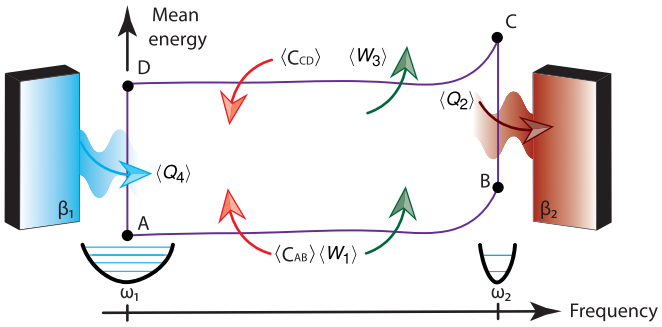

FIG. 1. Scheme of the Otto refrigerator in the energy-frequency domain. A cycle includes a compression stroke (AB) of duration τ, an instant isochoric stroke (BC) with the system coupled to a hot bath, an expansion stroke (CD) of duration τ, and another instant isochoric stroke with a cold bath (DA). Refrigeration corresponds to the withdrawal of heat $\langle Q_4\rangle$ from the cold bath. The energetic cost of the cycle is the sum of the work contributions $\langle W_1\rangle$ and $\langle W_3\rangle$, along with the energetic cost of the shortcut-to-adiabaticity (STA) driving, $\langle C_{AB}\rangle$ and $\langle C_{CD}\rangle$.

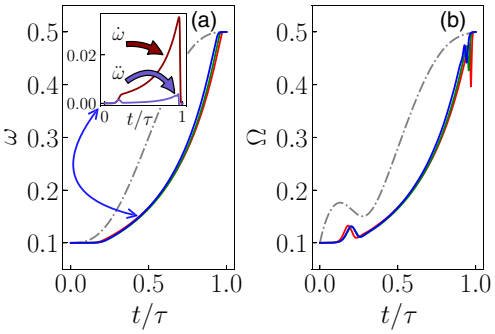

FIG. 2. (a) Examples of frequency profiles $\omega(t)$ (appearing in blue, green, and red), discovered by the differentiable programming (DP) machine learning (ML) scheme normalized by the compression or expansion stroke duration τ. The gray dashed-dotted line displays the polynomial ansatz, Eq. (7). The solid lines are the neural network results. The inset displays the first and second derivatives of $\omega(t)$ for one realization, showing compliance with the shortcut-to-adiabaticity (STA) conditions, Eq. (2). The fact that both derivatives approximately follow each other allows for the minimization of the energetic cost of the STA process (see Fig. 4). (b) The function $\Omega(t)$ corresponding to the frequency of the effective counterdiabatically driven harmonic oscillator. Parameters are $\omega_1 = 0.1$, $\omega_2 = 0.5$, $\beta_1 = 1$, and $\beta_2 = 0.75$.

to exploit ML in quantum dynamics and many-body physics [51]. Here, we adopt differentiable programming (DP) [52,53] to find optimal refrigeration schemes for the quantum Otto cycle under STA conditions (as depicted in Fig. 1). Like an RL perspective, in which an agent plays a game, the time-dependent frequency $\omega(t)$ of the harmonic oscillator acting as the working medium of the refrigerator can be varied in the time interval $t \in [0, \tau]$. For each attempted strategy $\omega(t)$, the agent receives a reward designed to minimize the energetic cost of the protocol while subjected to the physical constraints imposed by the LCD condition (both aspects are elaborated later in this paper). In our deep learning (DL)-based scheme, we optimize the devised strategies by using exact gradients with respect to the variational parameters. The driving profiles that are discovered by this scheme, exemplified in Fig. 2, are superior to previously proposed protocols [54,55]. Furthermore, the ML approach helps uncover a fundamental problem with a previously suggested energetic cost metric, which under some conditions violates basic physical principles (Carnot bound). We show a systematic optimization path based on state-of-the-art ML tools, which permits a search in a large multidimensional variational parameter space; the space consists of all functions fulfilling STA conditions. The advantage of the DP-ML scheme is derived in using the exact gradients of the quantity of interest with respect to variational parameters, hence reducing the number of required iterations to reach an extremum [51,56].

II. QUANTUM OTTO REFRIGERATORS

Prime examples of thermal machines are heat engines and refrigerators [9,43,57–59]. While the first performs work by utilizing heat current from a hot reservoir, refrigerators extract heat from a cold bath using external work. As a thermodynamic process, refrigerators attain their maximal (Carnot) cooling efficiency $\epsilon_C = \frac{T_1}{T_2 - T_1}$ (with $T_{1,2}$ as the temperatures of the cold and hot reservoirs, respectively) for an infinitely slow (adiabatic) process. However, for such processes, the power output, defined as the extracted heat over cycle time $J_c = \frac{\langle Q_4\rangle}{2\tau}$, is null due to the infinitely long cycle time $\tau \to \infty$. For a finite-time cycle, the efficiency decreases, and the power output increases. Therefore, the core question of finite-time thermodynamics is: *What is the optimal cycle for a figure of merit given by the cooling efficiency times output power?*

The quantum Otto refrigerator is depicted in Fig. 1. We choose a working medium consisting of a harmonic oscillator governed by the time-dependent Hamiltonian [12]:

$$H_0(t) = \frac{1}{2m}p^2 + \frac{m\omega(t)^2}{2}x^2. \tag{1}$$

The cycle consists of an isothermal compression stroke where the frequency $\omega(t)$ increases from ω_1 at $t = 0$ to ω_2 at $t = \tau$. Then the engine thermalizes with a hot bath in an isochoric stroke, followed by an isothermal expansion of duration τ back to the frequency ω_1 and an isochoric stroke in which heat $\langle Q_4\rangle$ is extracted from a cold bath. The thermalization strokes are assumed instantaneous.

III. STA AND COUNTERDIABATIC DRIVING

The goal of the STA driving is to speed up the compression and expansion strokes, thus enhancing the figure of merit. By adding the nonadiabatic driving $H_{\rm STA}(t)$ to Eq. (1), the final state of the system after a time evolution from $t = 0$ to τ exactly matches the outcome of an adiabatic approximation-based time evolution of $H_0(t)$ [54,60]. A further canonical transformation of $H_{\rm STA}(t)$ leads to the LCD Hamiltonian of a harmonic oscillator [61,62] with frequency $\Omega(t)^2 \equiv \omega(t)^2 - \frac{3\dot{\omega}(t)^2}{4\omega(t)^2} + \frac{\ddot{\omega}(t)}{2\omega(t)}$ [61]. This modified driving should fulfill the following conditions [54]:

$$\begin{aligned} &\omega(0) = \omega_1, \quad \dot{\omega}(0) = 0, \quad \ddot{\omega}(0) = 0, \\ &\omega(\tau) = \omega_2, \quad \dot{\omega}(\tau) = 0, \quad \ddot{\omega}(\tau) = 0, \end{aligned} \tag{2}$$

which ensure that the final state of the system is identical (phase included) to the state resulting from an adiabatic time evolution of $H_0(t)$.

During compression (AB) and expansion (CD) strokes (Fig. 1), the system is thermally isolated, and work is applied. Using the adiabatic solution of the time-dependent Schrödinger equation [63,64], the mean value of work is

$$\langle W_1\rangle = \frac{\hbar\omega_2}{2}\left(1-\frac{\omega_1}{\omega_2}\right)\coth\left(\frac{\beta_1\hbar\omega_1}{2}\right), \tag{3}$$

and similarly for $\langle W_3\rangle$ by replacing $1 \leftrightarrow 2$. Furthermore, the mean heat extracted during the DA stroke is [12]

$$\langle Q_4\rangle = \frac{\hbar\omega_1}{2}\left[\coth\left(\frac{\beta_1\hbar\omega_1}{2}\right)-\coth\left(\frac{\beta_2\hbar\omega_2}{2}\right)\right]. \tag{4}$$

We estimate the energetic cost of the STA driving with the time-averaged Schmidt norm of $H_{\rm STA}(t)$ [65,66]:

$$\langle C_{AB}\rangle = \coth\left(\frac{\beta_1\hbar\omega_1}{2}\right)\frac{\hbar\sqrt{3}}{4\tau}\int_0^\tau dt \times \frac{1}{\omega(t)}\left|-\frac{3\dot\omega_t^2}{4\omega(t)^2}+\frac{\ddot\omega_t}{2\omega(t)}\right|. \tag{5}$$

Here, $\langle C_{CD}\rangle$ is obtained by switching the temperature and frequency β_1,ω_1 to β_2,ω_2, respectively, see Fig. 1. Our derivation of Eq. (5) is included in Ref. [61]. Below, we show that the energetic cost, Eq. (5), preserves the physical (Carnot) bound, which is missed by other suggested cost metrics.

IV. OPTIMIZATION PROCEDURE

Our goal is to enhance the figure of merit χ defined as the product of the cooling efficiency ϵ with the heat extracted per cycle J_c:

$$\chi \equiv \epsilon J_c = \frac{\langle Q_4\rangle}{\langle W_1\rangle+\langle W_3\rangle+\langle C_{\rm AB}\rangle+\langle C_{\rm CD}\rangle}\frac{\langle Q_4\rangle}{2\tau}. \tag{6}$$

Motivated by Ref. [54], a possible way to boost the figure of merit could be by using a polynomial ansatz, which by construction satisfies the initial conditions of Eq. (2):

$$\omega(t) = \omega_1 + \Delta\omega\sum_{n=3}^{N_{\rm max}}\alpha_n\left(\frac{t}{\tau}\right)^n. \tag{7}$$

Here, $\Delta\omega=\omega_2-\omega_1$. A widely used ansatz which satisfies Eq. (2) consists of $(\alpha_3,\alpha_4,\alpha_5)=(10,-15,6)$ and all the other $\alpha=0$, depicted as the dashed-dotted line in Fig. 2. We use it throughout this paper as a benchmark.

The only quantity that depends on the transient values of $\omega(t)$ is the energetic cost function. Therefore, once the physical parameters (β_i,ω_i) are set, the optimal cooling protocol minimizes the energetic cost $\langle C_i\rangle$ [Eq. (5)]. Thus, we devise a cost function that includes $\langle C_i\rangle$, along with penalties for deviating from the STA constraints, Eq. (2). Details are given in Ref. [61]. For generality, we represent $\omega(t)$ as a NN whose parameters are optimized using automatic differentiation (AD), which allows us to compute exact gradients with respect to the parameters of the NN. We use ADAM [67], a first-order gradient-based optimization algorithm, to optimize our objective function. This process is performed for a large ensemble of 1000 initial conditions for the NN, out of which the optimal strategies are selected.

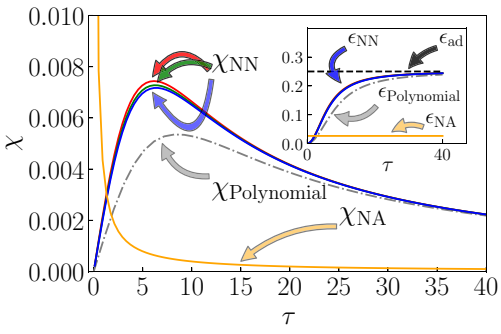

FIG. 3. Figure of merit of the Otto refrigerator χ, Eq. (6), depicted as a function of the expansion or compression stroke times τ for the frequency ramps $\omega(t)$ of Fig. 2. The gray dashed-dotted line displays the polynomial ansatz, Eq. (7). Solid lines represent neural network (NN) results. For comparison, the nonadiabatic sudden-frequency change $\chi_{\rm NA}$ is plotted in light orange. The NN optimization scheme almost doubles the figure of merit compared with the polynomial ansatz. The inset shows the corresponding cooling efficiency and the ideal adiabatic limit $\epsilon_{\rm ad}$. The optimal NN strategies approach the adiabatic efficiency faster than the polynomial benchmark.

V. RESULTS

Examples of optimal expansion profiles $\omega(t)$ are depicted in Fig. 2; the compression stroke is a time-reversed version of it $\omega(\tau-t)$. We note that, during our optimization process, which penalizes for deviations from the initial and final-time conditions, one finds local minima in which the latter are not met. To satisfy those, frequency profiles were stretched, in addition to being smoothened to become physically realizable [61]. In Fig. 2(a), we observe that optimal strategies share a similar feature of a late-bloomer; hence, they are very different from the polynomial ansatz of Eq. (7) depicted as a dashed-dotted line in Fig. 2. The inset shows the first and second derivatives of one of these profiles. The two derivatives rise together, which results in the minimization of the energetic cost function, as we discuss in the next section. Figure 2(b) displays the corresponding frequency ramp $\Omega(t)$, see text above Eq. (2). This frequency would have to be followed to get the same final state as would be achieved by an *adiabatic* driving, which follows $\omega(t)$.

In Fig. 3, we compare the figure of merit χ, Eq. (6), from the different strategies. Overall, we find an almost twofold improvement of NN profiles over the benchmark. The peak for the NN-based strategies occurs earlier than for the polynomial, at $\sim\tau=6$. Further, we compare these performances with the nonadiabatic step function strategy [68]. As expected, the latter strategy could be beneficial for short stroke times $\tau<2$ because the energetic cost of maintaining STA is high. However, it is inferior at longer cycles. We further plot the cooling efficiency ϵ as a function of stroke duration (inset). The black dashed line is the ideal adiabatic efficiency $\epsilon_{\rm ad}$, where the heat $\langle Q_4\rangle$ and work $\langle W_1\rangle$, $\langle W_3\rangle$ attain their adiabatic values and the associated driving energetic cost is neglected

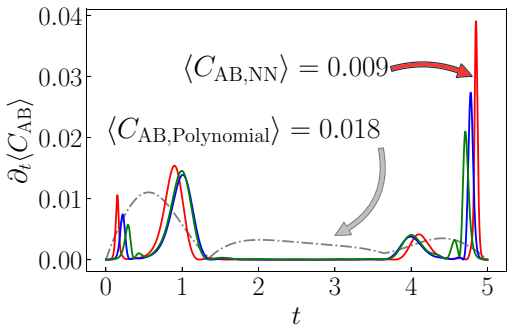

FIG. 4. Mean instantaneous energetic cost of the shortcut-to-adiabaticity (STA) driving as a function of time $\partial_t\langle C\rangle$, defined by the Schmidt norm of $H_{\mathrm{STA}}(t)$, Eq. (5). The dashed line corresponds to the polynomial ansatz, Eq. (7). Solid lines correspond to the optimized neural network (NN) of Figs. 2 and 3 with $\tau = 5$. The NN optimization can reduce the energetic cost by a factor of two compared with the benchmark: $\int_0^\tau \partial_t\langle C_{\mathrm{AB,Polynomial}}\rangle dt / \int_0^\tau \partial_t\langle C_{\mathrm{AB,NN}}\rangle dt = 2.05$.

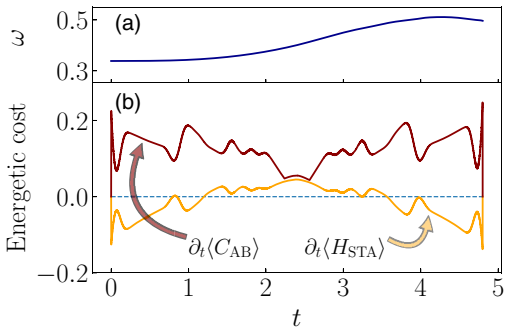

FIG. 5. Exemplifying the failure of the energetic cost definition $\langle H_{\mathrm{STA}}\rangle$. (a) Frequency ramp profile $\omega(t)$ optimized to minimize $\langle H_{\mathrm{STA}}\rangle$. The initial frequency ω_1 is set at 0.34 such that the sum of the average work needed for the expansion and compression strokes $\langle W_1 + W_3\rangle$ is relatively small. (b) Instantaneous energetic cost as a function of time $\partial_t\langle C_{\mathrm{AB}}\rangle$ of Eq. (5) (light line) and $\langle H_{\mathrm{STA}}\rangle$, which is based on the time-averaged mean shortcut-to-adiabaticity (STA) driving (dark line). For the latter, the overall input work plus energetic cost is negative, hence unphysical. In contrast, using $\langle C(t)\rangle$ as the energetic cost yields the cooling efficiency $\epsilon = 0.47 < \epsilon_{\mathrm{ad}} \leqslant \epsilon_C = 3$; $\tau = 4.8$, $\omega_1 = 0.34$, $\omega_2 = 0.5$, $\beta_1 = 1$, and $\beta_2 = 0.75$.

(note that the relation to the Carnot cooling efficiency is $\epsilon_C \geqslant \epsilon_{\mathrm{ad}}$). NN-based strategies can obtain a higher efficiency, and in turn, like the polynomial benchmark, they approach the adiabatic limit at large τ. The cooling protocols minimize the cost metric. In Fig. 4, we show the mean instantaneous energetic cost of the STA drive as a function of time $\partial_t\langle C_{AB}\rangle$. The NN optimization can reduce the energetic cost by a factor of two compared with the polynomial ansatz. Recall that, in Eq. (5), the first and second derivative terms appear with opposite signs. Hence, to reduce the energetic cost, one should envisage functions in which both derivatives follow similar temporal features. This is where the power of NN-based optimization techniques comes into play. The DP-ML method greatly reduces the *instantaneous* cost by realizing functions with this property [see inset in Fig. 2(a)].

Next, we discuss different energetic cost metrics and their adherence to thermodynamical principles. While we employed Eq. (5), previous studies suggested the time average of the mean STA driving $\langle H_{\mathrm{STA}}\rangle = 1/\tau \int_0^\tau dt\langle H_{\mathrm{STA}}(t)\rangle$ as the energetic cost of STA [69–71]. We show that this expression can lead to unphysical results: The system is a refrigerator $\langle Q_4\rangle > 0$, yet the total work plus associated STA cost is negative $\langle W_1 + W_3\rangle + \langle H_{\mathrm{STA}}\rangle < 0$, thus yielding a negative efficiency.

We compare cost metrics in Fig. 5. We use parameters close to the edge of the cooling window [determined by the condition $\beta_1\omega_1 < \beta_2\omega_2$, see Eq. (4)]. This choice leads to relatively small values of $\langle Q_4\rangle$ and $\langle W_1 + W_3\rangle$ compared with parameters used in Fig. 2, and it allows us to demonstrate the incentive for devising a different energetic cost function for STA protocols. NN optimization with the cost metric $\langle H_{\mathrm{STA}}\rangle$ yields profiles that allow cooling, see Fig. 5(a) for an example, yet give an overall *negative* energetic cost; in Fig. 5(b), we show the instantaneous contribution $\partial_t\langle H_{\mathrm{STA}}\rangle$, which is mostly negative. In contrast, the metric $\langle C\rangle$ of Eq. (5) is positive throughout. The cooling efficiency for these parameters becomes negative for the $\langle H_{\mathrm{STA}}\rangle$ energetic cost, which is unphysical: By fine-tuning parameters while maintaining the cooling condition, one can achieve an efficiency that exceeds Carnot [72]. In contrast, the Schmidt norm-based definition recovers a positive value for the overall energetic cost and an efficiency $\epsilon \leqslant \epsilon_C$, in compliance with thermodynamical laws.

VI. DISCUSSION

We demonstrated the potential and advantage of NN combined with AD in the field of quantum control of finite-time thermodynamics. Our method allowed us to discover driving protocols of the strokes of an Otto engine that perform better by a factor of 1.5 than previously conceived solutions. This scheme could find a nontrivial family of functions in which the first and second derivatives follow each other; from our results, we conclude that it is a crucial property of the cost function. In Ref. [61], we discuss an attempt to optimize a simpler ansatz, which can mimic the late blooming strategy of the NN. Our conclusion is that NN-based results are difficult to generate with a simple analytical form. Furthermore, employing the DP-ML optimization scheme enabled us to uncover a flaw in a previous definition for the cost of STA driving. In contrast, our modified definition provided physical results, which obeyed the Carnot bound on one hand and reached the adiabatic limit on the other hand. We point out that, among the plethora of energetic cost metrics suggested in the literature [70,73–77], we do not aim here to find which one is the most appropriate. However, ML optimization naturally identified violations to thermodynamical laws. Since our optimization method is general, it can be easily adapted to optimize other cost functions and figures of merits with little effort.

Our framework could be directly applied in other control problems, such as entropy reduction in closed systems [78], dynamical decoherence control [16], steering chemical reactions [79], and for the design of quantum electronic and thermal machines [44]. We paved the way for solving hard-constrained problems using state-of-the-art ML tools by orchestrating an objective as the minimum of a cost function. More generally, in this paper, we show that ML has an

advantage over some standard theoretical tools in designing quantum devices, thus making them favorable for an experimental realization. However, a benchmark against other contemporary optimal control methods, e.g., CRAB [11,80], is necessary to better understand the power of ML-based tools.

ACKNOWLEDGMENTS

We acknowledge fruitful discussions with Adolfo del Campo, Obinna Abah, Rodrigo A. Vargas-Hernández, Junjie Liu, and Aharon Brodutch. The work of I.K. was supported by the Centre for Quantum Information and Quantum Control at the University of Toronto. J.C. acknowledges support from the Natural Sciences and Engineering Research Council of Canada (NSERC), the Shared Hierarchical Academic Research Computing Network, Compute Canada, Google Quantum Research Award, the Canadian Institute for Advanced Research AI chair program, and companies sponsoring the Vector Institute. D.S. acknowledges support from an NSERC Discovery Grant and the Canada Research Chair program.

[1] C. Lanczos, *The Variational Principles of Mechanics* (University of Toronto Press, Toronto, 1970).

[2] A. Lipson, S. G. Lipson, and H. Lipson, *Optical Physics* (Cambridge University Press, Cambridge, 2010), 4th ed.

[3] D. A. Fedorov, B. Peng, N. Govind, and Y. Alexeev, VQE method: A short survey and recent developments Mater. Theory **6**, 2 (2022).

[4] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics: Emergence of Thermodynamic Behavior within Composite Quantum Systems* (Springer, Berlin, Heidelberg, 2009), 2nd ed., Vol. 784.

[5] S. C. Kaushik, S. K. Tyagi, and P. Kumar, *Finite Time Thermodynamics of Power and Refrigeration Cycles* (Springer International Publishing, Cham, Switzerland, 2018).

[6] M. Esposito, R. Kawai, K. Lindenberg, and C. Van den Broeck, Efficiency at Maximum Power of Low-Dissipation Carnot Engines, Phys. Rev. Lett. **105**, 150603 (2010).

[7] M. Esposito, R. Kawai, K. Lindenberg, and C. Van den Broeck, Finite-time thermodynamics for a single-level quantum dot, Europhys. Lett. **89**, 20003 (2010).

[8] M. I. Kamien and N. L. Schwartz, *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*, 2nd ed. (Dover Publications, Mineola, 2013), https://books.google.ca/books?id=liLCAgAAQBAJ.

[9] Y. Rezek, P. Salamon, K. H. Hoffmann, and R. Kosloff, The quantum refrigerator: The quest for absolute zero, Europhys. Lett. **85**, 30008 (2009).

[10] K. H. Hoffmann, P. Salamon, Y. Rezek, and R. Kosloff, Time-optimal controls for frictionless cooling in harmonic traps, Europhys. Lett. **96**, 60015 (2011).

[11] P. Doria, T. Calarco, and S. Montangero, Optimal Control Technique for Many-Body Quantum Dynamics, Phys. Rev. Lett. **106**, 190501 (2011).

[12] O. Abah and E. Lutz, Optimal performance of a quantum Otto refrigerator, Europhys. Lett. **113**, 60002 (2016).

[13] T. Fösel, P. Tighineanu, T. Weiss, and F. Marquardt, Reinforcement Learning with Neural Networks for Quantum Feedback, Phys. Rev. X **8**, 031084 (2018).

[14] Z. T. Wang, Y. Ashida, and M. Ueda, Deep Reinforcement Learning Control of Quantum Cartpoles, Phys. Rev. Lett. **125**, 100401 (2020).

[15] M. Y. Niu, S. Boixo, V. N. Smelyanskiy, and H. Neven, Universal quantum control through deep reinforcement learning, npj Quantum Inf. **5**, 33 (2019).

[16] R. Porotti, D. Tamascelli, M. Restelli, and E. Prati, Coherent transport of quantum states by deep reinforcement learning, Commun. Phys. **2**, 61 (2019).

[17] M. Dalgaard, F. Motzoi, J. J. Sørensen, and J. Sherson, Global optimization of quantum dynamics with AlphaZero deep exploration, npj Quantum Inf. **6**, 6 (2020).

[18] V. Nguyen, S. Orbell, D. T. Lennon, H. Moon, F. Vigneau, L. C. Camenzind, L. Yu, D. M. Zumbühl, G. A. D. Briggs, M. A. Osborne, and N. Ares, Deep reinforcement learning for efficient measurement of quantum devices, npj Quantum Inf. **7**, 100 (2021).

[19] V. B. Sørdal and J. Bergli, Deep reinforcement learning for quantum Szilard engine optimization, Phys. Rev. A **100**, 042314 (2019).

[20] N. D. Pozza, L. Buffoni, S. Martina, and F. Caruso, Quantum reinforcement learning: The maze problem, arXiv:2108.04490.

[21] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne *et al.*, JAX: composable transformations of Python+NumPy programs (2018), http://github.com/google/jax.

[22] A. Choromanska, M. Henaff, M. Mathieu, G. B. Arous, and Y. LeCun, The loss surfaces of multilayer networks, arXiv:1412.0233.

[23] F. Hioe, Theory of generalized adiabatic following in multilevel systems, Phys. Lett. A **99**, 150 (1983).

[24] T. A. Laine and S. Stenholm, Adiabatic processes in three-level systems, Phys. Rev. A **53**, 2501 (1996).

[25] A. A. Melnikov, H. Poulsen Nautrup, M. Krenn, V. Dunjko, M. Tiersch, A. Zeilinger, and H. J. Briegel, Active learning machine learns to create new quantum experiments, Proc. Natl. Acad. Sci. USA **115**, 1221 (2018).

[26] M. Bukov, A. G. R. Day, D. Sels, P. Weinberg, A. Polkovnikov, and P. Mehta, Reinforcement Learning in Different Phases of Quantum Control, Phys. Rev. X **8**, 031086 (2018).

[27] J. Wallnöfer, A. A. Melnikov, W. Dür, and H. J. Briegel, Machine learning for long-distance quantum communication, PRX Quantum **1**, 010301 (2020).

[28] C. Beeler, U. Yahorau, R. Coles, K. Mills, S. Whitelam, and I. Tamblyn, Optimizing thermodynamic trajectories using evolutionary and gradient-based reinforcement learning, Phys. Rev. E **104**, 064128 (2021).

[29] A. Jasinski, J. Montaner, R. C. Forrey, B. H. Yang, P. C. Stancil, N. Balakrishnan, J. Dai, R. A. Vargas-Hernández, and R. V. Krems, Machine learning corrected

quantum dynamics calculations, Phys. Rev. Research **2**, 032051 (2020).
[30] S. Deffner and S. Campbell, *Quantum Thermodynamics: An Introduction to the Thermodynamics of Quantum Information* (Morgan & Claypool Publishers, San Rafael, 2019).
[31] S. Bhattacharjee and A. Dutta, Quantum thermal machines and batteries, Eur. Phys. J. B **94**, 239 (2021).
[32] R. Kosloff and A. Levy, Quantum heat engines and refrigerators: continuous devices, Annu. Rev. Phys. Chem. **65**, 365 (2014).
[33] S. Vinjanampathy and J. Anders, Quantum thermodynamics, Contemp. Phys. **57**, 545 (2016).
[34] A. Das and V. Mukherjee, Quantum-enhanced finite-time Otto cycle, Phys. Rev. Research **2**, 033083 (2020).
[35] J. Klatzow, J. N. Becker, P. M. Ledingham, C. Weinzetl, K. T. Kaczmarek, D. J. Saunders, J. Nunn, I. A. Walmsley, R. Uzdin, and E. Poem, Experimental Demonstration of Quantum Effects in the Operation of Microscopic Heat Engines, Phys. Rev. Lett. **122**, 110601 (2019).
[36] R. B. S, V. Mukherjee, U. Divakaran, and A. del Campo, Universal finite-time thermodynamics of many-body quantum machines from Kibble-Zurek scaling, Phys. Rev. Research **2**, 043247 (2020).
[37] X. Chen, A. Ruschhaupt, S. Schmidt, A. del Campo, D. Guéry-Odelin, and J. G. Muga, Fast Optimal Frictionless Atom Cooling in Harmonic Traps: Shortcut to Adiabaticity, Phys. Rev. Lett. **104**, 063002 (2010).
[38] J. Deng, Q.-h. Wang, Z. Liu, P. Hänggi, and J. Gong, Boosting work characteristics and overall heat-engine performance via shortcuts to adiabaticity: quantum and classical systems, Phys. Rev. E **88**, 062122 (2013).
[39] E. Torrontegui, S. Ibáñez, S. Martínez-Garaot, M. Modugno, A. del Campo, D. Guéry-Odelin, A. Ruschhaupt, X. Chen, and J. G. Muga, Chapter 2: Shortcuts to adiabaticity, Adv. At. Mol. Opt. Phys. **62**, 117 (2013).
[40] X. Chen and J. G. Muga, Transient energy excitation in shortcuts to adiabaticity for the time-dependent harmonic oscillator, Phys. Rev. A **82**, 053403 (2010).
[41] J. G. Muga, X. Chen, S. Ibáñez, I. Lizuain, and A. Ruschhaupt, Transitionless quantum drivings for the harmonic oscillator, J. Phys. B: At., Mol. Opt. Phys. **43**, 085509 (2010).
[42] Y.-Y. Cui, X. Chen, and J. G. Muga, Transient particle energies in shortcuts to adiabatic expansions of harmonic traps, J. Phys. Chem. A **120**, 2962 (2016).
[43] O. Abah, J. Roßnagel, G. Jacob, S. Deffner, F. Schmidt-Kaler, K. Singer, and E. Lutz, Single-Ion Heat Engine at Maximum Power, Phys. Rev. Lett. **109**, 203006 (2012).
[44] A. C. Santos and M. S. Sarandy, Superadiabatic controlled evolutions and universal quantum computation, Sci. Rep. **5**, 15775 (2015).
[45] R. Kosloff and Y. Rezek, The quantum harmonic otto cycle, Entropy **19**, 136 (2017).
[46] K. Funo, N. Lambert, B. Karimi, J. P. Pekola, Y. Masuyama, and F. Nori, Speeding up a quantum refrigerator via counterdiabatic driving, Phys. Rev. B **100**, 035407 (2019).
[47] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, Energy consumption for shortcuts to adiabaticity, Rev. Mod. Phys. **91**, 045001 (2019).
[48] H. Zhou, Y. Ji, X. Nie, X. Yang, X. Chen, J. Bian, and X. Peng, Experimental Realization of Shortcuts to Adiabaticity in a Nonintegrable Spin Chain by Local Counterdiabatic Driving, Phys. Rev. Applied **13**, 044059 (2020).
[49] K. Ono, S. N. Shevchenko, T. Mori, S. Moriyama, and F. Nori, Analog of a Quantum Heat Engine Using a Single-Spin Qubit, Phys. Rev. Lett. **125**, 166802 (2020).
[50] L. Dupays, D. C. Spierings, A. M. Steinberg, and A. del Campo Delta-kick cooling, time-optimal control of scale-invariant dynamics, and shortcuts to adiabaticity assisted by kicks, Phys. Rev. Research **3**, 033261 (2021).
[51] J. Carrasquilla, Machine learning for quantum matter, Adv. Phys.: X **5**, 1797528 (2020).
[52] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, Automatic differentiation in machine learning: a Survey, J. Mach. Learning Res. **18**, 1 (2018).
[53] F. Schäfer, M. Kloc, C. Bruder, and N. Lörch, A differentiable programming method for quantum control, Mach. Learning: Sci. Technology **1**, 035009 (2020).
[54] A. del Campo, Shortcuts to Adiabaticity by Counterdiabatic Driving, Phys. Rev. Lett. **111**, 100502 (2013).
[55] M. Beau, J. Jaramillo, and A. Del Campo, Scaling-up quantum heat engines efficiently via shortcuts to adiabaticity, Entropy **18**, 168 (2016).
[56] L. Coopmans, D. Luo, G. Kells, B. K. Clark, and J. Carrasquilla, Protocol discovery for the quantum control of Majoranas by differentiable programming and natural evolution strategies, PRX Quantum **2**, 020332 (2021).
[57] J. Roßnagel, O. Abah, F. Schmidt-Kaler, K. Singer, and E. Lutz, Nanoscale Heat Engine Beyond the Carnot Limit, Phys. Rev. Lett. **112**, 030602 (2014).
[58] J. Roßnagel, S. T. Dawkins, K. N. Tolazzi, O. Abah, E. Lutz, F. Schmidt-Kaler, and K. Singer, A single-atom heat engine, Science **352**, 325 (2016).
[59] Y. Rezek and R. Kosloff, Irreversible performance of a quantum harmonic heat engine, New J. Phys. **8**, 83 (2006).
[60] M. V. Berry, Transitionless quantum driving, J. Phys. A: Math. Theor. **42**, 365303 (2009).
[61] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.4.L012029 for derivation of the energetic cost function, discussion of a additional ansatzes and a description of the neural network optimization.
[62] O. Abah and E. Lutz, Performance of shortcut-to-adiabaticity quantum engines, Phys. Rev. E **98**, 032121 (2018).
[63] K. Husimi, Miscellanea in elementary quantum mechanics, II, Prog. Theor. Phys. **9**, 381 (1953).
[64] M. A. Lohe, Exact time dependence of solutions to the time-dependent Schrödinger equation, J. Phys. A: Math. Theor. **42**, 35307 (2009).
[65] Y. Zheng, S. Campbell, G. De Chiara, and D. Poletti, Cost of counterdiabatic driving and work output, Phys. Rev. A **94**, 042132 (2016).
[66] S. Campbell and S. Deffner, Quantum Thermodynamics, Phys. Rev. Lett. **118**, 100601 (2017).
[67] D. P. Kingma and J. Ba (2014), ADAM: a method for stochastic optimization, arXiv:1412.6980 (2017).
[68] We use the nonadiabatic parameter $Q^*_{\rm NA} = \frac{\omega_1^2+\omega_2^2}{2\omega_1\omega_2}$, which alters Eqs. (3) and (4) (see Refs. [12,43] for details).

[69] O. Abah and E. Lutz, Energy efficient quantum machines, Europhys. Lett. **118**, 40005 (2017).

[70] O. Abah and M. Paternostro, Shortcut-to-adiabaticity Otto engine: a twist to finite-time thermodynamics, Phys. Rev. E **99**, 022110 (2019).

[71] O. Abah, M. Paternostro, and E. Lutz, Shortcut-to-adiabaticity quantum Otto refrigerator, Phys. Rev. Research **2**, 023120 (2020).

[72] Criticisms to this approach, albeit from a different aspect, were raised in Ref. [32].

[73] A. Del Campo, J. Goold, and M. Paternostro, More bang for your buck: Super-adiabatic quantum engines, Sci. Rep. **4**, 6208 (2014).

[74] K. Funo, J.-N. Zhang, C. Chatou, K. Kim, M. Ueda, and A. del Campo, Universal Work Fluctuations During Shortcuts to Adiabaticity by Counterdiabatic Driving, Phys. Rev. Lett. **118**, 100602 (2017).

[75] E. Torrontegui, I. Lizuain, S. González-Resines, A. Tobalina, A. Ruschhaupt, R. Kosloff, and J. G. Muga, Energy consumption for shortcuts to adiabaticity, Phys. Rev. A **96**, 022133 (2017).

[76] A. Tobalina, J. Alonso, and J. G. Muga, Energy consumption for ion-transport in a segmented Paul trap, New J. Phys. **20**, 065002 (2018).

[77] A. Tobalina, I. Lizuain, and J. G. Muga, Vanishing efficiency of a speeded-up ion-in-Paul-trap Otto engine, Europhys. Lett. **127**, 20005 (2019).

[78] P. Sgroi, G. M. Palma, and M. Paternostro, Reinforcement Learning Approach to Nonequilibrium Quantum Thermodynamics, Phys. Rev. Lett. **126**, 020601 (2021).

[79] A. F. de Almeida, R. Moreira, and T. Rodrigues, Synthetic organic chemistry driven by artificial intelligence, Nature Rev. Chem. **3**, 589 (2019).

[80] M. M. Müller, R. S. Said, F. Jelezko, T. Calarco, and S. Montangero, One decade of quantum optimal control in the chopped random basis, arXiv:2104.07687 (2021).