Editors' Suggestion

Quantum dynamics via score matching on Bohmian trajectories

Lei Wang *

Beijing National Laboratory for Condensed Matter Physics and Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China

(Received 6 May 2026; revised 3 August 2026; accepted 17 September 2026; published 9 November 2026)

We solve the time-dependent Schrödinger equation by learning the score function, the gradient of the log-probability density, on Bohmian trajectories. In Bohm's formulation of quantum mechanics, particles follow deterministic paths under the classical potential supplemented by a quantum potential depending on the score function of the evolving density. These noncrossing Bohmian trajectories form a continuous normalizing flow governed by the score. We parametrize the score with a neural network and minimize a self-consistent Fisher divergence between the network and the score of the resulting density. We prove that the zero-loss minimizer of this self-consistent objective recovers Schrödinger dynamics for nodeless wave functions, a condition met by localized vibrational wave packets of atoms before anharmonic dephasing sets in. We demonstrate the approach on wave-packet splitting in a double-well potential and anharmonic vibrations of a Morse chain. By recasting real-time quantum dynamics as a self-consistent score-driven normalizing flow, this framework opens the time-dependent Schrödinger equation to the rapidly advancing toolkit of modern generative modeling.

DOI: 10.1103/lmhf-ws7v

I. INTRODUCTION

The time-dependent Schrödinger equation (TDSE) for quantum particles of mass m with configuration $\boldsymbol{x} \in \mathbb{R}^d$ and potential $V(\boldsymbol{x})$,

$$i\hbar \frac{\partial \Psi(\boldsymbol{x}, t)}{\partial t} = \left[-\frac{\hbar^2}{2m} \nabla^2 + V(\boldsymbol{x}) \right] \Psi(\boldsymbol{x}, t), \tag{1}$$

governs quantum dynamics across electronic, atomic, and molecular physics. Grid-based methods scale exponentially with the number of degrees of freedom d, motivating alternatives that expand the wave function in a time-adaptive basis [1], parametrize it with neural networks [2], or represent it through particle trajectories [3–5]—the approach we build on here.

The trajectory framework originates from the Madelung decomposition [6], which writes the wave function as $\Psi = \sqrt{\rho}\, e^{iS/\hbar}$ and recasts Eq. (1) as two coupled Eulerian field equations: a

*Contact author: wanglei@iphy.ac.cn

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

continuity equation for the probability density ρ and a quantum Hamilton-Jacobi equation for the phase S,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \left(\rho \frac{\nabla S}{m} \right) = 0, \tag{2}$$

$$\frac{\partial S}{\partial t} + \frac{|\nabla S|^2}{2m} + V + Q = 0, \tag{3}$$

where the quantum potential $Q = -(\hbar^2/2m)\nabla^2\sqrt{\rho}/\sqrt{\rho}$ encodes all nonclassical effects. De Broglie [7] introduced the pilot-wave picture in which a particle is guided by the wave with velocity $\boldsymbol{v} = \nabla S/m$, and Bohm [8,9] later independently developed an alternative *interpretation* of quantum mechanics in terms of hidden variables: The quantum potential Q guides an ensemble of particles along well-defined trajectories that reproduce all predictions of the Schrödinger equation.

In this article, we turn Bohmian mechanics from an interpretive framework into a practical computational tool for quantum dynamics by revealing its connection to modern machinery in generative modeling, including continuous normalizing flows [10,11] and score matching [12,13].

To begin, we define the score function $\boldsymbol{s} \equiv \nabla \ln \rho$, in terms of which the quantum potential takes the form

$$Q = -\frac{\hbar^2}{4m}\left[\nabla \cdot \boldsymbol{s} + \frac{1}{2}|\boldsymbol{s}|^2 \right]. \tag{4}$$

Writing $\ln \Psi = \frac{1}{2}\ln \rho + iS/\hbar$, one sees that $\boldsymbol{s}$ and $\boldsymbol{v}$ are (up to constants) the gradients of the real and imaginary parts of $\ln \Psi$. Both are therefore gradient fields by construction.

To follow individual particles in the Lagrangian frame, we apply the material derivative $d/dt = \partial/\partial t + \boldsymbol{v} \cdot \nabla$ to the Madelung system, obtaining the Bohmian trajectory equations [8,9]:

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{v}, \quad m\frac{d\boldsymbol{v}}{dt} = -\nabla(V + Q). \tag{5}$$

This is formally identical to classical mechanics with effective potential $V + Q$, except that $\boldsymbol{v} = \nabla S/m$ is a single-valued field on $(\boldsymbol{x}, t)$ rather than an independent degree of freedom. By the uniqueness theorem, trajectories cannot cross or pass through nodes (see Sec. 3.3.8 of Ref. [14]). The flow map $\boldsymbol{x}(0) \mapsto \boldsymbol{x}(t)$ is therefore a diffeomorphism for nodeless wave functions, a continuous normalizing flow [10], with $\det F \neq 0$, where $F(t) = \partial \boldsymbol{x}(t)/\partial \boldsymbol{x}(0)$ is the deformation gradient (flow Jacobian). Since $\boldsymbol{v} = \nabla S/m$ is a gradient field, the flow is more specifically a Monge-Ampère flow [11]. Each particle, initialized at $\boldsymbol{x}(0) \sim \rho_0 \equiv \rho(\cdot, 0)$ with velocity $\boldsymbol{v}(0) = \nabla S(\boldsymbol{x}, 0)/m$, follows a deterministic trajectory, and the density evolves exactly as $\rho(\boldsymbol{x}(t), t) = \rho_0(\boldsymbol{x}(0))/|\det F(t)|$. Figure 1 illustrates this flow for a wave-packet splitting in a one-dimensional double well: The unimodal initial density (the base distribution of the normalizing flow) is deformed into multimodal densities, and noncrossing Bohmian trajectories launched from random samples of the initial density trace the full density evolution in between.

In Eq. (5), the phase S no longer appears as a dynamical field: $\boldsymbol{v}$ is propagated directly and Q depends only on the score $\boldsymbol{s}$ via Eq. (4), so the complex phase need not be represented along the trajectory [15,16]. The phase enters only through the initial condition $\boldsymbol{v}(0) = \nabla S(\boldsymbol{x}, 0)/m$ (trivially zero when the initial wave function is real), and S can be reconstructed *post hoc* by integrating Eq. (3) along trajectories, if desired. The practical difficulty, however, is that integrating Eq. (5) requires the quantum force $-\nabla Q$, which depends on the score $\boldsymbol{s}$ of the evolving density. This score is itself unknown and must be determined self-consistently. References [3,4] pioneered quantum trajectory methods by representing the density on a moving mesh and propagating the hydrodynamic equations directly. References [15,16] eliminated the mesh by fitting $\boldsymbol{s}$ from the particle ensemble via a polynomial basis and a functional equivalent to Hyvärinen score matching [17]. Reference [18] used a Bayesian Gaussian mixture for the same purpose. These methods fit the score pointwise at each time step using a limited basis (polynomials or Gaussian mixtures), without a global variational principle connecting successive times.

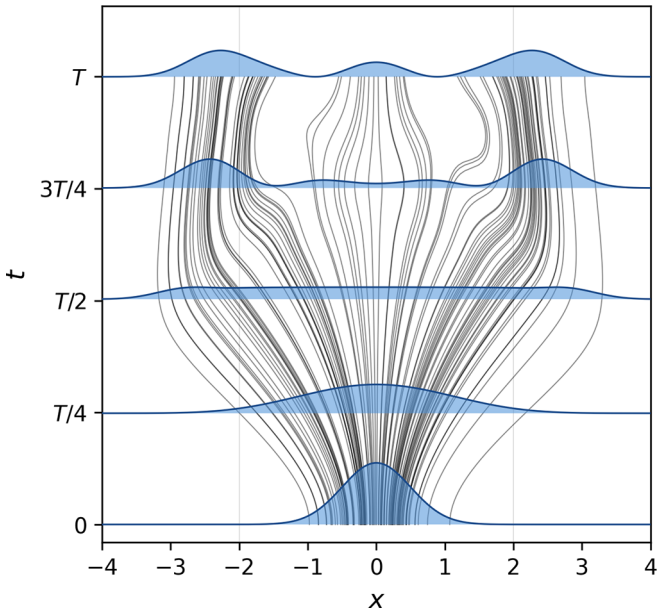


FIG. 1. Quantum dynamics as a continuous normalizing flow. The black curves are deterministic, non-crossing Bohmian trajectories, launched from particle positions sampled from the initial density and tracking the time-dependent densities shown by the blue waterfall. The blue waterfall plots the exact density $|\Psi(x,t)|^2$ at evenly spaced snapshots, showing an initial Gaussian wave packet splitting into multimodal densities in a symmetric double-well potential.

In this article, we instead parametrize the score with a neural network $\boldsymbol{s}_{\boldsymbol{\theta}}(\boldsymbol{x},t)$ and train it end-to-end by minimizing a single Fisher-divergence objective, simultaneously lifting the basis-set restriction and enforcing spacetime self-consistency. The resulting quantum potential drives a continuous normalizing flow from the initial density, and any zero-loss minimizer of the objective recovers the Schrödinger solution.

II. THEORY

The central idea is to train the score network by minimizing the mismatch between $\boldsymbol{s}_{\boldsymbol{\theta}}$ and the score of its own implied density. The training objective is the Fisher divergence:

$$L[\boldsymbol{s}_{\boldsymbol{\theta}}] = \int_0^T dt \int_{\mathbb{R}^d} d\boldsymbol{x}\, \rho_{\boldsymbol{\theta}}\, |\boldsymbol{s}_{\boldsymbol{\theta}} - \boldsymbol{\nabla} \ln \rho_{\boldsymbol{\theta}}|^2, \tag{6}$$

where T is the time horizon over which the wave function is propagated. Both integrals are estimated on the trajectory itself: The spatial integral becomes the per-time expectation $\mathbb{E}_{\boldsymbol{x}\sim\rho_{\boldsymbol{\theta}}(\cdot,t)}[|\boldsymbol{s}_{\boldsymbol{\theta}} - \boldsymbol{\nabla} \ln \rho_{\boldsymbol{\theta}}|^2]$ on the particle ensemble, and the time integral becomes a sum over trajectory checkpoints. Crucially, $\rho_{\boldsymbol{\theta}}$ itself depends on $\boldsymbol{s}_{\boldsymbol{\theta}}$: The score drives the trajectories that determine the density. Since $L \geqslant 0$ with equality iff $\boldsymbol{s}_{\boldsymbol{\theta}} = \boldsymbol{\nabla} \ln \rho_{\boldsymbol{\theta}}$, minimizing L drives the system toward self-consistency. This self-consistency is precisely what remains of the coupling between the Madelung equations once the phase has been eliminated: The quantum potential $Q[\boldsymbol{s}_{\boldsymbol{\theta}}]$ must agree with the score of the density its own flow generates, and L is a computable measure of the residual. Because L is an expectation over $\rho_{\boldsymbol{\theta}}$, it concentrates on regions where the physical density has support: The score need not be accurate where no particles are sampled.

Theorem. Suppose the score $\boldsymbol{s}_{\boldsymbol{\theta}}$ achieves (a) $L[\boldsymbol{s}_{\boldsymbol{\theta}}] = 0$ and (b) induces a Bohmian flow whose density remains strictly positive on $[0,T]$. Then, $\boldsymbol{s}_{\boldsymbol{\theta}}(\boldsymbol{x},t) = \boldsymbol{\nabla} \ln |\Psi(\boldsymbol{x},t)|^2$, where Ψ is the unique

TABLE I. Derivative chain in the Fisher loss gradient. Rows 1–5 are successive spatial derivatives of $\varphi_{\boldsymbol{\theta}}(\boldsymbol{x}, t)$ of increasing order; row 6 is the parameter gradient obtained by BPTT through the trajectory.

Order	Quantity	Role
First	$\boldsymbol{s} = \nabla\varphi$	Score function
Second	$\nabla \cdot \boldsymbol{s}$	Quantum potential Q, Eq. (4)
Third	$-\nabla Q$	Quantum force, Eq. (5)
Fourth	$\mathrm{Hess}(V + Q)$	Drives (F, G), Eq. (8)
Fifth	$\nabla_{\boldsymbol{x}(0)} \ln \lvert\det F\rvert$	Score transport, Eq. (7)
Sixth	$dL/d\boldsymbol{\theta}$	Parameter gradient (BPTT)

solution of the TDSE (1) evolving from the prescribed initial wave function; in particular, the wave function (and hence its density and phase) is fully recovered. The proof is detailed in Appendix A. The theorem characterizes the zero-loss minimizers; it supplies no existence, convergence, or stability result, and no error bound at finite loss. In particular, the Fisher divergence controls $\boldsymbol{s}$ in $L^2(\rho_{\boldsymbol{\theta}})$, whereas the quantum force involves two further derivatives of the score, so a small loss does not by itself bound the error in ∇Q without additional regularity assumptions.

Evaluating the loss function Eq. (6) requires $\nabla \ln \rho_{\boldsymbol{\theta}}$ along each trajectory. Differentiating both sides of $\ln \rho_{\boldsymbol{\theta}} = \ln \rho_0 - \ln \lvert\det F\rvert$ with respect to $\boldsymbol{x}(0)$ and transforming to the current frame yields

$$\nabla_{\boldsymbol{x}(t)} \ln \rho_{\boldsymbol{\theta}} = F^{-\mathsf{T}} \cdot [\nabla_{\boldsymbol{x}(0)} \ln \rho_0 - \nabla_{\boldsymbol{x}(0)} \ln \lvert\det F\rvert]. \tag{7}$$

Note that the $d \times d$ deformation gradient F is needed to obtain the target score $\nabla \ln \rho_{\boldsymbol{\theta}}$. Differentiating the equations of motion (5) with respect to the initial position $\boldsymbol{x}(0)$ yields the evolution of the deformation gradient F and its velocity counterpart $G = \partial \boldsymbol{v}(t)/\partial \boldsymbol{x}(0)$:

$$\frac{dF}{dt} = G, \quad m\frac{dG}{dt} = -\mathrm{Hess}(V + Q) \cdot F, \tag{8}$$

with initial conditions $F(0) = I$ and $G(0) = \partial \boldsymbol{v}(0)/\partial \boldsymbol{x}(0)$, which reduces to $G(0) = 0$ whenever the initial wave function is real ($\boldsymbol{v}(0) = 0$), as in both applications below. Here, $[\mathrm{Hess}(V + Q)]_{ij} = \partial^2(V + Q)/\partial x_i \partial x_j$ is the Hessian matrix of second derivatives. F and G track how an infinitesimal neighborhood of initial conditions deforms under the flow, and are co-integrated alongside the particle trajectories using a symplectic leapfrog scheme.

III. IMPLEMENTATION

We parametrize the score function as $\boldsymbol{s}_{\boldsymbol{\theta}}(\boldsymbol{x}, t) = \nabla_{\boldsymbol{x}} \varphi_{\boldsymbol{\theta}}(\boldsymbol{x}, t)$, the gradient of the scalar function $\varphi_{\boldsymbol{\theta}}$, ensuring that $\boldsymbol{s}_{\boldsymbol{\theta}}$ is a gradient field by construction. The loss (6) depends on $\boldsymbol{\theta}$ through the score $\boldsymbol{s}_{\boldsymbol{\theta}}$, the flow Jacobian F, and the sampling distribution $\rho_{\boldsymbol{\theta}}$. Since initial positions $\boldsymbol{x}_i(0) \sim \rho_0$ are fixed, gradients flow end-to-end through the trajectory ordinary differential equation (ODE) via the chain rule. The full gradient $dL/d\boldsymbol{\theta}$ is computed by backpropagation through time (BPTT) along the discretized leapfrog trajectory, with gradient checkpointing [19] to manage memory. The algorithm chains five spatial-derivative orders of $\varphi_{\boldsymbol{\theta}}$, plus the parameter gradient via BPTT. The full derivative ladder is given in Appendix B, Table I. Gradient stability is aided by the symplecticity of the leapfrog integrator, which constrains the growth of Jacobian products. Figure 2 summarizes the resulting computational loop.

IV. APPLICATION

We apply the method to a chain of d coupled Morse oscillators $V = \sum_{i=0}^{d-1} D(1 - e^{-\beta x_i})^2 + \lambda \sum_{i=0}^{d-2} x_i x_{i+1}$, with $D = 12.5$, $\beta = 0.2$, $\lambda = 0.3$, and the first mode displaced by $x_0 = 1$. The score is decomposed as $\boldsymbol{s}_{\boldsymbol{\theta}} = \boldsymbol{s}_{\mathrm{base}} + \nabla_{\boldsymbol{x}} \varphi_{\boldsymbol{\theta}}$, where $\boldsymbol{s}_{\mathrm{base}}$ is the analytical score from the harmonic

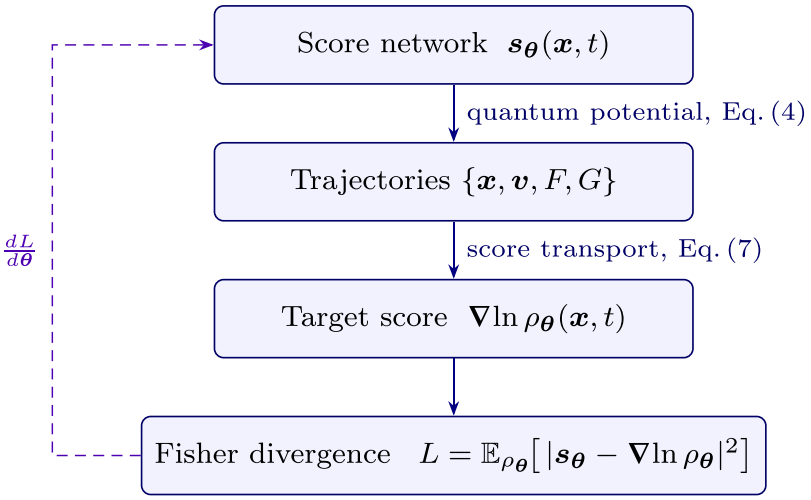


FIG. 2. Training loop for the self-consistent score matching on Bohmian trajectories. The score network s_θ determines the quantum potential Q, which together with the classical potential V drives the Bohmian trajectories via Eq. (5). Along each trajectory, we co-integrate the deformation gradient F and its velocity counterpart G via Eq. (8). The Fisher divergence [Eq. (6)] measures the mismatch between s_θ and the score target [Eq. (7)], and gradients $dL/d\theta$ propagate end to end through the trajectory via BPTT.

approximation to the potential, providing a warm start. The correction φ_θ is a multilayer perceptron with feature-wise linear modulation (FiLM) [20] time conditioning, which modulates each hidden layer by a learned scale and shift that depend on t. Training uses $M = 5000$ particles, $T = \pi$, $dt = 0.01$ (314 leapfrog steps). Exact Bohmian dynamics on nodeless wave functions admits no caustics: The quantum potential Q provides the antifocusing force that keeps $\det F$ bounded away from zero, and the flow map is a diffeomorphism (see Sec. 3.3.8 of Ref. [14]). Early in training, however, the learned score yields an insufficient Q and classical trajectory focusing appears transiently. We mask near-caustic time steps ($|\det F| < 0.1$) in the loss to avoid divergent score targets. As training progresses, the learned quantum force $-\nabla Q$ prevents caustic formation and the masked fraction drops to zero. This regularization therefore does not affect the final accuracy.

Figure 3 shows the training convergence over 45 000 epochs. The occasional spikes visible in panels (a) and (b) are caustic events: $|\det F|$ transiently becomes small, so the score target $F^{-\mathsf{T}}[\nabla_{x(0)} \ln \rho_0 - \nabla_{x(0)} \ln |\det F|]$ in Eq. (7) inflates and the network takes an outsized gradient step. The system self-heals within tens of epochs: The learned Q strengthens, $|\det F|$ recovers, and both the loss and energy error return to baseline. The Fisher loss [panel (a)] decreases steadily apart from these events; being a self-consistency residual rather than an error, its magnitude serves as a self-consistency diagnostic, while accuracy is quantified in panel (b) and against the reference solution below. Panel (b) reports the mean-energy error $|\langle E\rangle - E_{\text{exact}}|$, where the per-particle energy $\frac{1}{2}m|v_i|^2 + V(x_i) + Q(x_i, t)$ is averaged over all ensemble particles and over the trajectory interval $[0, T]$. For the exact Bohmian flow, $\langle E\rangle$ at each time equals the conserved expectation value, so this scalar measures the time-and-ensemble-averaged bias relative to E_{exact}; a temporal drift whose excursions cancel would not show up in it, and we monitor that separately through the temporal standard deviation of $\langle E\rangle(t)$, which is $\approx 5\times10^{-3}$ at convergence. The energy error decreases from $\sim$0.2 to a median of $\sim 4\times10^{-3}$ at convergence, a relative error of 0.17% against $E_{\text{exact}} = 2.3236$ obtained from a four-dimensional (4D) split-operator fast-Fourier-transform (FFT) reference on a 64^4 grid.

To verify that the trained score reproduces the correct quantum dynamics, we propagate an independent ensemble of $M = 20\,000$ particles with the learned score. Figure 4 compares the resulting

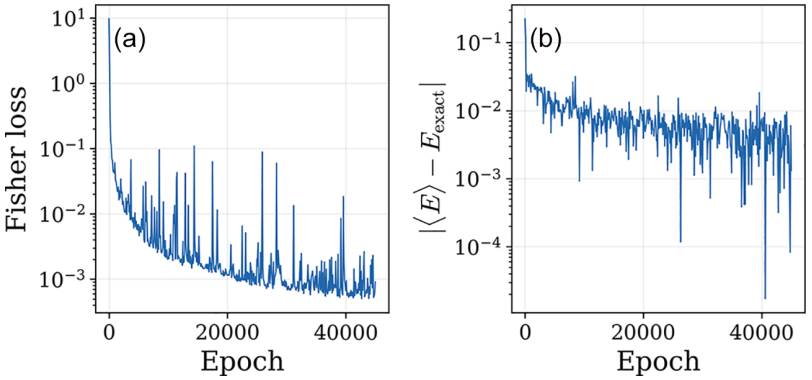


FIG. 3. Training convergence on the $d = 4$ Morse chain over 45 000 epochs (raw per-epoch values, log scale). (a) Fisher loss vs epoch. (b) Mean-energy error $|\langle E\rangle - E_{\text{exact}}|$ vs epoch. $\langle E\rangle$ is averaged over the $M = 5000$ particles and the trajectory time $[0, T]$, and $E_{\text{exact}} = 2.3236$ is obtained from a 4D split-operator FFT reference.

per-mode mean $\langle x_i\rangle(t)$ and width $\sigma_i(t) = \sqrt{\langle x_i^2\rangle - \langle x_i\rangle^2}$ against the exact FFT reference. Panel (a) shows dipolelike oscillations of the mode positions: The initially displaced $i = 0$ mode oscillates back through and past equilibrium under its restoring force, while the bilinear coupling $\lambda\, x_i x_{i+1}$ transmits this excitation along the chain, driving progressively delayed and smaller-amplitude motion of modes $i = 1, 2, 3$. Panel (b) shows mode-specific breathing of the widths: The Morse anharmonicity makes the local restoring force amplitude dependent, so each marginal is squeezed and stretched as it samples regions of varying curvature, with $i = 0$ breathing most strongly because of its larger excursion and the others inheriting weaker breathing through their coupled motion. The network reproduces both effects in phase and amplitude across all four modes.

To visualize what the network has learned, Fig. 5 shows the score field $\boldsymbol{s}_{\boldsymbol{\theta}}(\boldsymbol{x}, t)$ on the (x_0, x_1) slice through the 4D configuration space, with nonplotted dimensions fixed at equilibrium $(x_2 = x_3 = 0)$. The background color shows the exact slice density $\rho(x_0, x_1, 0, 0) = |\Psi(x_0, x_1, 0, 0)|^2$ on a logarithmic scale, computed from the FFT wave function. Streamlines depict the learned score field $\boldsymbol{s}_{\boldsymbol{\theta}}$, which the network must produce self-consistently to satisfy $\boldsymbol{s}_{\boldsymbol{\theta}} = \nabla \ln \rho_{\boldsymbol{\theta}}$.

At $t = 0$ [panel (a)], the score field is approximately radial, reflecting the Gaussian initial state. At $t = \pi$ [panel (b)], the density has deformed significantly under the anharmonic coupled potential:

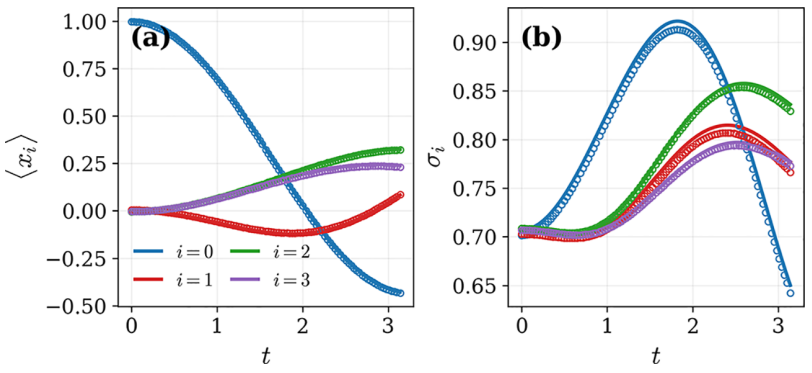


FIG. 4. Mean $\langle x_i\rangle(t)$ [panel (a)] and width $\sigma_i(t)$ [panel (b)] for each of the four modes of the $d = 4$ Morse chain. Solid lines: exact FFT reference. Open circles: $M = 20\,000$ particles propagated with the trained score network.

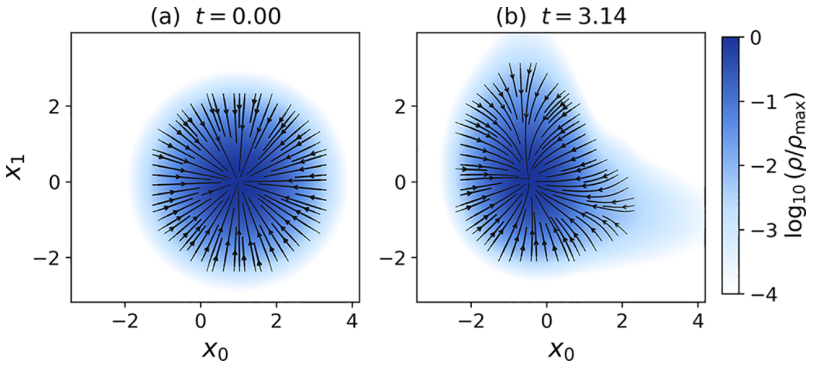


FIG. 5. Learned score field $\boldsymbol{s}_{\boldsymbol{\theta}}(\boldsymbol{x}, t)$ (streamlines) on the (x_0, x_1) slice of the $d = 4$ Morse chain ($x_2 = x_3 = 0$). Background: Exact slice density on a log scale. (a) $t = 0$: Gaussian initial state with approximately radial score. (b) $t = \pi$: anharmonic coupling deforms the density and the score field. The learned streamlines remain aligned with the exact density gradient across the wave packet.

The wave packet elongates along x_0 and loses its near-rotational symmetry. The score field tracks this deformation, curving to follow the density gradient of the non-Gaussian distribution. The network captures the full 4D score on these trajectories: Even though only a 2D projection is displayed, $\boldsymbol{s}_{\boldsymbol{\theta}}$ is a four-component vector field that must simultaneously be correct in all dimensions to maintain the self-consistent density evolution. As a direct density-level check, we evaluate the reverse Kullback-Leibler divergence $\mathrm{KL}(\rho_{\boldsymbol{\theta}} \| \rho_{\mathrm{exact}})(t) \approx M^{-1} \sum_i [\ln \rho_{\boldsymbol{\theta}}(\boldsymbol{x}_i(t), t) - \ln \rho_{\mathrm{exact}}(\boldsymbol{x}_i(t), t)]$ as a per-particle Monte Carlo average, using $\ln \rho_{\boldsymbol{\theta}}(\boldsymbol{x}_i(t), t) = \ln \rho_0(\boldsymbol{x}_i(0)) - \ln |\det F_i(t)|$ from the change of variables and $\rho_{\mathrm{exact}} = |\Psi|^2$ multilinearly interpolated from the 4D FFT reference. At $M = 20\,000$, the KL remains near the interpolation floor throughout the trajectory ($\sim 10^{-2}$ nats at $t = 0$ and $t = \pi$), confirming that the learned density matches the Schrödinger solution at the full density level, not only in low-order moments.

V. DISCUSSION

Bohm's theory was famously regarded as a conceptual framework to interpret quantum mechanics [8,9]. Self-consistent score matching on Bohmian trajectories turns it into a practical computational tool for the TDSE, leveraging machinery from modern generative modeling. The link between score matching and quantum mechanics in fact predates that machinery itself. Already in 2003, Garashchuk and Rassolov [15,16] fitted the quantum potential on trajectories by minimizing $\mathbb{E}_\rho[\frac{1}{2}|\boldsymbol{s}|^2 + \nabla \cdot \boldsymbol{s}] \propto \langle Q \rangle$ in a polynomial basis—the implicit score-matching loss of Hyvärinen [17] up to constants. What modern generative modeling contributes—which we leverage here—is a flexible neural parametrization of the score, automatic differentiation through long trajectories, and a globally enforced self-consistent objective over the entire time horizon, rather than an instantaneous least-squares fit at each time step.

The parallel between generative modeling and quantum mechanics extends from deterministic trajectories to stochastic ones. Nelson's stochastic mechanics [21,22] is a Schrödinger-equivalent formulation in which particles follow a drift-diffusion process rather than the deterministic Bohmian path; the drift combines the Bohmian velocity with an osmotic term proportional to the score, $\boldsymbol{v} + (\hbar/2m)\nabla \ln \rho$. In modern generative modeling, flow-based models transport densities along deterministic ODEs, while diffusion-based models [12,13] use stochastic differential equations (SDEs) with a score-proportional drift, the two linked by a probability-flow ODE that reproduces the SDE's marginals. The Bohmian ODE is precisely the probability-flow ODE of Nelson's SDE: Bohmian mechanics is to flow-based generative models what Nelson's stochastic mechanics is to diffusion-based ones. The same viewpoint underlies recent neural methods for the Fokker-Planck

equation [23–27], which replace SDE simulation by integration of a deterministic ODE driven by a self-consistently learned score or velocity field. We extend this construction to Schrödinger dynamics. The broader toolkit of modern generative modeling, including denoising score matching and flow matching [28–30], may extend naturally to quantum dynamics. Orlova *et al.* [31] have taken a first step on the stochastic side with deep stochastic mechanics: They parametrize *both* the current and osmotic velocities of Nelson's process as neural networks, sample trajectories self-consistently by Euler-Maruyama integration of the learned SDE, and train the networks by minimizing partial differential equation (PDE) residuals of the corresponding velocity equations on those samples. The sampling strategy is thus self-generated as in the present method, but the governing dynamics are stochastic rather than deterministic, two velocity fields are learned instead of a single score, and the loss is a PDE residual rather than the Fisher divergence.

Other neural approaches to the TDSE differ more sharply. Time-dependent variational Monte Carlo (tVMC) [2,32], among the most established neural-variational methods for quantum dynamics, parametrizes the complex wave function and evolves its parameters in time by projecting the TDSE onto the tangent space of the parameter manifold via the time-dependent variational principle [33,34]. This projection requires inverting an often ill-conditioned quantum geometric tensor at every integrator step, together with Monte Carlo sampling from the time-dependent wave function. The present method trades nodal flexibility for a real-valued, density-only representation in which the complex phase is never parametrized—only its gradient $\boldsymbol{v} = \nabla S/m$ is propagated, and every sampled quantity enters the ensemble average with a strictly positive weight—replaces the instantaneous tangent-space projection by a single global loss over the time horizon, replaces Metropolis sampling of particle coordinates at every step by a deterministic particle ensemble transported along the Bohmian trajectories, and needs no metric inversion at all.

Physics-informed neural networks (PINNs) [35] are another widely used neural framework for solving PDEs that has been applied to the TDSE. PINNs parametrize the complex space-time wave function $\Psi_{\boldsymbol{\theta}}(\boldsymbol{x}, t)$ directly and minimize the TDSE residual on a set of collocation points drawn from a user-prescribed distribution over configuration space and time. As in the present method, the optimization is a single global loss over the time horizon, but the collocation distribution is chosen by hand rather than by the dynamics: Equal weight is given to regions the physical state never visits, and accuracy degrades in the tails where $|\Psi|^2$ is small. A recent application of PINNs to the fermionic TDSE [36] samples collocation from $|\Psi_{\boldsymbol{\theta}}|^2$ at each fixed t to bring the training distribution onto the physical support. A related global-in-time scheme [37] expands the complex wave function in a fixed neural-quantum-state basis with time-dependent coefficients, and likewise samples from $|\Psi_{\boldsymbol{\theta}}|^2$ at each time-integration node. The present method drops the complex phase via the Madelung passage and replaces user-prescribed collocation or $|\Psi_{\boldsymbol{\theta}}|^2$ sampling by a self-generated particle ensemble transported along the Bohmian trajectories, automatically allocating the training effort where the density has support.

Closer in spirit to the present method is the parametrized Wasserstein Hamiltonian flow (WHF) of Ref. [38], which likewise eliminates the complex wave function and casts Schrödinger dynamics as a continuous normalizing flow on the probability density. Like tVMC, however, WHF projects the TDSE onto the parameter manifold at each instant and integrates the resulting ODE for the neural-network parameters step by step. The present method replaces this integrator-driven projection by a global optimization of the Fisher divergence over the time horizon, endowed with a built-in self-consistency diagnostic ($L = 0$ iff self-consistent).

Nodes also arise dynamically: Exact solutions of the linear TDSE create and annihilate vortices in pairs of opposite circulation [39,40], and in that regime the de Broglie-Bohm flow turns singular and its trajectories become chaotic [41,42]. Since a diffeomorphism transports the zero level set of its base [43], a flow cannot create or reconnect nodes; this is a boundary of the smooth diffeomorphic flow representation used here, and the long-time or split-and-recombine regimes in $d \geqslant 2$ where vortices proliferate lie outside the present scope.

Fermionic wave functions generically have nodes: Antisymmetry under particle exchange forces $\Psi = 0$ on a set that partitions configuration space into sign-disconnected cells [44]. These nodes

violate the $\rho_\theta > 0$ hypothesis of the present analysis on two counts. First, at a node, the score $\boldsymbol{s} = \nabla \ln \rho$ diverges and Bohmian trajectories cannot cross. A smooth score network $\boldsymbol{s}_\theta = \nabla \varphi_\theta$ cannot represent the resulting cell structure. Second, nodes resurrect the Wallstrom objection [45]: The Madelung system recovers the Schrödinger equation only under the quantization condition $\oint \boldsymbol{v} \cdot d\boldsymbol{\ell} = 2\pi n\hbar/m$, $n \in \mathbb{Z}$, on every closed loop in configuration space, a constraint automatic for a single-valued Ψ but invisible to a score-only formulation. A possible route that resolves both issues is to employ the fermion flow architecture [46–48], which parametrizes many-electron wave functions as a permutation-equivariant normalizing flow from a Slater determinant. A diffeomorphism maps nodal cells to nodal cells, so within each cell $\rho > 0$ and the dynamics of Eq. (5) without explicit phase parametrization applies. The sign across cells is read off the Slater determinant rather than learned by the score network. The Slater determinant is also single-valued with integer phase winding fixed by its orbitals, and a diffeomorphism preserves that winding on every closed loop, so the Wallstrom condition is automatically satisfied. This addresses the sign and the winding, but not topology change: Mapping nodal cells to nodal cells is precisely what prevents a diffeomorphism from altering the nodal topology on its own, which would instead have to be carried by the determinant, for instance, through time-dependent orbitals. Combining such fermionic flow architectures with self-consistent score matching is therefore a promising route toward real-time quantum dynamics of many-electron systems rather than a solved construction.

ACKNOWLEDGMENTS

The author acknowledges discussions with Tao Wang, Yuan Wan, Kai Zhou, Han Wang, Linfeng Zhang, Zongyue Liu, Jinguo Liu, and Marín Bukov. This work was supported by the National Natural Science Foundation of China under Grants No. T2225018, No. 12188101, and No. T2121001, and the Cross-Disciplinary Key Project of Beijing Natural Science Foundation (No. Z250005).

DATA AVAILABILITY

The code and trained model weights that support the findings of this article are openly available [49].

APPENDIX A: PROOF OF THE EXACTNESS THEOREM

We prove the Exactness Theorem by constructing the wave function $\Psi = \sqrt{\rho_\theta}\, e^{iS/\hbar}$ explicitly from $\boldsymbol{s}_\theta$ and verifying that it satisfies the TDSE (1). The construction also provides the practical recipe for recovering the density and phase from a trained score.

(i) *Self-consistency*. The network $\boldsymbol{s}_\theta$ and the score $\nabla \ln \rho_\theta$ of the density ρ_θ that it induces through the flow driven by $Q[\boldsymbol{s}_\theta]$ are, *a priori*, distinct fields. Since $L \geqslant 0$ with equality iff its integrand vanishes on the support of ρ_θ, hypothesis (a) forces $\boldsymbol{s}_\theta(\boldsymbol{x}, t) = \nabla \ln \rho_\theta(\boldsymbol{x}, t)$ wherever $\rho_\theta > 0$, identifying $Q[\boldsymbol{s}_\theta]$ with the Madelung quantum potential $Q[\nabla \ln \rho_\theta]$.

(ii) *Density.* The Bohmian flow driven by $\boldsymbol{s}_\theta$ pushes ρ_0 forward to $\rho_\theta(\boldsymbol{x}(t), t) = \rho_0(\boldsymbol{x}(0))/|\det F(t)|$—the recipe for the density. Probability is therefore exactly conserved on Lagrangian particles, and the Madelung continuity equation (2) holds by construction.

(iii) *Existence of a phase.* The initial velocity $\boldsymbol{v}(0) = \nabla S(\boldsymbol{x}, 0)/m$ is irrotational. Expanding the material derivative, the Bohmian equation $m\, d\boldsymbol{v}/dt = -\nabla(V + Q)$ reads $\partial_t \boldsymbol{v} = -\nabla(V + Q)/m - (\boldsymbol{v} \cdot \nabla)\boldsymbol{v}$. Taking the curl and using the vorticity identity $\nabla \times [(\boldsymbol{v} \cdot \nabla)\boldsymbol{v}] = (\boldsymbol{v} \cdot \nabla)\boldsymbol{\omega} - (\boldsymbol{\omega} \cdot \nabla)\boldsymbol{v} + \boldsymbol{\omega}(\nabla \cdot \boldsymbol{v})$ with $\boldsymbol{\omega} \equiv \nabla \times \boldsymbol{v}$ yields the standard vorticity equation $d\boldsymbol{\omega}/dt = (\boldsymbol{\omega} \cdot \nabla)\boldsymbol{v} - \boldsymbol{\omega}(\nabla \cdot \boldsymbol{v})$, whose right-hand side is linear in $\boldsymbol{\omega}$. Hence, $\boldsymbol{\omega}(0) = 0$ is preserved along trajectories. A single-valued phase S with $\boldsymbol{v} = \nabla S/m$ therefore exists for all $t \in [0, T]$.

(iv) *Quantum Hamilton-Jacobi equation.* Substituting $\boldsymbol{v} = \nabla S/m$ into $\partial_t \boldsymbol{v} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = -\nabla(V + Q)/m$ and using the irrotational identity $(\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = \frac{1}{2}\nabla|\boldsymbol{v}|^2$ (from the general identity $(\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = \frac{1}{2}\nabla|\boldsymbol{v}|^2 + (\nabla \times \boldsymbol{v}) \times \boldsymbol{v}$ together with $\nabla \times \boldsymbol{v} = 0$) gives $\nabla[\partial_t S + |\nabla S|^2/(2m) + V + Q] = 0$,

so Eq. (3) holds up to a gauge function $C(t)$ absorbed by $S \to S - \int_0^t C(t')\,dt'$. Numerically, S is reconstructed along each trajectory by integrating Eq. (3) from the initial value $S(\boldsymbol{x}, 0)$—the recipe for the phase.

By hypothesis (b), $\rho_{\boldsymbol{\theta}} > 0$ on $[0, T]$, so Ψ is nodeless and the phase S just constructed is globally single-valued. The Madelung equations [6] [Eqs. (2)–(3)] are both satisfied; hence, Ψ solves the TDSE (1), and step (i) gives $\boldsymbol{s}_{\boldsymbol{\theta}} = \nabla \ln \rho_{\boldsymbol{\theta}} = \nabla \ln |\Psi|^2$. ■

APPENDIX B: DERIVATIVE CHAIN AND DIMENSIONAL SCALING

Computing the Fisher loss [Eq. (6)] and its parameter gradient $dL/d\boldsymbol{\theta}$ chains five orders of spatial differentiation of the learnable scalar potential $\varphi_{\boldsymbol{\theta}}(\boldsymbol{x}, t)$, followed by a BPTT pass through the trajectory. Table I summarizes the cascade.

For d degrees of freedom, each Bohmian trajectory carries its position $\boldsymbol{x} \in \mathbb{R}^d$ along with the $d \times d$ matrices F and G, co-integrated via Eq. (8). The dominant cost per leapfrog step is forming $\mathrm{Hess}(V + Q) \cdot F$ in the G update—one Hessian-vector product per column of F, $O(d^2)$ per trajectory for a fixed-width network, against $O(d)$ for a bare position update. The score-transport correction $\nabla_{\boldsymbol{x}(0)} \ln |\det F|$ in Eq. (7) is then obtained by automatic differentiation through this co-integrated ODE at the same $O(d^2)$ per-step cost.

This $O(d^2)$ cost is intrinsic to second-order dynamics. In first-order probability flows for the Fokker-Planck equation [23–27], the velocity field is a local function of $(\boldsymbol{x}, t)$ that depends linearly on the score, so the log-density evolves through the divergence $\nabla \cdot \boldsymbol{v}$—a trace of a local network Jacobian estimable by Hutchinson's trick [10,50] in $O(d)$ per step. Bohmian dynamics differs structurally: The score enters *nonlinearly* through the quantum potential $Q \propto \nabla \cdot \boldsymbol{s} + \frac{1}{2}|\boldsymbol{s}|^2$ in the acceleration $m\ddot{\boldsymbol{x}} = -\nabla(V + Q[\boldsymbol{s}_{\boldsymbol{\theta}}])$, and density transport in Eq. (7) couples to the full deformation gradient F through $F^{-\mathsf{T}}$ (not merely $\det F$). No trace shortcut can reduce the cost below $O(d^2)$.

[1] H.-D. Meyer, U. Manthe, and L. S. Cederbaum, The multi-configurational time-dependent Hartree approach, Chem. Phys. Lett. **165**, 73 (1990).

[2] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science **355**, 602 (2017).

[3] C. L. Lopreore and R. E. Wyatt, Quantum wave packet dynamics with trajectories, Phys. Rev. Lett. **82**, 5190 (1999).

[4] F. S. Mayor, A. Askar, and H. A. Rabitz, Quantum fluid dynamics in the Lagrangian representation and applications to photodissociation problems, J. Chem. Phys. **111**, 2423 (1999).

[5] R. E. Wyatt, *Quantum Dynamics with Trajectories: Introduction to Quantum Hydrodynamics, Interdisciplinary Applied Mathematics* (Springer, New York, 2005), Vol. 28.

[6] E. Madelung, Quantentheorie in hydrodynamischer Form, Z. Phys. **40**, 322 (1927).

[7] L. de Broglie, La mécanique ondulatoire et la structure atomique de la matière et du rayonnement, J. Phys. Radium **8**, 225 (1927).

[8] D. Bohm, A suggested interpretation of the quantum theory in terms of "hidden" variables. I, Phys. Rev. **85**, 166 (1952).

[9] D. Bohm, A suggested interpretation of the quantum theory in terms of "hidden" variables. II, Phys. Rev. **85**, 180 (1952).

[10] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud, Neural ordinary differential equations, in *Advances in Neural Information Processing Systems*, edited by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Curran Associates, Inc., Red Hook, NY, 2018), Vol. 31.

[11] L. Zhang, W. E, and L. Wang, Monge-Ampère flow for generative modeling, arXiv:1809.10188.

[12] Y. Song and S. Ermon, Generative modeling by estimating gradients of the data distribution, in *Advances in Neural Information Processing Systems*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., Red Hook, NY, 2019), Vol. 32.

[13] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, Score-based generative modeling through stochastic differential equations, in *International Conference on Learning Representations* (OpenReview.net, 2021).
[14] P. R. Holland, *The Quantum Theory of Motion: An Account of the de Broglie–Bohm Causal Interpretation of Quantum Mechanics* (Cambridge University Press, Cambridge, 1993).
[15] S. Garashchuk and V. A. Rassolov, Quantum dynamics with Bohmian trajectories: Energy conserving approximation to the quantum potential, Chem. Phys. Lett. **376**, 358 (2003).
[16] S. Garashchuk and V. A. Rassolov, Energy conserving approximations to the quantum potential: Dynamics with linearized quantum force, J. Chem. Phys. **120**, 1181 (2004).
[17] A. Hyvärinen, Estimation of non-normalized statistical models by score matching, J. Mach. Learn. Res. **6**, 695 (2005).
[18] J. B. Maddox and E. R. Bittner, Estimating Bohm's quantum force using Bayesian statistics, J. Chem. Phys. **119**, 6465 (2003).
[19] T. Chen, B. Xu, C. Zhang, and C. Guestrin, Training deep nets with sublinear memory cost, arXiv:1604.06174.
[20] E. Perez, F. Strub, H. de Vries, V. Dumoulin, and A. Courville, FiLM: Visual reasoning with a general conditioning layer, in *Proceedings of the AAAI Conference on Artificial Intelligence* (2018), Vol. 32.
[21] I. Fényes, Eine wahrscheinlichkeitstheoretische Begründung und Interpretation der Quantenmechanik, Z. Phys. **132**, 81 (1952).
[22] E. Nelson, Derivation of the Schrödinger equation from Newtonian mechanics, Phys. Rev. **150**, 1079 (1966).
[23] D. Maoutsa, S. Reich, and M. Opper, Interacting particle solutions of Fokker–Planck equations through gradient–log–density estimation, Entropy **22**, 802 (2020).
[24] Z. Shen, Z. Wang, S. Kale, A. Ribeiro, A. Karbasi, and H. Hassani, Self-consistency of the Fokker–Planck equation, in *Proceedings of the 35th Conference on Learning Theory* (PMLR, 2022), Vol. 178, pp. 817–841.
[25] N. M. Boffi and E. Vanden-Eijnden, Probability flow solution of the Fokker–Planck equation, Mach. Learn.: Sci. Technol. **4**, 035012 (2023).
[26] L. Li, S. Hurault, and J. Solomon, Self-consistent velocity matching of probability flows, in *Advances in Neural Information Processing Systems*, edited by A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (Curran Associates, Inc., Red Hook, NY, 2023), Vol. 36, pp. 57038–57057.
[27] X. Wu and Q. Liao, Adaptive probability flow residual minimization for high-dimensional Fokker–Planck equations, arXiv:2512.19196.
[28] Y. Lipman, R. T. Q. Chen, H. Ben-Hamu, M. Nickel, and M. Le, Flow matching for generative modeling, in *International Conference on Learning Representations* (OpenReview.net, 2023).
[29] X. Liu, C. Gong, and Q. Liu, Flow straight and fast: Learning to generate and transfer data with rectified flow, in *International Conference on Learning Representations* (OpenReview.net, 2023).
[30] M. S. Albergo and E. Vanden-Eijnden, Building normalizing flows with stochastic interpolants, in *International Conference on Learning Representations* (OpenReview.net, 2023).
[31] E. Orlova, A. Ustimenko, R. Jiang, P. Y. Lu, and R. Willett, Deep stochastic mechanics, in *Proceedings of the 41st International Conference on Machine Learning* (PMLR, 2024), Vol. 235, pp. 38779–38814.
[32] G. Carleo, F. Becca, M. Schiró, and M. Fabrizio, Localization and glassy dynamics of many-body quantum systems, Sci. Rep. **2**, 243 (2012).
[33] P. A. M. Dirac, Note on exchange phenomena in the Thomas atom, Math. Proc. Cambridge Philos. Soc. **26**, 376 (1930).
[34] A. D. McLachlan, A variational solution of the time-dependent Schrödinger equation, Mol. Phys. **8**, 39 (1964).
[35] M. Raissi, P. Perdikaris, and G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. **378**, 686 (2019).

[36] E. Hou, Y. Liu, L. Zhang, D. Ye, L. Wang, and H. Wang, A global spacetime optimization approach to the real-space time-dependent Schrödinger equation, Mach. Learn.: Sci. Technol. **7**, 035012 (2026).
[37] A. Sinibaldi, D. Hendry, F. Vicentini, and G. Carleo, Time-dependent neural Galerkin method for quantum dynamics, Phys. Rev. Lett. **136**, 120402 (2026).
[38] H. Wu, S. Liu, X. Ye, and H. Zhou, A parameterized Wasserstein Hamiltonian flow approach to solving nonlinear Schrödinger equations, J. Comput. Phys. **562**, 114986 (2026).
[39] I. Bialynicki-Birula, Z. Bialynicka-Birula, and C. Śliwa, Motion of vortex lines in quantum mechanics, Phys. Rev. A **61**, 032110 (2000).
[40] T. Chiueh, T.-P. Woo, H.-Y. Jian, and H.-Y. Schive, Vortex turbulence in linear Schrödinger wave mechanics, J. Phys. B: At. Mol. Opt. Phys. **44**, 115101 (2011).
[41] D. A. Wisniacki and E. R. Pujals, Motion of vortices implies chaos in Bohmian mechanics, Europhys. Lett. **71**, 159 (2005).
[42] C. Efthymiopoulos, C. Kalapotharakos, and G. Contopoulos, Origin of chaos near critical points of the quantum flow, Phys. Rev. E **79**, 036203 (2009).
[43] J. Stokes, B. Chen, and S. Veerapaneni, Numerical and geometrical aspects of flow-based variational quantum Monte Carlo, Mach. Learn.: Sci. Technol. **4**, 021001 (2023).
[44] D. M. Ceperley, Fermion nodes, J. Stat. Phys. **63**, 1237 (1991).
[45] T. C. Wallstrom, Inequivalence between the Schrödinger equation and the Madelung hydrodynamic equations, Phys. Rev. A **49**, 1613 (1994).
[46] H. Xie, L. Zhang, and L. Wang, *Ab-initio* study of interacting fermions at finite temperature with neural canonical transformation, J. Mach. Learn. **1**, 38 (2022).
[47] H. Xie, L. Zhang, and L. Wang, m^* of two-dimensional electron gas: A neural canonical transformation study, SciPost Phys. **14**, 154 (2023).
[48] Z. Li, H. Xie, X. Dong, and L. Wang, Deep variational free energy calculation of hydrogen Hugoniot, Phys. Rev. Lett. **136**, 076504 (2026).
[49] L. Wang, BohmianFlow: Code and data for "Quantum dynamics via score matching on Bohmian trajectories," GitHub repository (2026), https://github.com/wangleiphy/BohmianFlow.
[50] W. Grathwohl, R. T. Q. Chen, J. Bettencourt, I. Sutskever, and D. Duvenaud, FFJORD: Free-form continuous dynamics for scalable reversible generative models, in *International Conference on Learning Representations* (OpenReview.net, 2019).