

Reinforcement Learning Assisted Quantum Simulation of Many-Body Excited States and Real-Time Dynamics

Jiaji Zhang ^{*}

*School of Optical Information and Energy Engineering,
Wuhan Institute of Technology, Wuhan, China*

Lipeng Chen 

Zhejiang Laboratory, Hangzhou 311100, China

Carlos L. Benavides-Riveros [†]

IQM Quantum Computers, Georg-Brauchle-Ring 23-25, 80992 Munich, Germany

(Dated: Last updated: May 19, 2026)

Abstract

The computation of electronic excited states and real-time quantum dynamics of many-fermion systems is among the most promising applications of near-term quantum computing. In this work, we generalize the reinforcement learning contracted quantum eigensolver (RL-CQE), previously developed for ground-state problems, to electronic excited states and real-time quantum dynamics, in which a deep Q-network agent adaptively selects the two-body operators at each iteration, yielding more compact ansätze and improved robustness with respect to critical hyperparameters. A key feature of the algorithm is a scalable state representation based on the ACSE residuals, whose dimension grows with the one-particle basis but remains independent of the number of targeted excited states. We also verify the equivalence of sign-free qubit operators in the excited-state setting, extending a result previously established for ground-state problems. Our RL-CQE for time evolution derives from a constant-scaling ansatz that represents the wave function with a fixed number of unitary transformations independent of simulation time t , enabled by the shared unitary structure of the purified ensemble treatment of excited states. Benchmarks on chemical systems demonstrate chemical accuracy with minimal operator counts across a range of bond lengths.

I. INTRODUCTION

Electronic excited states play a central role in a wide range of physical and chemical processes that lie beyond the reach of ground-state or equilibrium descriptions, including photoexcited dynamics in light-harvesting complexes [1–3], non-equilibrium charge transport driven by laser fields [4–7], and nonlinear response phenomena in advanced quantum materials [8–10]. Computing these states accurately poses fundamental challenges: they often exhibit strong multi-configurational character [11, 12] and near-degeneracy at conical intersections [13, 14], features that render standard single-reference methods unreliable. Implementing robust excited-state solvers on near-term quantum devices therefore represents one of the most compelling and practically urgent applications of quantum algorithms.

Among the algorithms developed to date, the contracted quantum eigensolver (CQE), which is based on the contracted Schrödinger equation, has attracted considerable attention

* Authors to whom correspondence should be addressed: zhang.jiaji.i92@kyoto-u.jp

† Authors to whom correspondence should be addressed: carlos.benavides@iqm.tech

[15–19]. A key advantage of CQE is that the wavefunction is parameterized entirely through two-body operators, yielding favorable scaling with the number of orbitals and electrons [15]. Although originally formulated for ground-state problems, CQE has been extended to excited states through a purified ensemble approach, which recasts the excited-state problem as an effective ground-state problem on an enlarged Hilbert space [19, 20]. While this ensemble formulation naturally inherits improvements developed for ground-state CQE, it exhibits slower convergence due to the simultaneous optimization of multiple states, and its numerical performance is sensitive to several critical hyperparameters (most notably the ensemble weight vector) which limits its applicability to generic systems [19].

The rapid development of artificial intelligence and deep learning has dramatically accelerated progress across the fundamental sciences [21–26]. Among its subfields, reinforcement learning (RL) trains an agent to make sequential decisions in complex, potentially black-box environments [27–30]. Unlike supervised and unsupervised learning, in which a model is passively trained on fixed datasets to minimize a predefined loss, an RL agent learns actively through exploration and exploitation, maximizing a cumulative reward without requiring prior knowledge of the problem structure. This model-free character has made RL a powerful tool across a broad range of real-world tasks, including game playing [31–33], robotic control [34–36], autonomous driving [37–39], and large language models [40–42].

The model-free nature of RL makes it particularly well suited for hybrid quantum-classical algorithms, where the quantum device can naturally be treated as a black-box environment whose internal details need not be explicitly modeled [26]. RL has already been applied in this context to quantum eigensolvers and real-time quantum control [43–49], and has been shown to discover optimal operator sequences requiring fewer circuit elements than conventional optimization methods [43]. Most recently, Wang and Mazziotti [45] proposed an RL-CQE algorithm for ground-state problems, reformulating the CQE update procedure as a Markov decision process and training a deep Q-network to select two-body operators at each iteration, demonstrating significant reductions in circuit depth relative to conventional CQE implementations. Despite these advances, the application of RL to the direct optimization of excited-state wavefunctions within a quantum eigensolver framework has, to the best of our knowledge, not yet been explored.

In this work, we generalize the RL-CQE algorithm to excited-state problems of quantum chemistry Hamiltonians, building on the purified ensemble approach in which the wavefunc-

tion is parameterized as a sequence of unitary transformations applied to a set of orthonormal reference states [20]. While sharing the same conceptual foundation as the ground-state RL-CQE of Ref. [45], our formulation introduces several key technical advances tailored to the excited-state setting. Specifically, it substantially reduces per-iteration measurement overhead, provides a scalable state representation whose complexity is independent of the number of targeted excited states, and yields solutions that are markedly more robust with respect to the ensemble weight vector. In addition, we establish for the first time the equivalence of sign-free qubit operators (previously validated only for ground-state problems [16]) in the excited-state setting. We benchmark the algorithm on H_2 and linear H_3^+ , demonstrating compact solutions with minimal operator counts while maintaining chemical accuracy across a range of bond lengths.

We further extend the RL-CQE framework to the quantum simulation of time-dependent wavefunctions, a long-standing challenge under limited quantum resources. Conventional approaches decompose the time-evolution operator into sequential short-time propagators, leading to circuit depths that grow without bound as the simulation time increases. To overcome this limitation, we introduce an algorithm that maintains a fixed ansatz size regardless of simulation time. The key insight is to expand the time-dependent wavefunction in the basis of Hamiltonian eigenstates, which are naturally shared across all targeted states within the purified ensemble framework. This shared structure allows the time-dependent wavefunction to be expressed as a fixed set of unitary transformations acting on a time-dependent reference state, which is itself compactly represented with the aid of RL. The resulting ansatz achieves a constant total operator count as a function of simulation time, opening new possibilities for scalable many-body dynamics on near-term quantum hardware.

The remainder of this paper is organized as follows. Section II introduces the theoretical background, covering the CQE formalism, its extension to excited states via the purified ensemble approach, the RL agent architecture, and the generalization to time evolution. Section III presents numerical results on the benchmark systems. Section IV summarizes the main findings and outlines directions for future research.

II. THEORY

We introduce the generalization of RL-CQE to excited state problem in Section II A, present the deep-Q network in II B, which are utilized as the main body of agent in this work, and discuss the generalization to the time evolution problem in Section II C.

A. RL-CQE for Excited States

We consider an electronic system described by the Hamiltonian

$$\hat{H} = \sum_{pq} h_{pq} \hat{c}_p^\dagger \hat{c}_q + \frac{1}{2} \sum_{pqkl} g_{pq,kl} \hat{c}_p^\dagger \hat{c}_q^\dagger \hat{c}_k \hat{c}_l, \quad (1)$$

where $\hat{c}_p^\dagger$ and $\hat{c}_p$ are the fermionic creation and annihilation operators for the p -th spin-orbital, and h_{pq} and $g_{pq,kl}$ are the one- and two-electron integrals, respectively. The ν -th eigenstate $|\psi_\nu\rangle$ can be obtained by enforcing the anti-Hermitian contracted Schrödinger equation (ACSE) [15],

$$\langle \psi_\nu | [\hat{\Gamma}_{kl}^{pq}, \hat{H}] | \psi_\nu \rangle = 0, \quad (2)$$

where $\hat{\Gamma}_{kl}^{pq} = \hat{c}_p^\dagger \hat{c}_q^\dagger \hat{c}_k \hat{c}_l$ is the two-body operator. To solve the ACSE iteratively, we adopt an ansatz of the form [15, 45]

$$|\psi_\nu^{(n+1)}\rangle = e^{\theta^{(n)} \hat{A}^{(n)}} |\psi_\nu^{(n)}\rangle, \quad (3)$$

where $\hat{A}^{(n)}$ is an anti-Hermitian operator and $\theta^{(n)}$ is a real-valued scalar at the n -th iteration. The initial reference states at $n = 0$ are chosen to be mutually orthogonal Hartree–Fock states satisfying $\langle \psi_\mu^{(0)} | \psi_\nu^{(0)} \rangle = \delta_{\mu\nu}$, a property that is preserved by all subsequent unitary transformations.

To target multiple eigenstates simultaneously, the Rayleigh–Ritz variational principle provides a variational lower bound for the weighted ensemble energy [19],

$$\sum_{\nu=0}^{K-1} w_\nu \langle \psi_\nu | [\hat{\Gamma}_{kl}^{pq}, \hat{H}] | \psi_\nu \rangle \geq 0, \quad (4)$$

where the weights $\{w_\nu\}$ are a fixed set of positive values that are normalized, $\sum_{\nu=0}^{K-1} w_\nu = 1$, and strictly decreasingly ordered, $w_\nu \geq w_{\nu+1} > 0$.

The excited-state CQE with exponential ansatz maps naturally onto an Markov decision process, which lies in the theoretical foundation of RL. Within this framework, the CQE

update procedure and its associated quantum algebra are regarded as the RL environment. The anti-Hermitian operators $\hat{A}^{(n)}$ applied at each updating iteration are regarded as RL action. In this work, we utilize the sign-free two-qubit operators as the RL action, defined in Ref. [16]:

$$\hat{\gamma}_{kl}^{pq} = \hat{\sigma}_p^+ \hat{\sigma}_q^+ \hat{\sigma}_k^- \hat{\sigma}_l^-, \quad (5)$$

where $\hat{\sigma}_p^\pm = \hat{\sigma}_p^x \pm i\hat{\sigma}_p^y$ is the raising/lowering operator on the p -th qubit, with $\hat{\sigma}^x$ and $\hat{\sigma}^y$ the standard Pauli matrices. These operators differ with the original two-body operators $\hat{\Gamma}_{kl}^{pq}$ up to a non-local sign factor that arises from the encoding of fermionic operators as qubits [50, 51]. Although $\hat{\gamma}_{kl}^{pq}$ does not strictly respect fermionic statistics, it has been demonstrated to yield the same convergence behavior as the original fermionic operators in the ground state problem [16]. Here we extend this equivalence to the more generalized excited state cases.

At each updating iteration, an RL agent follows a policy that will be parameterized as a neural network to select one of the operators $\hat{\gamma}_{kl}^{pq}$ as the RL action based on the RL state. The RL state, which fully characterizes the CQE, is represented by the vector of CSE residuals $\mathbf{r}^{(n)} = \{r_{pq,kl}^{(n)}\}$, which is defined through $\hat{\gamma}_{kl}^{pq}$ as

$$r_{pq,kl}^{(n)} = \sum_{\nu=0}^{K-1} w_\nu \langle \psi_\nu^{(n)} | [\hat{\gamma}_{kl}^{pq}, \hat{H}] | \psi_\nu^{(n)} \rangle. \quad (6)$$

Crucially, the dimension of $\mathbf{r}^{(n)}$ depends only on the one-particle basis size and is entirely independent of the number of excited states K . This is a non-trivial scalability advantage over any representation based directly on the K wavefunctions $\{|\phi_\nu^{(n)}\rangle\}$, whose combined description would grow with K . In addition, the variational principle for excited states, defined in Eq.(4), can be generalized to the sign-free operators $\hat{\gamma}_{kl}^{pq}$, as already shown in the ground state problem [16], which provides a variational lower bound for CSE residuals. Once the action $\hat{\gamma}^{(n)}$ has been selected, we solve the scalar factor $\theta^{(n)}$ from the one-dimensional optimization,

$$\theta^{(n)} = \operatorname{argmin}_\theta \sum_{\nu=0}^{K-1} w_\nu \langle \psi_\nu^{(n)} | e^{-\theta \hat{A}^{(n)}} \hat{H} e^{\theta \hat{A}^{(n)}} | \psi_\nu^{(n)} \rangle, \quad (7)$$

with $\hat{A}^{(n)} = \hat{\gamma}^{(n)} - \hat{\gamma}^{(n)\dagger}$ to ensure unitarity.

B. Deep-Q Network

The RL policy introduced in the previous subsection is parameterized using a deep neural network, placing our algorithm within the framework of deep reinforcement learning [29]. Specifically, we adopt the deep Q-network (DQN) [52–54], one of the most successful value-based deep RL algorithms, which has already been applied to quantum computing tasks [43–45]. DQN employs a neural network to approximate the state-action value function, also known as the Q-function, within the standard Q-learning framework.

We denote by $\pi(a_n | s_n)$ the probability of selecting action a_n when the environment is in state s_n at step n . The corresponding Q-function is defined as the expected cumulative reward under policy π ,

$$Q(s, a) = \mathbb{E}_\pi[R_{n+1} + \eta Q(s_{n+1}, a_{n+1}) | s_n = s, a_n = a], \quad (8)$$

where R_n is the immediate reward at step n and $0 < \eta < 1$ is a discount factor that controls the weight assigned to future rewards and ensures convergence of the recursion. The Q-function is updated iteratively during training according to

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \left[R_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right], \quad (9)$$

where α is the learning rate. In DQN, the Q-function is approximated by a neural network $Q_\theta(s, a)$ with learnable parameters θ , trained by minimizing the mean-squared Bellman error implied by Eq. (9) via gradient descent.

The reward signal is defined as

$$R_n = - \sum_{\nu=0}^{K-1} w_\nu \langle \psi_\nu^{(n)} | \hat{H} | \psi_\nu^{(n)} \rangle - \lambda \|\mathbf{r}^{(n)}\|, \quad (10)$$

where $\|\cdot\|$ denotes the Euclidean norm of the residual vector, and $\lambda \geq 0$ is an adjustable regularization parameter that controls the relative weight between energy minimization and residual suppression.

C. RL-CQE for Time Evolution

Encoding the time evolution $|\Psi(t)\rangle = e^{-i\hat{H}t}|\Psi(0)\rangle$ as a sequence of unitary transformations has long been a fundamental challenge in quantum simulation. Conventional

algorithms based on Trotterization decompose the propagator into sequential short-time unitaries, leading to circuit depths that can grow without bound as the simulation time increases. In this subsection, we introduce a time-evolution algorithm based on the excited-state RL-CQE framework that requires only a fixed number of unitaries, so that the total computational cost remains constant in t .

We begin from the iterative ansatz of Eq. (3) and assume that the Hamiltonian eigenstates can be obtained by sequentially applying N unitary transformations to the reference Hartree–Fock states, $|\psi_\nu\rangle = \hat{U}_N \cdots \hat{U}_1 |\psi_\nu^{(0)}\rangle$, where $\hat{U}_n$ is the n -th unitary defined by the right-hand side of Eq. (3). We then expand the time-dependent wavefunction in the basis of Hamiltonian eigenstates,

$$|\Psi(t)\rangle = \sum_{\nu=0}^{K-1} c_\nu(t) |\psi_\nu\rangle = \hat{U}_N \cdots \hat{U}_1 |\phi^{(0)}(t)\rangle, \quad (11)$$

where

$$c_\nu(t) = e^{-iE_\nu t} \langle \psi_\nu | \Psi(0) \rangle \quad (12)$$

are the time-dependent expansion coefficients, and

$$|\phi^{(0)}(t)\rangle = \sum_{\nu=0}^{K-1} c_\nu(t) |\psi_\nu^{(0)}\rangle \quad (13)$$

is a time-dependent superposition of the reference Hartree–Fock states. Equation (11) is the central object of our time-dependent treatment: because all eigenstates can be obtained simultaneously using the *same* set of unitary transformations $\{\hat{U}_n\}$, that shared set can be reused directly in the construction of the time-dependent wavefunction. This is in sharp contrast to state-specific treatments of excited states, in which each eigenstate is associated with a different set of unitaries and orthogonality is enforced by additional penalty terms. The number of required unitaries N is constant in t and is determined solely by the number of eigenstates K that have non-zero overlap with the initial state $|\Psi(0)\rangle$.

The ansatz in Eq. (11) requires prior knowledge of all relevant Hamiltonian eigenstates and the explicit construction of the reference state $|\phi^{(0)}(t)\rangle$, which is, as a result, inconvenient in practice. To arrive at a more practical form, we introduce a secondary set of time-dependent unitaries that prepares $|\phi^{(0)}(t)\rangle$ from a simple fixed reference state, chosen here to be the Hartree–Fock ground state $|\phi_0\rangle$,

$$|\phi^{(0)}(t)\rangle = \hat{U}'_{N'}(t) \cdots \hat{U}'_1(t) |\phi_0\rangle. \quad (14)$$

Because $|\phi^{(0)}(t)\rangle$ is a superposition of simple Hartree–Fock states, the number of secondary unitaries N' is only affected by the number K , but is nearly independent of t . Combining Eqs. (11) and (14), we conclude that any $|\Psi(t)\rangle$ can be prepared from the fixed reference $|\phi_0\rangle$ by a total of $N + N'$ unitary transformations, a number that remains constant throughout the simulation.

In practice, both sets of unitaries are determined within the RL-CQE framework with a minor modification to the reward function. Adopting the same ansatz form as Eq. (3),

$$|\phi^{(n;t)}\rangle = e^{\theta^{(n;t)}\hat{A}^{(n;t)}} |\phi^{(n-1;t)}\rangle, \quad |\phi^{(0;t)}\rangle \equiv |\phi_g\rangle, \quad (15)$$

we proceed as follows. Given the wavefunction $|\Psi(t)\rangle$ at time t , we first construct the target state

$$|\chi\rangle = e^{-i\hat{H}\delta_t} |\Psi(t)\rangle, \quad (16)$$

where δ_t is a small time step. The operators $\hat{A}^{(n,t+\delta_t)}$ and scalar parameters $\theta^{(n,t+\delta_t)}$ are then selected by the same RL-CQE procedure detailed in the preceding subsections, including the same choice of RL state and RL action, but with the reward replaced by the fidelity,

$$R'_n = |\langle\chi|\phi^{(n;t)}\rangle|^2. \quad (17)$$

The algorithm is run until the fidelity approaches unity, at which point the wavefunction at the next time step is obtained as $|\Psi(t + \delta_t)\rangle = |\phi^{(M;t)}\rangle$, where M is the number of iterations required. Based on the analysis above, we can conclude that $M \leq N + N'$, and likewise remains constant in t . The utilization of RL allows us to obtain more efficient form with less number of unitaries. We note that the ACSE residual of Eq. (6) remains a valid state descriptor for the time-evolution problem, as established in the context of the time-dependent contracted Schrödinger equation [55].

Before closing this subsection, we briefly compare the present RL-CQE time-evolution algorithm with the recently developed correlation-based efficient time-evolution (CETE) algorithm of Rose et al. [56]. In Ref. [56], the time-dependent contracted Schrödinger equation [55] is adopted as the theoretical foundation, and the time-evolution problem is recast as a sequence of time-independent problems. By employing an iterative ansatz analogous to Eq. (15) and updating it via fidelity gradients in analogy with conventional CQE, CETE achieves an ansatz whose computational cost remains constant in t . The present work arrives at an equivalent constant-scaling ansatz form from a different theoretical starting point,

namely the purified ensemble treatment of excited states, which naturally provides a shared set of unitary transformations across all targeted eigenstates. Beyond the theoretical difference, the RL-assisted update strategy employed here offers several practical advantages over the gradient-based approach of CETE: it requires no gradient evaluation, is less susceptible to local minima, and adaptively selects operators based on the global structure of the optimization landscape, as demonstrated for the ground- and excited-state problems in the preceding subsections.

III. RESULTS AND DISCUSSION

In this section, we benchmark RL-CQE on the H_2 and linear equidistant H_3^+ molecules for both excited-state energies and time evolution, which represent two of the most widely studied benchmark systems for the quantum simulation of electronic excited states [57]. The one- and two-electron integrals are computed using the Slater-type orbital basis STO-6G as implemented in PySCF [58]. The molecular Hamiltonians are mapped to 4 and 6 qubits, respectively, via the Jordan–Wigner transformation, and all quantum simulations are performed using a noiseless statevector simulator.

The DQN is implemented as a feedforward neural network with 8 linear layers, 512 hidden channels, and Gaussian error linear unit (GELU) activation functions. Training follows the standard DQN workflow with an experience replay buffer of capacity 10^6 transitions, a discount factor $\eta = 0.99$, and a residual regularization weight $\lambda = 0.5$ in the reward function of Eq. (10). The maximum number of RL steps per episode is set to 5 unless otherwise specified. The network is optimized using the AdamW optimizer with a learning rate of 2×10^{-4} , a batch size of 256, and 3000 training episodes. All remaining hyperparameters follow the default values of the PyTorch library [59]. The source code for this work is publicly available on GitHub.

A. Excited state of H_2

We first benchmark RL-CQE on the H_2 molecule, targeting the lowest four eigenstates within the spin sector $\langle \hat{S}_z \rangle = 0$. Following Ref. [19], we adopt the weight vector $\mathbf{w} \propto [9, 9, 1, 1]$ prior to normalization. While these weights were reported to be optimal for con-

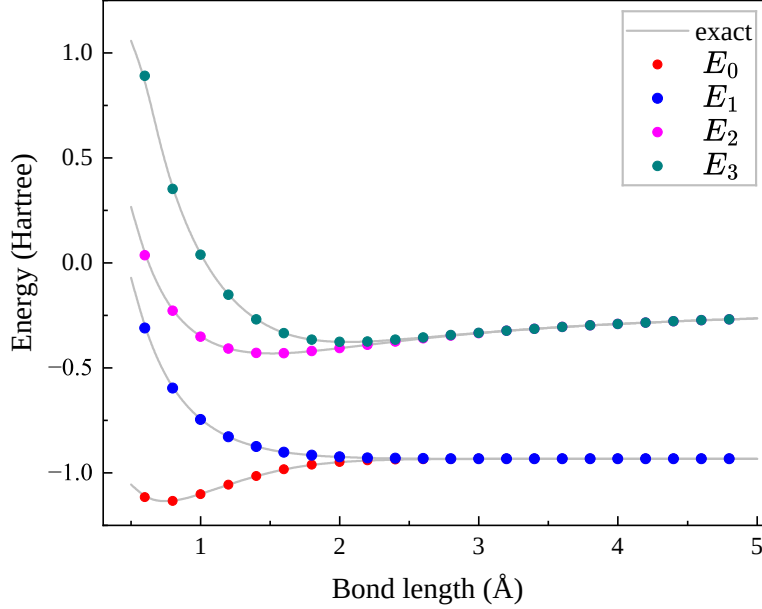


FIG. 1. The energy eigenvalues of H_2 within spin sector $\langle \hat{S}^z \rangle = 0$ at bond length 0.5 Å to 5.0 Å. The results from RL-CQE are represented by the colored dot, while the exact results from full CI are represented by the corresponding colored lines.

ventional CQE, we will show that RL-CQE is largely insensitive to this choice. Figure 1 shows the potential energy curves of H_2 at various bond lengths, with exact results obtained from full configuration interaction (FCI) calculations. With at most 5 unitary transformations, RL-CQE achieves energies within 10^{-3} Hartree of the FCI reference across all geometries, demonstrating both the accuracy of the algorithm and the effectiveness of sign-free qubit operators as the building blocks of the exponential ansatz.

One of the key features of RL-CQE is its ability to minimize the number of operators by adjusting the upper limit of RL steps. To illustrate this, we select four representative bond lengths: (i) 0.5 Å (repulsive region); (ii) 0.7 Å (near equilibrium); (iii) 1.2 Å (intermediate region); and (iv) 3.0 Å (dissociation limit). For all four cases, RL-CQE calculations are performed with a maximum of 2 steps, with all other settings unchanged from the previous subsection.

Figure 2 shows the Euclidean norm of the ACSE residual $\|\mathbf{r}^{(n)}\|$ as a function of the RL step n , as defined in Eq. (3). In all cases, RL-CQE reproduces the correct result using only 2 operators, and further increasing the number of steps yields no improvement, as the residual

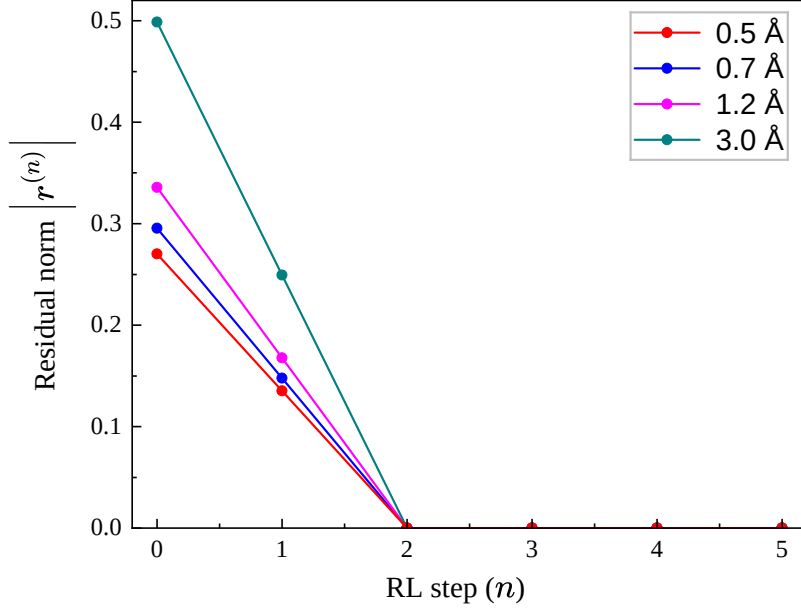


FIG. 2. Norm of CSE residual $|\mathbf{r}^{(n)}|$ against RL steps at different bond length. All the parameters are the same as Fig.1.

norm drops to near zero ($\sim 10^{-8}$), at which point the calculation can be terminated.

This compactness can be understood by contrasting RL-CQE with conventional CQE, which follows a fixed algorithmic update that applies all two-body operator coefficients simultaneously. This rigid update strategy may occasionally deviate from the direction of steepest descent of the energy and ACSE norm, resulting in a less compact ansatz that requires more operators. In contrast, the RL-based approach adaptively selects operators at each step, consistently following the steepest descent direction and thereby producing near-optimal ansätze with a minimal operator count.

Another key feature of RL-CQE is its robustness with respect to the choice of the ensemble weight vector. In the calculations above, we adopted the weight vector $\mathbf{w} \propto [9, 9, 1, 1]$, which was determined by analyzing the point-group symmetry of the system and shown to be optimal for conventional CQE. However, such optimal values are not always straightforward to determine for more complex systems, and suboptimal choices can require significantly more operators to converge in conventional CQE.

To assess the sensitivity of RL-CQE to this choice, we fix the bond length at 0.7 Å and apply RL-CQE with the following weight vectors prior to normalization: (i) $[9, 9, 1, 1]$; (ii)

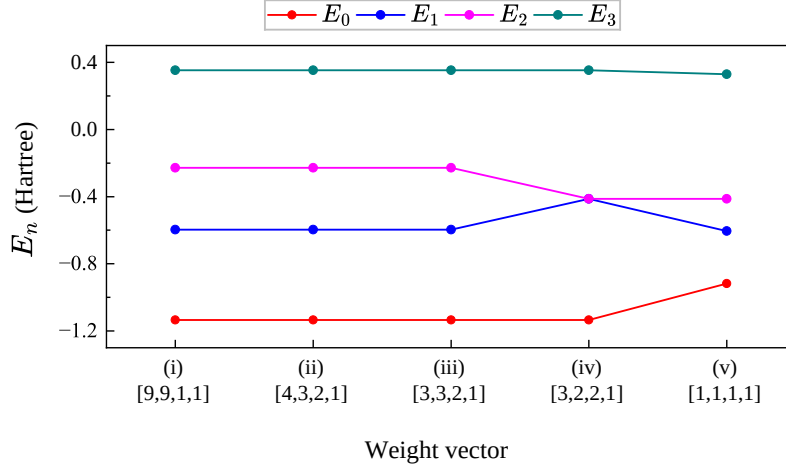


FIG. 3. The energy eigenvalues obtained from RL-CQE with different weight vectors. Using weight vector (i), (ii), and (iii) lead to correct results, whereas some of the eigenvalues in (iv) and (v) are incorrect due to the improper choice of weight value.

[4, 3, 2, 1]; (iii) [3, 3, 2, 1]; (iv) [3, 2, 2, 1]; and (v) [1, 1, 1, 1]. The converged energy eigenvalues for each case are shown in Fig. 3.

We find that a simple strictly descending choice such as (ii) [4, 3, 2, 1] yields accuracy comparable to the optimal vector (i) [9, 9, 1, 1]. Assigning identical weights to states that share a symmetry-protected degeneracy, as in case (iii) [3, 3, 2, 1], can slightly improve training stability without compromising the final accuracy. However, assigning identical weights to states that do not share such a degeneracy, as in case (iv) [3, 2, 2, 1] where the 2nd and 3rd states receive the same weight, causes RL-CQE to converge to a mixture of the corresponding eigensubspace rather than the individual eigenstates. In this situation, a subspace diagonalization must be performed to recover the correct eigenstates. Similarly, the uniform vector (v) [1, 1, 1, 1] induces mixing across all eigensubspaces, leading to deviations in all eigenvalues. These results suggest that a strictly descending weight vector such as [4, 3, 2, 1] is a robust and practical default choice for RL-CQE, requiring no prior knowledge of the system symmetry.

B. Excited state of H_3^+

We next benchmark RL-CQE on the linear equidistant H_3^+ molecule, a widely used test case for strong electron correlation. We target the lowest four eigenstates within the spin sector $\langle \hat{S}_z \rangle = 0$, using the weight vector $\mathbf{w} \propto [4, 3, 2, 1]$ prior to normalization and a maximum of 15 RL steps. Fig. 4(a) shows the potential energy curve at various H–H bond lengths after training for 10^6 episodes, with exact results obtained from numerical diagonalization. RL-CQE yields energies in close agreement with the exact reference across all geometries. Fig. 4(b) shows the convergence of the ACSE residual norm $\|\mathbf{r}^{(n)}\|$ as a function of RL step n at two representative geometries: (i) 1.7 Å (near equilibrium) and (ii) 2.9 Å (dissociation limit). The first five operators $\hat{\gamma}^{(1)}, \dots, \hat{\gamma}^{(5)}$ and their corresponding parameters $\theta^{(1)}, \dots, \theta^{(5)}$ (see Eq. (3)) are listed explicitly in Table III B. In both cases, RL-CQE selects distinct operator sequences adapted to the local electronic structure at each geometry, while maintaining high accuracy and a compact operator representation in both regimes.

RL step (n)	1	2	3	4	5
1.7 Å, $\hat{\gamma}^{(n)}$	$\hat{\gamma}_{56}^{12}$	$\hat{\gamma}_{36}^{14}$	$\hat{\gamma}_{36}^{23}$	$\hat{\gamma}_{45}^{04}$	$\hat{\gamma}_{56}^{34}$
1.7 Å, $\theta^{(n)}$	0.1216	-0.2420	0.2762	0.1692	0.2631
2.9 Å, $\hat{\gamma}^{(n)}$	$\hat{\gamma}_{56}^{12}$	$\hat{\gamma}_{45}^{23}$	$\hat{\gamma}_{45}^{14}$	$\hat{\gamma}_{36}^{23}$	$\hat{\gamma}_{56}^{34}$
2.9 Å, $\theta^{(n)}$	0.3758	-0.6163	0.780	0.3177	0.8851

TABLE I. First 5 operators and factors of ansatz Eq.(3) obtained from RL-CQE at H-H distance 1.7 Å and 2.9 Å, respectively.

C. Time evolution problem

Finally, we apply RL-CQE to the time evolution of H_2 and H_3^+ . The initial state $|\Psi(0)\rangle$ is chosen as a superposition of Hartree–Fock states within the spin sector $\langle \hat{S}_z \rangle = 0$ with randomly chosen coefficients. The time evolution is evaluated over $t \in [0, 20]$ a.u. with a time step $\delta_t = 0.05$ a.u., following the iterative RL-CQE procedure described in Section II C. The reference wavefunction $|\Psi_{\text{ref}}(t)\rangle$ is obtained from exact numerical diagonalization of the Hamiltonian. Fig. 5 shows the fidelity $F_n(t) = |\langle \Psi_{\text{ref}}(t) | \phi^{(n;t)} \rangle|$, as a function of RL step n

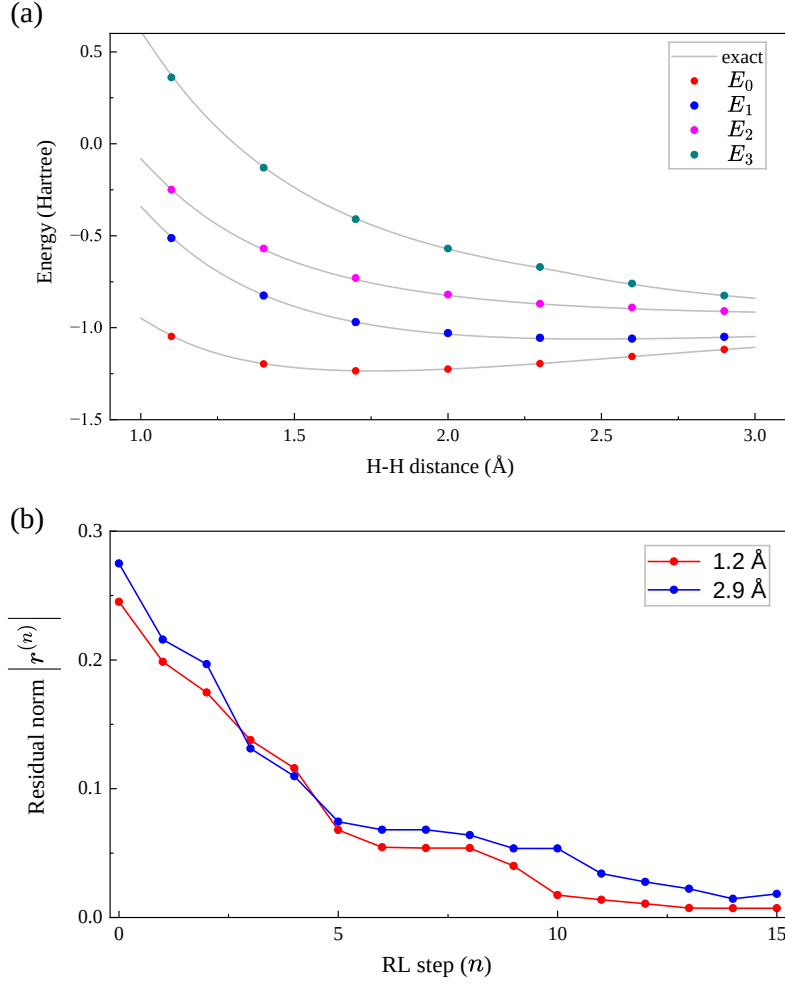


FIG. 4. (a) The energy eigenvalues of linear H_3^+ as a function of H-H distance at 1.0 Å to 3.0 Å. Only lowest 4 singlet states with $\langle \hat{S}^z \rangle = 0$ are presented. (b) Norm of CSE residual $|\mathbf{r}^{(n)}|$ against RL steps at different bond length.

at selected time points. Across the entire time range, $F_n(t)$ converges to near unity within (a) 5 steps for H_2 and (b) 20 steps for H_3^+ . Crucially, the number of steps required remains constant throughout the simulation, and is comparable to that needed for the convergence of the excited-state calculations. We note that $F_n(t)$ exhibits sharp jumps at intermediate RL steps for H_3^+ , which arise because only the absolute value of the overlap is considered, with the global phase factor neglected. These results demonstrate the ability of RL-CQE to generate compact operator sequences of fixed length for time-dependent wavefunctions, consistent with the theoretical analysis of Section II C.

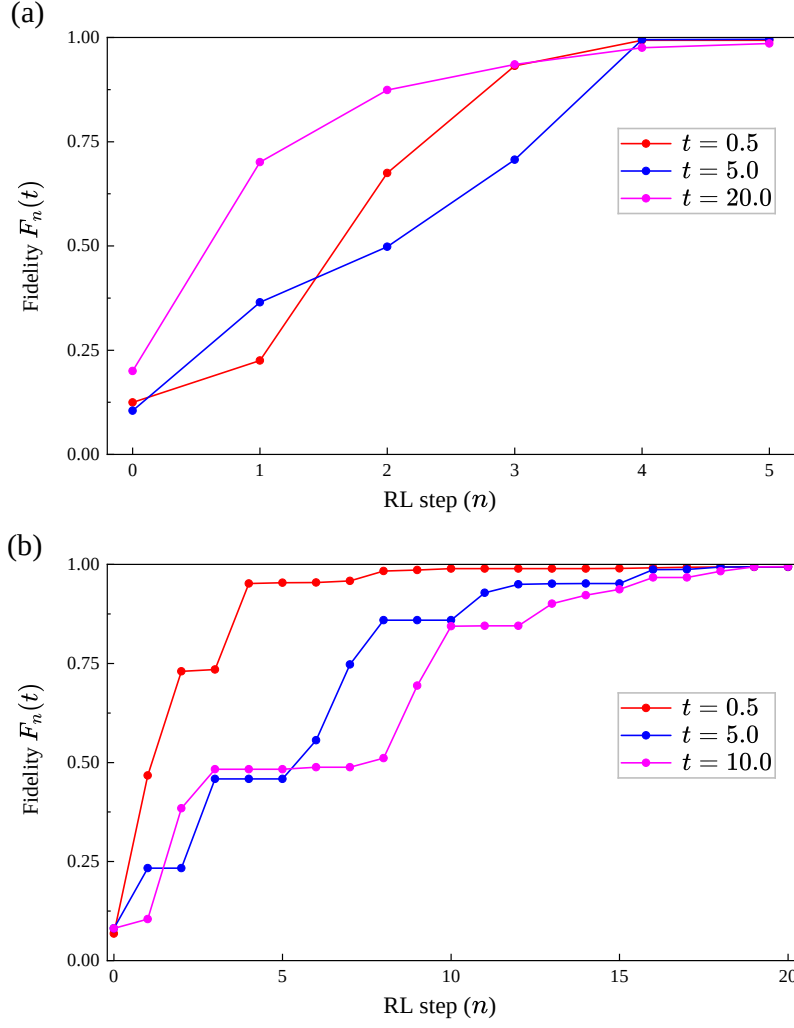


FIG. 5. The time-dependent fidelity $F_n(t)$ against RL step n at different time points for (a) H_2 and (b) H_3^+ molecules, respectively.

IV. CONCLUSION

In this work, we have generalized the reinforcement learning contracted quantum eigensolver (RL-CQE), previously developed for ground-state problems, to electronic excited states and real-time quantum dynamics. The algorithm updates the wavefunction through a sequence of two-body exponential transformations, with an RL agent selecting the operator to be applied at each step from a fixed operator pool. This one-operator-per-step strategy yields more compact ansätze and more robust convergence with respect to the ensemble weight vector compared to conventional CQE, which updates all operator coefficients.

cients simultaneously. We have also verified the equivalence of sign-free qubit operators in the excited-state setting, extending a result previously established for ground-state problems, and introduced a scalable state representation based on the ACSE residuals whose dimension is independent of the number of targeted states.

Building on the excited-state formulation, we have further generalized RL-CQE to the simulation of time-dependent wavefunctions. The key insight is to leverage the shared unitary structure of the purified ensemble treatment, in which all targeted eigenstates are represented by the same set of unitary transformations. This allows the time-dependent wavefunction to be expressed with a fixed number of operators independent of the simulation time t , in contrast to conventional Trotterization-based methods whose circuit depth grows without bound.

Benchmarks on molecular systems across a range of bond lengths demonstrate that RL-CQE achieves chemical accuracy with minimal operator counts for both excited-state energies and time-evolved wavefunctions. The number of operators required for time evolution remains constant throughout the simulation and is comparable to that needed for the excited-state calculations, confirming the theoretical constant-scaling prediction. Taken together, these results demonstrate that RL provides an effective and flexible framework for minimizing quantum resource requirements in hybrid quantum-classical algorithms, with implications for near-term quantum simulation beyond what is accessible by conventional methods.

Several directions remain open for future work. One natural extension is the generalization to arbitrary quantum state preparation [60, 61], where designing efficient and resource-minimal protocols remains a significant challenge. The adaptive operator selection strategy of RL-CQE is well suited to address this problem, as it navigates complex optimization landscapes without requiring gradient information or prior knowledge of the problem structure. More broadly, the present framework can in principle be extended to open quantum systems [62], non-equilibrium dynamics driven by external fields, fermion-boson mixtures [63], and larger molecular systems where conventional methods become intractable, making RL-assisted quantum eigensolvers a promising avenue for future development.

ACKNOWLEDGMENTS

J.Z. and L.C. acknowledge the support from the National Natural Science Foundation of China (Grant No. 22473101). C.L.B.R. gratefully acknowledges the financial support from the Royal Society of Chemistry and the European Union’s Horizon Europe Research and Innovation program under the Marie Skłodowska-Curie Grant Agreement No. 101065295-RDMFTforbosons.

DATA AVAILABILITY

The data and codes that support the findings of this article are openly available at the following URL/DOI: <https://github.com/jiaji-zhang/RL-CSE>. All codes to reproduce, examine, and improve our proposed analysis are freely available online.

-
- [1] R. E. Blankenship, *Molecular Mechanisms of Photosynthesis* (Wiley, 2002).
 - [2] J. M. Artes Vivancos, I. H. M. van Stokkum, F. Saccon, Y. Hontani, M. Klotz, A. Ruban, R. van Grondelle, and J. T. M. Kennis, Unraveling the excited-state dynamics and light-harvesting functions of xanthophylls in light-harvesting complex ii using femtosecond stimulated raman spectroscopy, *Journal of the American Chemical Society* **142**, 17346–17355 (2020).
 - [3] G. D. Scholes, G. R. Fleming, A. Olaya-Castro, and R. van Grondelle, Lessons from nature about solar light harvesting, *Nature Chemistry* **3**, 763–774 (2011).
 - [4] J. E. Greenwald, J. Cameron, N. J. Findlay, T. Fu, S. Gunasekaran, P. J. Skabara, and L. Venkataraman, Highly nonlinear transport across single-molecule junctions via destructive quantum interference, *Nature Nanotechnology* **16**, 313–317 (2020).
 - [5] L. Bhan, C. L. Covington, and K. Varga, Laser-driven petahertz electron ratchet nanobubbles, *Nano Letters* **22**, 4240–4245 (2022).
 - [6] L. Conrad, B. Paulus, and J. C. Tremblay, Non-equilibrium charge transport through molecular junctions as stochastic many-electron dynamics, *The Journal of Chemical Physics* **164**, 054102 (2026).
 - [7] Q. Zeng, B. Chen, S. Zhang, D. Kang, H. Wang, X. Yu, and J. Dai, Full-scale ab initio simulations of laser-driven atomistic dynamics, *npj Computational Materials* **9**, 213 (2023).

- [8] K. Yang, Y. Zhang, K.-Y. Li, K.-Y. Lin, S. Gopalakrishnan, M. Rigol, and B. L. Lev, Phantom energy in the nonlinear response of a quantum many-body scar state, [Science](#) **385**, 1063–1067 (2024).
- [9] A. Sazhin, V. N. Gladilin, A. Erglis, G. Hellmann, F. Vewinger, M. Weitz, M. Wouters, and J. Schmitt, Observation of nonlinear response and onsager regression in a photon bose-einstein condensate, [Nature Communications](#) **15**, 4730 (2024).
- [10] K. E. Dorfman, F. Schlawin, and S. Mukamel, Nonlinear optical signals and spectroscopy with quantum light, [Review of Modern Physics](#) **88**, 045008 (2016).
- [11] K. Sugisaki, S. Yamamoto, S. Nakazawa, K. Toyota, K. Sato, D. Shiomi, and T. Takui, Quantum chemistry on quantum computers: A polynomial-time quantum algorithm for constructing the wave functions of open-shell molecules, [The Journal of Physical Chemistry A](#) **120**, 6459–6466 (2016).
- [12] B. Shi and Y.-M. Lu, Deciphering the nonlocal entanglement entropy of fracton topological orders, [Phys. Rev. B](#) **97**, 144106 (2018).
- [13] W. Domcke and D. R. Yarkony, Role of conical intersections in molecular spectroscopy and photoinduced chemical dynamics, [Annual Review of Physical Chemistry](#) **63**, 325–352 (2012).
- [14] W. Domcke, D. R. Yarkony, and H. Köppel, *Conical Intersections: Electronic Structure, Dynamics and Spectroscopy* (WORLD SCIENTIFIC, 2004).
- [15] S. E. Smart and D. A. Mazziotti, Quantum solver of contracted eigenvalue equations for scalable molecular simulations on quantum computing devices, [Physical Review Letters](#) **126**, 070504 (2021).
- [16] S. E. Smart and D. A. Mazziotti, Many-fermion simulation from the contracted quantum eigensolver without fermionic encoding of the wave function, [Physical Review A](#) **105**, 062424 (2022).
- [17] Y. Wang and D. A. Mazziotti, Electronic excited states from a variance-based contracted quantum eigensolver, [Physical Review A](#) **108**, 022814 (2023).
- [18] S. E. Smart, D. M. Welakuh, and P. Narang, Many-body excited states with a contracted quantum eigensolver, [Journal of Chemical Theory and Computation](#) **20**, 3580–3589 (2024).
- [19] C. L. Benavides-Riveros, Y. Wang, S. Warren, and D. A. Mazziotti, Quantum simulation of excited states from parallel contracted quantum eigensolvers, [New Journal of Physics](#) **26**, 033020 (2024).

- [20] C. L. Benavides-Riveros, L. Chen, C. Schilling, S. Mantilla, and S. Pittalis, Excitations of quantum many-body systems via purified ensembles: A unitary-coupled-cluster-based approach, [Physical Review Letters](#) **129**, 066401 (2022).
- [21] A. C. Mater and M. L. Coote, Deep learning in chemistry, [Journal of Chemical Information and Modeling](#) **59**, 2545–2559 (2019).
- [22] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, Physics-informed machine learning, [Nature Reviews Physics](#) **3**, 422–440 (2021).
- [23] L. Zhao and H. Zong, Ai-driven decoding of material dynamics: From machine learning potentials and interpretability to generative prediction, [Advanced Materials](#) **1**, e14626.
- [24] P. O. Dral, Ai in computational chemistry through the lens of a decade-long journey, [Chemical Communications](#) **60**, 3240–3258 (2024).
- [25] J. Zhang, C. L. Benavides-Riveros, and L. Chen, Artificial-intelligence-based surrogate solution of dissipative quantum dynamics: Physics-informed reconstruction of the universal propagator, [The Journal of Physical Chemistry Letters](#) **15**, 3603–3610 (2024).
- [26] Y. Alexeev, M. H. Farag, T. L. Patti, M. E. Wolf, N. Ares, A. Aspuru-Guzik, S. C. Benjamin, Z. Cai, S. Cao, C. Chamberland, Z. Chandani, F. Fedele, I. Hamamura, N. Harrigan, J.-S. Kim, E. Kyoseva, J. G. Lietz, T. Lubowe, A. McCaskey, R. G. Melko, K. Nakaji, A. Peruzzo, P. Rao, B. Schmitt, S. Stanwyck, N. M. Tubman, H. Wang, and T. Costa, Artificial intelligence for quantum computing, [Nature Communications](#) **16**, 10829 (2025).
- [27] A. G. B. Richard S. Sutton, *Reinforcement Learning: An Introduction* (MIT Press, 1998).
- [28] W. B. Powell, *Reinforcement Learning and Stochastic Optimization: A Unified Framework for Sequential Decisions* (Wiley, 2022).
- [29] X. Wang, S. Wang, X. Liang, D. Zhao, J. Huang, X. Xu, B. Dai, and Q. Miao, Deep reinforcement learning: A survey, [IEEE Transactions on Neural Networks and Learning Systems](#) **35**, 5064 (2024).
- [30] M. Ghasemi, A. H. Moosavi, and D. Ebrahimi, [A comprehensive survey of reinforcement learning: From algorithms to practical challenges](#) (2025), [arXiv:2411.18892 \[cs.AI\]](#).
- [31] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, Mastering the game of go with deep neural networks and tree search, [Nature](#) **529**, 484–489

- (2016).
- [32] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis, A general reinforcement learning algorithm that masters chess, shogi, and go through self-play, [Science](#) **362**, 1140 (2018).
 - [33] O. Vinyals, I. Babuschkin, W. M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, D. H. Choi, R. Powell, T. Ewalds, P. Georgiev, J. Oh, D. Horgan, M. Kroiss, I. Danihelka, A. Huang, L. Sifre, T. Cai, J. P. Agapiou, M. Jaderberg, A. S. Vezhnevets, R. Leblond, T. Pohlen, V. Dalibard, D. Budden, Y. Sulsky, J. Molloy, T. L. Paine, C. Gulcehre, Z. Wang, T. Pfaff, Y. Wu, R. Ring, D. Yogatama, D. Wünsch, K. McKinney, O. Smith, T. Schaul, T. Lillicrap, K. Kavukcuoglu, D. Hassabis, C. Apps, and D. Silver, Grandmaster level in starcraft ii using multi-agent reinforcement learning, [Nature](#) **575**, 350–354 (2019).
 - [34] J. Kober, J. A. Bagnell, and J. Peters, Reinforcement learning in robotics: A survey, [The International Journal of Robotics Research](#) **32**, 1238–1274 (2013).
 - [35] C. Tang, B. Abbatematteo, J. Hu, R. Chandra, R. Martín-Martín, and P. Stone, Deep reinforcement learning for robotics: A survey of real-world successes, [Annual Review of Control, Robotics, and Autonomous Systems](#) **8**, 153–188 (2025).
 - [36] Y. Li, X. Ma, J. Xu, Y. Cui, Z. Cui, Z. Han, L. Huang, T. Kong, Y. Liu, H. Niu, W. Peng, J. Qiao, Z. Ren, H. Shi, Z. Su, J. Tian, Y. Xiao, S. Zhang, L. Zheng, H. Li, and Y. Wu, [Gr-rl: Going dexterous and precise for long-horizon robotic manipulation](#) (2025), [arXiv:2512.01801 \[cs.RO\]](#).
 - [37] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. A. Sallab, S. Yogamani, and P. Pérez, Deep reinforcement learning for autonomous driving: A survey, [IEEE Transactions on Intelligent Transportation Systems](#) **23**, 4909 (2022).
 - [38] J. Dinneweth, A. Boubezoul, R. Mandiau, and S. Espié, Multi-agent reinforcement learning for autonomous vehicles: a survey, [Autonomous Intelligent Systems](#) **2**, 27 (2022).
 - [39] M. Tobisawa, K. Matsuda, T. Suzuki, T. Harada, J. Hoshino, Y. Itoh, K. Kumagae, J. Matsuoka, and K. Hattori, Reinforcement learning-based autonomous driving control for efficient road utilization in lane-less environments, [Artificial Life and Robotics](#) **30**, 276–288 (2025).
 - [40] D. M. Ziegler, N. Stiennon, J. Wu, T. B. Brown, A. Radford, D. Amodei, P. Christiano, and G. Irving, [Fine-tuning language models from human preferences](#) (2020), [arXiv:1909.08593](#)

[cs.CL].

- [41] Q. Liu, Z. Song, Y. Liang, Z. Xie, S. Zhang, J. Zhang, and Y. Li, Corlhf: Reinforcement learning from human feedback with cooperative policy-reward optimization for llms, [Expert Systems with Applications](#) **301**, 130113 (2026).
- [42] N. Lambert, [Reinforcement learning from human feedback](#) (2026), [arXiv:2504.12501 \[cs.LG\]](#).
- [43] J. Yao, L. Lin, and M. Bukov, Reinforcement learning for many-body ground-state preparation inspired by counterdiabatic driving, [Physical Review X](#) **11**, 031070 (2021).
- [44] P. Peng, X. Huang, C. Yin, L. Joseph, C. Ramanathan, and P. Cappellaro, Deep reinforcement learning for quantum hamiltonian engineering, [Physical Review Applied](#) **18**, 024033 (2022).
- [45] Y. Wang and D. A. Mazziotti, Quantum many-body simulations from a reinforcement-learned exponential ansatz, [Physical Review A](#) **112**, 022403 (2025).
- [46] M. Bukov, A. G. R. Day, D. Sels, P. Weinberg, A. Polkovnikov, and P. Mehta, Reinforcement learning in different phases of quantum control, [Physical Review X](#) **8**, 031086 (2018).
- [47] T. Fösel, P. Tighineanu, T. Weiss, and F. Marquardt, Reinforcement learning with neural networks for quantum feedback, [Physical Review X](#) **8**, 031084 (2018).
- [48] M. Y. Niu, S. Boixo, V. N. Smelyanskiy, and H. Neven, Universal quantum control through deep reinforcement learning, [npj Quantum Information](#) **5**, 33 (2019).
- [49] K. Reuer, J. Landgraf, T. Fösel, J. O’Sullivan, L. Beltrán, A. Akin, G. J. Norris, A. Remm, M. Kerschbaum, J.-C. Besse, F. Marquardt, A. Wallraff, and C. Eichler, Realizing a deep reinforcement learning agent for real-time quantum feedback, [Nature Communications](#) **14**, 7138 (2023).
- [50] M. A. Nielsen and I. L. Chuang, [Quantum Computation and Quantum Information: 10th Anniversary Edition](#) (Cambridge University Press, 2012).
- [51] S. B. Bravyi and A. Y. Kitaev, Fermionic quantum computation, [Annals of Physics](#) **298**, 210 (2002).
- [52] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, [Playing atari with deep reinforcement learning](#) (2013), [arXiv:1312.5602 \[cs.LG\]](#).
- [53] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, Human-level control through deep reinforcement learning, [Nature](#) **518**, 529–533 (2015).

- [54] Z. Wang, T. Schaul, M. Hessel, H. Van Hasselt, M. Lanctot, and N. De Freitas, Dueling network architectures for deep reinforcement learning, in *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48*, ICML'16 (JMLR.org, 2016) p. 1995–2003.
- [55] H. Nakatsuji, Equation for the direct determination of the density matrix: Time-dependent density equation and perturbation theory, [Theoretical Chemistry Accounts: Theory, Computation, and Modeling \(Theoretica Chimica Acta\)](#) **102**, 97–104 (1999).
- [56] M. Rose and D. A. Mazziotti, [Many-body time evolution from a correlation-efficient quantum algorithms](#) (2025), [arXiv:2511.13871 \[cs.AI\]](#).
- [57] C. Cianci, L. F. Santos, and V. S. Batista, Subspace-search quantum imaginary time evolution for excited state computations, [Journal of Chemical Theory and Computation](#) **20**, 8940 (2024).
- [58] Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksen, Y. Gao, S. Guo, J. Hermann, M. R. Hermes, K. Koh, P. Koval, S. Lehtola, Z. Li, J. Liu, N. Mardirossian, J. D. McClain, M. Motta, B. Mussard, H. Q. Pham, A. Pulkin, W. Purwanto, P. J. Robinson, E. Ronca, E. R. Sayfutyarova, M. Scheurer, H. F. Schurkus, J. E. T. Smith, C. Sun, S.-N. Sun, S. Upadhyay, L. K. Wagner, X. Wang, A. White, J. D. Whitfield, M. J. Williamson, S. Wouters, J. Yang, J. M. Yu, T. Zhu, T. C. Berkelbach, S. Sharma, A. Y. Sokolov, and G. K.-L. Chan, Recent developments in the pyscf program package, [The Journal of Chemical Physics](#) **153**, 024109 (2020).
- [59] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimeshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, Pytorch: an imperative style, high-performance deep learning library, in *Proceedings of the 33rd International Conference on Neural Information Processing Systems* (Curran Associates Inc., Red Hook, NY, USA, 2019).
- [60] C. Feniou, O. Adjoua, B. Claudon, J. Zylberman, E. Giner, and J.-P. Piquemal, Sparse quantum state preparation for strongly correlated systems, [The Journal of Physical Chemistry Letters](#) **15**, 3197–3205 (2024).
- [61] J. Iaconis, S. Johri, and E. Y. Zhu, Quantum state preparation of normal distributions using matrix product states, [npj Quantum Information](#) **10**, 15 (2024).

- [62] L. H. Delgado-Granados, T. J. Krogmeier, L. M. Sager-Smith, I. Avdic, Z. Hu, M. Sajjan, M. Abbasi, S. E. Smart, P. Narang, S. Kais, A. W. Schlimgen, K. Head-Marsden, and D. A. Mazziotti, Quantum algorithms and applications for open quantum systems, [Chemical Reviews](#) **125**, 1823 (2025).
- [63] S. Warren, Y. Wang, C. L. Benavides-Riveros, and D. A. Mazziotti, Quantum algorithm for polaritonic chemistry based on an exact ansatz, [Quantum Science and Technology](#) **10**, 02LT02 (2025).