

Quantum Ground States from Reinforcement Learning

Ariel Barr

ARB225@CAM.AC.UK

Willem Gispen

WG265@CAM.AC.UK

Austen Lamacraft

AL200@CAM.AC.UK

TCM Group, Cavendish Laboratory, University of Cambridge, J. J. Thomson Ave., Cambridge CB3 0HE, UK

Abstract

Finding the ground state of a quantum mechanical system can be formulated as an optimal control problem. In this formulation, the drift of the optimally controlled process is chosen to match the distribution of paths in the Feynman–Kac (FK) representation of the solution of the imaginary time Schrödinger equation. This provides a variational principle that can be used for reinforcement learning of a neural representation of the drift. Our approach is a drop-in replacement for path integral Monte Carlo, learning an optimal importance sampler for the FK trajectories. We demonstrate the applicability of our approach to several problems of one-, two-, and many-particle physics.

Keywords: Quantum Mechanics, Feynman–Kac Formula, Optimal Control, Reinforcement Learning

1. Introduction

Quantum mechanics takes place in infinite dimensional Hilbert space. Naturally, any numerical approach to solving the equations of quantum mechanics – or any other physical system with a continuum description – involves a finite truncation of this space. When we turn to *many-body* quantum mechanics, the dimension of Hilbert space necessarily grows exponentially with the number of particles relative to the finite truncation used for a single particle. To be specific, if a single particle is described by a wavefunction $\psi(\mathbf{r})$ defined on a real-space grid of linear size L (with L^3 points in three dimensions), the wavefunction of N particles $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ is defined on a grid in $3N$ dimensions of L^{3N} points.

Traditionally, this challenge has been dealt with by considering many body wavefunctions of restricted form. For example, the Hartree–Fock method employs factorized wavefunctions¹

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \psi_1(\mathbf{r}_1) \dots \psi_N(\mathbf{r}_N).$$

This reduces the memory cost to linear in the number of particles,² but represents a drastic simplification that performs especially poorly when the interaction between particles is strong. Over the years many *post–Hartree–Fock* hand-crafted improvements of the many-body wavefunction have been introduced, including Jastrow factors, and the coupled cluster and configuration interactions methods (Foulkes et al. (2001)).

The exponential growth in the complexity of many-body quantum mechanics closely parallels the curse of dimensionality encountered in computer vision and other traditional applications of machine learning. Therefore it is natural that neural methods that have recently proven successful in the latter domain be applied to quantum mechanical calculations.

1. We leave aside the issue of the statistics of indistinguishable particles for the moment.

2. The computational complexity of the Hartree–Fock method scales as the cube of the number of basis functions used to represent the $\psi_i(\mathbf{r})$.

1.1. Deep Learning Approaches to Quantum Ground States

Beginning with the work of [Carleo and Troyer \(2017\)](#), the past few years have seen numerous attempts to leverage the expressive power of deep networks to solve the quantum many-body problem.³

This initial work dealt with lattice models of spins that are a mainstay of quantum condensed matter physics, representing the wavefunction using a Restricted Boltzmann Machine. Since then neural representations of many body wavefunctions have multiplied to include lattice systems of fermions ([Nomura et al. \(2017\)](#)) and bosons ([Saito \(2017\)](#)), particles in continuous space ([Ruggeri et al. \(2018\)](#); [Kessler et al. \(2019\)](#)), systems with symmetry ([Choo et al. \(2018\)](#)), and the many-electron problem ([Han et al. \(2019\)](#); [Pfau et al. \(2019\)](#); [Hermann et al. \(2019\)](#)).

All of the above-mentioned applications of neural networks to quantum mechanics have used the Schrödinger picture, in which a representation is sought of the wavefunction. There are several mathematically equivalent formulations of quantum mechanics that provide alternatives to the Schrödinger equation, and hence potential alternative routes for the application of machine learning. In this work we show how *reinforcement learning* may be used to solve problems in quantum mechanics via the *path integral* representation.

1.2. Feynman–Kac Representation of the Ground State

Feynman’s path integral formulation of quantum mechanics starts from the Schrödinger equation ([Feynman and Hibbs \(1965\)](#))

$$i \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = [H\psi](\mathbf{r}, t), \quad (1)$$

where for a single particle the Hamiltonian operator describing motion in a potential $V(\mathbf{r})$ has the form

$$[H\phi](\mathbf{r}) = -\frac{1}{2}\nabla^2\phi(\mathbf{r}) + V(\mathbf{r})\phi(\mathbf{r}). \quad (2)$$

The wavefunction at time $t_2 > t_1$ can be expressed in terms of the wavefunction at time t_1 by

$$\psi(\mathbf{r}_2, t_2) = \int d\mathbf{r}_1 \mathcal{K}(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) \psi(\mathbf{r}_1, t_1), \quad (3)$$

where the *propagator* $\mathcal{K}(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)$ is the Green’s function of the Schrödinger equation

$$\left(i \frac{\partial}{\partial t_2} - H_{\mathbf{r}_2} \right) \mathcal{K}(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) = i\delta(\mathbf{r}_1 - \mathbf{r}_2)\delta(t_1 - t_2). \quad (4)$$

$\mathcal{K}(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)$ has the physical interpretation of the *probability amplitude* to move from \mathbf{r}_1 at time t_1 to \mathbf{r}_2 at time t_2 . The *Born rule* then states that the square modulus of the probability amplitude gives the probability for this event.

The path integral is a representation of the propagator as a ‘sum over paths’, written formally as

$$\mathcal{K}(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) = \int_{\substack{\mathbf{r}(t_1)=\mathbf{r}_1 \\ \mathbf{r}(t_2)=\mathbf{r}_2}} \mathcal{D}\mathbf{r}(t) \exp \left(i \int_{t_1}^{t_2} L(\mathbf{r}, \dot{\mathbf{r}}) dt \right) \quad (5)$$

3. There is earlier work on few-body quantum mechanics from before the deep learning era ([Lagaris et al. \(1997\)](#)). A separate strand of work is concerned with efficient parameterization of interatomic potentials that have been calculated by other means, see e.g. ([Behler \(2016\)](#)).

where $L(\mathbf{r}, \mathbf{v}) = \frac{1}{2}\mathbf{v}^2 - V(\mathbf{r})$ is the classical Lagrangian function of the system, and the domain of the integral is all paths satisfying the stated endpoint conditions.

Despite being successfully wielded by physicists for decades Feynman’s original idea has never found rigorous mathematical formulation due to the difficulties of defining a suitable measure on the path space. Kac (1949) discovered, however, that a fully rigorous path integral formula exists for the heat-type equations

$$\frac{\partial \psi(\mathbf{r}, t)}{\partial t} = -[H\psi](\mathbf{r}, t), \quad (6)$$

also known as the imaginary time Schrödinger equation. Moving to imaginary time makes the exponent in (5) real. Kac observed that the part of the exponent arising from the kinetic energy could be interpreted as a measure on Brownian paths, leading to the *Feynman–Kac (FK) formula* for the solution of (6)

$$\psi(\mathbf{r}_2, t_2) = \mathbb{E}_{\mathbf{r}(t_2)=\mathbf{r}_2} \left[\exp \left(- \int_{t_1}^{t_2} V(\mathbf{r}(t)) dt \right) \psi(\mathbf{r}(t_1), t_1) \right], \quad (7)$$

where the expectation is over Brownian paths finishing at \mathbf{r}_2 at time t_2 . In this way quantum mechanics is brought into the realm of stochastic processes, albeit in ‘imaginary time’. Whilst this formulation is therefore not of direct utility in studying quantum dynamics, it provides a very useful tool for studying *ground states*. This is because the propagator $K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)$ for (6) has a spectral representation in terms of the eigenfunctions φ_n and eigenenergies E_n of the time independent Schrödinger equation $H\varphi_n = E_n\varphi_n$ as

$$K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) = \sum_n \varphi_n(\mathbf{r}_2) \varphi_n^*(\mathbf{r}_1) e^{-E_n(t_2-t_1)} \quad (8)$$

$$\longrightarrow \varphi_0(\mathbf{r}_2) \varphi_0^*(\mathbf{r}_1) e^{-E_0(t_2-t_1)} \quad \text{as } t_2 - t_1 \rightarrow \infty. \quad (9)$$

Thus, as $t_2 - t_1 \rightarrow \infty$, only the ground state contributes.

The FK formula defines a new path measure \mathbb{P}_{FK} that differs from the Brownian measure \mathbb{P}_0 by the Radon–Nikodym derivative

$$\frac{d\mathbb{P}_{\text{FK}}}{d\mathbb{P}_0} = \mathcal{N} \exp \left(- \int_{t_1}^{t_2} V(\mathbf{r}(t)) dt \right) \quad (10)$$

where \mathcal{N} is a normalization factor. Intuitively, (10) describes paths that spend more time in the attractive regions of the potential ($V(\mathbf{r}) < 0$) and less time in the repulsive regions ($V(\mathbf{r}) > 0$). It is natural to conjecture that if $t_2 = -t_1 = T/2$ with $T \rightarrow \infty$, the distribution of $\mathbf{r}(0)$ under this measure coincides with the ground state probability distribution $|\varphi_0(\mathbf{r})|^2$ from the Born rule. To see that this is the case, consider a path that passes through $(\mathbf{r}_-, -T/2)$, $(\mathbf{r}, 0)$ and $(\mathbf{r}_+, T/2)$ for some arbitrary initial and final point \mathbf{r}_\pm . The overall propagator is then

$$K(\mathbf{r}_+, T/2; \mathbf{r}, 0) K(\mathbf{r}, 0; \mathbf{r}_-, -T/2) \sim |\varphi_0(\mathbf{r})|^2 \varphi_0(\mathbf{r}_+) \varphi_0^*(\mathbf{r}_-) e^{-E_0 T}. \quad (11)$$

Apart from a normalization factor that depends on \mathbf{r}_\pm and T ,⁴ this is just the expected ground state distribution $|\varphi_0(\mathbf{r})|^2$. Thus, the ability to sample from the FK measure for long trajectories would also allow us to sample from the ground state distribution.

4. In particular, we see that $\mathcal{N} \sim e^{E_0 T}$ in the long time limit.

1.3. Quantum Mechanics and Optimal Control

Long before Feynman, the path measure was studied for finite T and the case of a free particle ($V(\mathbf{r}) = 0$) by Schrödinger (1931, 1932) in an early exploration of the connection between his equation and the heat equation. Schrödinger sought the path measure that interpolates between given marginal distributions at the initial and final times. The generalizations of this question are now known as the *Schrödinger problem*: see Léonard (2014) for an insightful recent review with a historical survey. One of the major conclusions of these works is that the path measure is *Markov* (Jamison (1974)) and that paths $\mathbf{r}(t)$ satisfy a stochastic differential equation (SDE)

$$d\mathbf{r}_t = d\mathbf{B}_t + \mathbf{v}(\mathbf{r}_t, t)dt, \tag{12}$$

where $\mathbf{B}_t \in \mathbb{R}^d$ is a standard Brownian motion and $\mathbf{v}(\mathbf{r}_t, t)$ is a drift that is determined by the potential $V(\mathbf{r})$, as well as the initial and final conditions. The problem of finding the drift can be given an optimal control formulation, with $\mathbf{v}(\mathbf{r}_t, t)$ achieving the minimum of the cost function (Holland (1977); Fleming (1977))

$$C_T[\mathbf{v}] = \frac{1}{T} \mathbb{E} \left[\int_0^T \left[\frac{1}{2}(\mathbf{v}(\mathbf{r}_t, t))^2 + V(\mathbf{r}_t) \right] dt \right], \tag{13}$$

where the expectation is over the process (12). When this cost is minimized the path measure of this process coincides with the FK measure. For the infinite horizon $T \rightarrow \infty$ case the optimal drift has no explicit time dependence and we obtain the ground state energy as $E_0 = \lim_{T \rightarrow \infty} \min_{\mathbf{v}} C_T[\mathbf{v}]$.

The existence of a variational principle allows the tools of deep learning to be brought to bear. In this work the drift function will be parameterized by a neural network with parameters θ : $\mathbf{v}(\mathbf{r}) = \mathbf{v}_\theta(\mathbf{r})$. Finding the optimal drift can be regarded as a *reinforcement learning* problem with the cost (13) as our (negative) reward function. Although the expectation in (13) is intractable, a Monte Carlo estimate may be made by generating a batch of solutions of the SDE using standard discretizations (Kloeden and Platen (2013)). Optimization of this estimate is then possible through automatic differentiation of the estimated cost with respect to θ . Derivatives with respect to the expectation are straightforward as the SDE increments are written in terms of standard Brownian increments, analogous to the reparameterization trick used in stochastic backpropagation through deep latent Gaussian models (Rezende et al. (2014)) or variational autoencoders (Kingma and Welling (2013)).

As well as providing a route to the evaluation of the ground state energy, (12) provides a sampler for the ground state probability distribution, as the stationary distribution of the SDE with the optimal drift coincides with $|\varphi_0|^2$.

The imaginary time Feynman path integral (or FK formula) is the basis of other numerical techniques in physics, notably the path integral Monte Carlo method (Ceperley (1995)). In this approach, the state of the Monte Carlo simulation corresponds to an entire Feynman trajectory, which is updated according to a Markov chain that ensures the paths sample the FK distribution in the stationary state. Our approach by contrast learns to sample the trajectories in the optimal way. We intend our method to be a drop-in replacement for path integral Monte Carlo as far as ground state properties are concerned.

1.4. The Many-Body Problem

All of the above considerations extend straightforwardly to systems of many identical particles, with one important caveat. The wavefunction of a system of identical particles must either be completely *symmetric* or *antisymmetric* under exchange of particles, corresponding to identical bosons

or fermions (the latter described by the Pauli exclusion principle). The overall ground state of an identical particle Hamiltonian is symmetric under very general conditions and therefore corresponds to the bosonic case (Feynman (1998)). Apart from some two-electron problems where the spin state of the electrons is antisymmetric, allowing the spatial wavefunction to be symmetric (see Section 4), the many-fermion problem is out of reach of our method for now.⁵

1.5. Outline

The outline of the remainder of this paper is as follows. In Section 2 we discuss the connection between quantum mechanics and optimal control in more detail, focusing on the variational principle that can be used to learn the optimal drift. In Section 3 we introduce our neural representation of the drift, with particular attention given to permutation equivariance which is a feature of systems of indistinguishable particles. We also describe the algorithm used to learn the drift. Section 4 describes experiments on some simple physical systems. Finally, in Section 5 we summarize our findings and provide an outlook to future work.

2. Quantum Mechanics and Optimal Control

2.1. Fokker–Planck to Schrödinger

The most direct way to establish a link between stochastic processes and quantum mechanics is via a standard mapping between the Fokker–Planck and Schrödinger equations. This is discussed for example in Risken (1996); Pavliotis (2014), but we repeat the main points here. Starting from a solution $p(\mathbf{r}, t)$ of the Fokker–Planck (FP) equation

$$\frac{\partial p(\mathbf{r}, t)}{\partial t} = \frac{1}{2} \nabla^2 p(\mathbf{r}, t) + \nabla \cdot (p(\mathbf{r}, t) \nabla U(\mathbf{r})), \quad (14)$$

with a drift $\mathbf{v}(\mathbf{r}) = -\nabla U(\mathbf{r})$ given in terms of some potential function $U(\mathbf{r})$, the stationary state $\pi(\mathbf{r})$ of this FP equation has the form of a Boltzmann distribution

$$\pi(\mathbf{r}) \propto \exp(-2U(\mathbf{r})). \quad (15)$$

If we define the function

$$\psi(\mathbf{r}, t) = \frac{p(\mathbf{r}, t)}{\sqrt{\pi(\mathbf{r})}}, \quad (16)$$

then $\psi(\mathbf{r}, t)$ satisfies the (imaginary time) Schrödinger equation

$$\frac{\partial \psi(\mathbf{r}, t)}{\partial t} = -H\psi(\mathbf{r}, t) \quad (17)$$

with Hamiltonian $H = \frac{1}{2} \nabla^2 + V_U(\mathbf{r})$, where the potential $V_U(\mathbf{r})$ has the form

$$V_U(\mathbf{r}) \equiv \frac{1}{2} [-\nabla^2 U + (\nabla U)^2] = \frac{1}{2} [\nabla \cdot \mathbf{v} + \mathbf{v}^2]. \quad (18)$$

The zero energy ground state wavefunction of this Hamiltonian is

$$\varphi_0(\mathbf{r}) = \sqrt{\pi(\mathbf{r})}. \quad (19)$$

5. The same restriction applies to the path integral Monte Carlo method, where it is known as the *sign problem*.

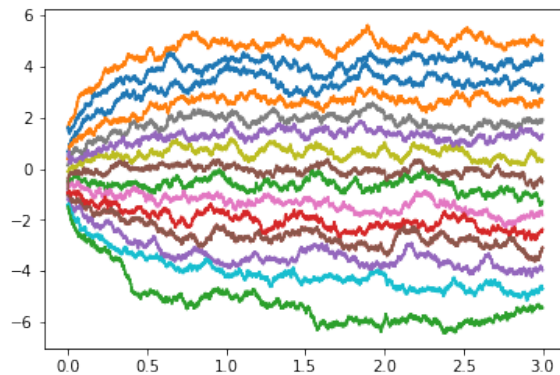


Figure 1: Trajectories of 15 particles evolving under the Dyson process (21).

The Fokker–Planck equation (14) describes the evolution of the probability density of a stochastic process described by the stochastic differential equation (SDE)

$$d\mathbf{r}_t = d\mathbf{B}_t + \mathbf{v}(\mathbf{r}_t)dt, \quad (20)$$

where \mathbf{B}_t is a standard Brownian motion. We emphasize that the quantum probability in the ground state $|\varphi_0(\mathbf{r})|^2$ coincides with the classical stationary distribution π_0 of this process.

2.1.1. EXAMPLE: CALOGERO–SUTHERLAND AND DYSON BROWNIAN MOTION

An instructive example of a many body problem and the associated stochastic process is provided by the Calogero–Sutherland model describing particles in one dimension in a harmonic potential interacting by an inverse square potential (Sutherland (1972))

$$H = \sum_i \frac{1}{2} \left[-\frac{\partial^2}{\partial x_i^2} + x_i^2 \right] + \lambda(\lambda - 1) \sum_{i < j} \frac{1}{(x_i - x_j)^2}.$$

In this case, the ground state is known exactly and has the form of the Gaussian ground state wavefunction of the harmonic oscillator multiplied by a pairwise Jastrow factor

$$\Phi_0(x_1, \dots, x_N) = \prod_{i < j} |x_i - x_j|^\lambda \exp \left(-\frac{1}{2} \sum_i x_i^2 \right).$$

The drift of the associated SDE is

$$v_i = \partial_i \log \Phi_0 = -x_i + \lambda \sum_{j \neq i} \frac{1}{x_i - x_j} \quad (21)$$

corresponding to Dyson’s Brownian motion (Dyson (1962), see Figure 1).

2.2. The Variational Principle

In order to turn the above connection into a calculational tool, we exploit a connection between the path measure \mathbb{P}_v of the SDE (20) and the corresponding quantity \mathbb{P}_{FK} from the FK formula. For a FP measure that relates the ground state distribution $\pi_0(\mathbf{r}_{0,T})$ at $t = 0$ and $t = T$, the log likelihood ratio (log Radon–Nikodym derivative) of the two measures obeys

$$\begin{aligned} \log \left(\frac{d\mathbb{P}_v}{d\mathbb{P}_{\text{FK}}} \right) &= \ell_T - E_0T + \log \left(\frac{\varphi_0(\mathbf{r}_0)}{\varphi_0(\mathbf{r}_T)} \right) \\ \ell_T &\equiv \int \mathbf{v}(\mathbf{r}_t) \cdot d\mathbf{B}_t + \int dt \left(\frac{1}{2} |\mathbf{v}(\mathbf{r}_t)|^2 + V(\mathbf{r}_t) \right), \end{aligned} \quad (22)$$

See Appendix A for a heuristic proof. The Kullback–Leibler (KL) divergence of the two measures is then

$$D_{\text{KL}}(\mathbb{P}_v \| \mathbb{P}_{\text{FK}}) = \mathbb{E}_{\mathbb{P}_v} \left[\ell_T - E_0T + \log \left(\frac{\varphi_0(\mathbf{r}_0)}{\varphi_0(\mathbf{r}_T)} \right) \right]. \quad (23)$$

Note that (23) is true for any \mathbf{r}_0 . If we additionally average over \mathbf{r}_0 drawn from the stationary distribution of the SDE, then the distributions of \mathbf{r}_T and \mathbf{r}_0 coincide, and the final term vanishes. Because the KL divergence is positive $D_{\text{KL}}(\mathbb{P}_v \| \mathbb{P}_{\text{FK}}) \geq 0$ we then have $\mathbb{E}_{\mathbb{P}_v}[\ell_T] \geq E_0T$, with equality when the two measures match. This is the variational principle that forms the basis of our approach. For $T \rightarrow \infty$ we can ignore the final term so that

$$\lim_{T \rightarrow \infty} \frac{\mathbb{E}_{\mathbb{P}_v}[\ell_T]}{T} \geq E_0, \quad (24)$$

irrespective of the initial state distribution (as $T \rightarrow \infty$ the final state distribution will be the stationary state of the SDE, assuming ergodicity). Using the fact that Brownian increments have zero expectation recovers Holland’s cost (13).

3. Reinforcement Learning

We have identified a variational principle for the drift with a time integrated cost function. Learning the optimal drift may be regarded as an entropy regularized reinforcement learning (RL) problem. The usual Markov decision process formulation of RL involves Markov dynamics in a state space together with a set of actions that can influence the dynamics within this space. We seek a *policy* – a distribution of actions to be taken at each step given the current state – optimal with respect to a time integrated reward function. In the present case, optimizing over all possible drift functions corresponds to complete controllability of the state, instead of a restriction to a (possibly finite) set of actions. As we have seen in Section 2.2, however, the optimization problem is regularized by a KL divergence that originates from the kinetic energy, and penalizes deviations of the stochastic process from the ‘passive’ Brownian dynamics. In RL the resulting cost gives rise to *linearly solvable Markov decision problems* (Todorov (2007, 2008)).⁶

In the remainder of this section we introduce a concrete algorithm to learn the optimal drift for a given quantum many-body problem.

6. In RL this entropy regularization has a practical utility in encouraging diverse policies (Levine (2018)).

3.1. Neural representation of the drift

In practice, one cannot optimize over arbitrary drifts. We introduce a neural parameterization $v_\theta(\mathbf{r})$ of the drift function, so the quality of our approximation is limited by the expressiveness of the network. Thus we are performing a variational approximation to the optimal control problem, with the cost bounded from below by the ground state energy.

For $N > 1$ particles, the drift of each is a function of all the particle coordinates: $\mathbf{v}_{i,\theta}(\mathbf{r}_1, \dots, \mathbf{r}_N)$ $i = 1, \dots, N$. For identical particles, this drift must be *permutation equivariant*, meaning that for any permutation P of the particles

$$\mathbf{v}_{i,\theta}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \mathbf{v}_{P(i),\theta}(\mathbf{r}_{P(1)}, \dots, \mathbf{r}_{P(N)}). \quad (25)$$

We now introduce two ways in which equivariance may be achieved.

3.1.1. DEEPSETS

[Guttenberg et al. \(2016\)](#); [Zaheer et al. \(2017\)](#) identified deep neural architectures that respect permutation equivariance, building on the much earlier work of [Shawe-Taylor \(1989\)](#). We use a variant of the DEEPSETS architecture of [Zaheer et al. \(2017\)](#), in which an equivariant hidden state is based on the sum-pooling operation

$$\mathbf{h}_i = \phi \left(\mathbf{\Gamma} \mathbf{r}_i + \mathbf{\Lambda} \sum_j \mathbf{r}_j \right)$$

(biases omitted for clarity) for weight matrices $\mathbf{\Gamma}, \mathbf{\Lambda} \in \mathbb{R}^{H \times d}$, where d is the spatial dimension and H the number of hidden units. ϕ is a nonlinear activation function. These layers can then be stacked to produce a more expressive representation.

3.1.2. PAIRDRIIFT

Alternatively, we may use a representation of single particle and pair features (c.f. [Pfau et al. \(2019\)](#))

$$\mathbf{h}_i = \boldsymbol{\sigma}_1(\mathbf{r}_i) + \sum_j \boldsymbol{\pi}_1(\mathbf{r}_i - \mathbf{r}_j) \quad (26)$$

$$\mathbf{h}_{ij} = \mathbf{\Pi}_1(\mathbf{r}_i - \mathbf{r}_j). \quad (27)$$

where $\boldsymbol{\sigma}, \boldsymbol{\pi} : \mathbb{R}^d \rightarrow \mathbb{R}^H$ and $\mathbf{\Pi} : \mathbb{R}^d \rightarrow \mathbb{R}^{H \times H}$ are neural networks. The single particle hidden state \mathbf{h}_i contains information about single particle and pair locations. This process may be continued

$$\tilde{\mathbf{h}}_i = \boldsymbol{\sigma}_2(\mathbf{h}_i) + \sum_j \boldsymbol{\pi}_2(\mathbf{h}_{ij}) \quad (28)$$

$$\tilde{\mathbf{h}}_{ij} = \mathbf{\Pi}_2(\mathbf{h}_{ij}). \quad (29)$$

Finally, the permutation equivariant drift function is written in terms of the last hidden state as

$$\mathbf{v}_i = \boldsymbol{\sigma}_3(\tilde{\mathbf{h}}_i) + \sum_j \boldsymbol{\pi}_3(\tilde{\mathbf{h}}_{ij}). \quad (30)$$

Note that if we had represented the drift simply as

$$\mathbf{v}_i = \boldsymbol{\sigma}(\mathbf{r}_i) + \sum_j \boldsymbol{\pi}(\mathbf{r}_i - \mathbf{r}_j),$$

this would correspond to a pairwise Jastrow factor in the many-body wavefunction via the formula $\mathbf{v}_i = \nabla_i \log \Psi$. Repeatedly merging the pair features into the single particle features enriches the representation. In our experiments, we find it sufficient to include a single hidden layer.

3.2. Integration of the SDE

From (22), we see that to evaluate the cost function we need to solve the system of SDEs

$$\begin{aligned} d\mathbf{r}_t &= d\mathbf{B}_t + \mathbf{v}(\mathbf{r}_t)dt, \\ d\ell_t &= \mathbf{v}(\mathbf{r}_t) \cdot d\mathbf{B}_t + \left(\frac{1}{2} |\mathbf{v}(\mathbf{r}_t)|^2 + V(\mathbf{r}_t) \right) dt. \end{aligned} \quad (31)$$

As $T \rightarrow \infty$, ℓ_T satisfies $\mathbb{E}_{\mathbb{P}_v}[\ell_T] \geq E_0 T$. Numerous techniques exist for the numerical solution of such problems, beginning with the Euler–Maruyama scheme

$$\begin{aligned} \mathbf{r}_{t+\Delta t} &= \mathbf{r}_t + \Delta\mathbf{B}_t + \mathbf{v}(\mathbf{r}_t)\Delta t, \\ \ell_{t+\Delta t} &= \ell_t + \mathbf{v}(\mathbf{r}_t) \cdot \Delta\mathbf{B}_t + \left(\frac{1}{2} |\mathbf{v}(\mathbf{r}_t)|^2 + V(\mathbf{r}_t) \right) \Delta t. \end{aligned} \quad (32)$$

where $\Delta\mathbf{B}_t \sim \mathcal{N}(0, t)$ are iid Gaussian increments.

Higher order methods offering improved accuracy with longer time steps Δt are described in Kloeden and Platen (2013). For most of our experiments we used the SOSRA method from Rackauckas and Nie (2018). Implementing these methods in a framework that supports automatic differentiation allows the cost to be optimized by gradient descent with respect to the parameters θ describing the drift.

As emphasized in recent work on neural ordinary differential equations Chen et al. (2018), one can regard integration schemes like (32) as a (recurrent) residual network. This affords a clean separation of the model (the drift $\mathbf{v}_\theta(\mathbf{r})$) from the implementation and allows us to adapt the integration scheme on the fly for different epochs of training, or at evaluation time.

We evolve a batch of trajectories for $T/\Delta t$ steps of size Δt , starting from an initial state equal to the final state of the previous batch. This is analogous to training a recurrent network by back-propagation through time on infinite sequences by unrolling for a finite number of steps, stopping the gradients from propagating into the infinite past. Starting each step from the final state of the previous step means that the batch tracks closely the stationary distribution of the stochastic process, so no burn-in is required before the integration of the cost function can begin.

This procedure leads to the following issue, arising from the presence of the boundary term in (22). Its expectation is zero for a stationary process where the marginal distributions of the initial and terminal points coincide. The gradients of this expectation with respect to the drift parameters are however nonzero, as changing these parameters changes the endpoint \mathbf{r}_T . Evaluating this gradient to get the gradient of the KL in (23) is intractable because the ground state wavefunction is not known, but it may be estimated from the current drift as

$$\nabla_\theta \log \left(\frac{\varphi_0(\mathbf{r}_0)}{\varphi_0(\mathbf{r}_T)} \right) = -\nabla_\theta \mathbf{r}_T \cdot \mathbf{v}(\mathbf{r}_T) \approx -\nabla_\theta \mathbf{r}_T \cdot \mathbf{v}_\theta(\mathbf{r}_T), \quad (33)$$

an approximation that improves as the true drift $\mathbf{v}(\mathbf{r}) = \nabla \log \varphi_0(\mathbf{r})$ is approached. Although this correction may be neglected in the $T \rightarrow \infty$ limit, we find it significantly improves convergence in our experiments.

3.3. Stochastic Backpropagation

Recall that the cost function (23) is the expectation of a time integral over the stochastic process. Although the expectation is intractable, a batch of B simulated trajectories $\mathbf{r}_t^{(b)}$ can be used to make a Monte Carlo estimate as follows

$$\ell_T[\mathbf{v}_\theta] \approx \frac{1}{BT} \sum_{b,t} \left[\mathbf{v}_\theta(\mathbf{r}_t^{(b)}) \cdot \Delta \mathbf{B}_t^{(b)} + \left(\frac{1}{2} \left(\mathbf{v}_\theta(\mathbf{r}_t^{(b)}) \right)^2 + V(\mathbf{r}_t^{(b)}) \right) \Delta t \right]. \quad (34)$$

Because \mathbf{r}_t^b have been generated by an SDE discretization such as (32), backpropagating through this estimated expectation is analogous to the reparameterization trick in deep latent Gaussian models (Rezende et al. (2014)) or variational autoencoders (Kingma and Welling (2013)). This similarity has been emphasized in recent work on neural stochastic differential equations (Tzen and Raginsky (2019)). The memory cost of backpropagation is proportional to $BT/\Delta t$, and is the main bottleneck in our approach.

Though the exact $T \rightarrow \infty$ value of the cost is an upper bound on the ground state energy, we have to bear in mind that the approximations of finite time, finite batch size, and any discretization error in the integration of the SDE mean that in practice the upper bound may be violated during training, particularly as optimization converges.

3.4. Related Work

We have already mentioned wavefunction based neural approaches to quantum mechanics in Section 1.1. Here we summarize work that is technically related to ours.

Learning SDEs using stochastic optimization was introduced in Ryder et al. (2018) to infer SDE parameters from trajectories and in Tzen and Raginsky (2019) as a generative model to map a simple initial distribution of $\mathbf{r}_{-T/2}$ to a given empirical distribution of $\mathbf{r}_{T/2}$.

The path integral point of view has been applied to RL in Kappen (2007); Theodorou et al. (2010). Also from the RL side, Todorov (2007) introduced *linearly solvable Markov decision problems*. Their reward function – as in the quantum case – is a sum of a cost on the state space and the KL divergence between the passive and controlled dynamics. Todorov showed that for these problems the nonlinear Bellman equation could be transformed to a linear equation by an exponential transformation. As we have seen in Section 2, for the quantum cost this linear equation is just the Schrödinger equation.

Finally, Han et al. (2018) solved semilinear parabolic PDEs in high dimension using a representation in terms of a backward SDE with a neural representation, but with a cost function based on matching to a given terminal condition.

4. Experiments

Our experiments use PAIRDRIIFT architectures described in Section 3.1, which we found to perform best, with a single hidden layer and HardTanh activation. To ensure the trajectories remain close

to the origin from the outset, these networks are initialized to give zero drift, but we add a skip connection to give a restoring drift of simple form. For example, a linear skip connection $\mathbf{W}_{\text{skip}}\mathbf{r}_i$ to the drift of the i^{th} particle ensures that the SDE is initially a 3D Ornstein–Uhlenbeck process.

Optimization of the estimated cost (34) uses Adam (Kingma and Ba (2015)). The gradient is corrected to account for the boundary terms as described in Section 3.2

$$\nabla_{\theta} \ell_T \rightarrow \nabla_{\theta} \ell_T - \nabla_{\theta} \mathbf{r}_T \cdot \mathbf{v}_{\theta}(\mathbf{r}_T) \tag{35}$$

All models were implemented in PyTorch. For hyperparameters see Appendix B. The code is available at <https://github.com/AustenLamacraft/QuaRL>

4.1. Hydrogen Atom

Hydrogen is an electrically neutral atom containing a nucleus consisting of a single proton, and a single electron which is bound to the nucleus via the Coulomb force. Consider a single non-relativistic electron subject to an isotropic Coulomb interaction force provided by the nucleus. In atomic units⁷ the explicit form of the Hamiltonian reads

$$H = -\frac{1}{2}\nabla^2 - \frac{1}{|\mathbf{r}|} \tag{36}$$

where \mathbf{r} is the position of the electron with respect to the center of the nucleus. The Schrödinger equation for hydrogen can be solved explicitly via separation of variables, yielding the ground state wave-function

$$\varphi_0(\mathbf{r}) = \frac{1}{\sqrt{\pi}}e^{-|\mathbf{r}|}. \tag{37}$$

As the ground state of hydrogen has an exact analytic form, it is also possible to derive an exact expression for the electron’s drift. The ground state drift velocity of the electron is given by

$$\mathbf{v} = \nabla \log \varphi_0(\mathbf{r}) = -\hat{\mathbf{r}}. \tag{38}$$

The exact ground state energy of Hydrogen is $E_0 = -0.5$. We used a skip connection with $\mathbf{W}_{\text{skip}} = -\mathbf{1}_3$, and found an energy of -0.497 , which is within 0.6% of the exact value. The learned drift is illustrated in Figure 2. From (38) and Figure 2, we see that the Hydrogen ground state drift has a cusp at the position of the proton.

4.2. Helium atom and Hydrogen molecule

The Helium atom consists of two electrons bound to a nucleus of charge $+2e$. The non-relativistic two-body Hamiltonian reads

$$H = -\frac{1}{2}(\nabla_{\mathbf{r}_1}^2 + \nabla_{\mathbf{r}_2}^2) - \frac{2}{|\mathbf{r}_1|} - \frac{2}{|\mathbf{r}_2|} + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{39}$$

where \mathbf{r}_1 and \mathbf{r}_2 denote the positions of the electrons with respect to the nucleus. An analytic solution to the Schrödinger equation does not exist for Helium. Accurate numerical calculations may be found in the recent paper Baskerville et al. (2019), together with a comparison with the

⁷ Atomic units denotes energies measured in Hartrees (27.211 eV) and distances measured in Bohr radii (0.529 Å).

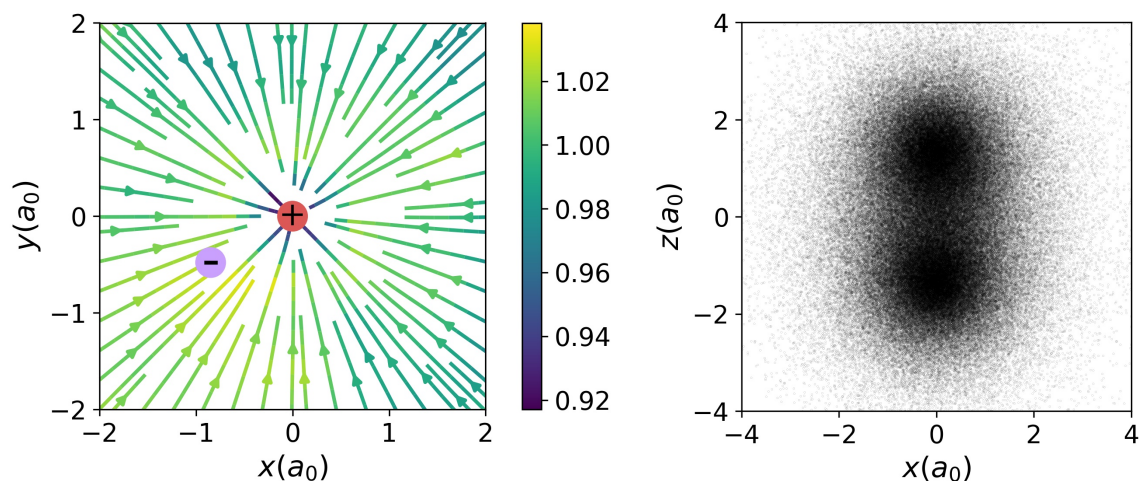


Figure 2: (Left) The learned drift field for the Hydrogen atom in the $x - y$ plane, colorscaled by its magnitude. (Right) The electron density distribution in the stretched H_2 molecule at proton separation $R = 2.8$.

Hartree–Fock method. In the Hartree–Fock (HF) approximation the wavefunction is taken as the product of two single particle wave-functions⁸

$$\Psi_{\text{HF}}(\mathbf{r}_1, \mathbf{r}_2) = \psi_{\text{HF}}(r_1) \psi_{\text{HF}}(r_2), \quad (40)$$

where the single particle wavefunction $\psi_{\text{HF}}(r)$ is optimized to minimize the energy. [Baskerville et al. \(2019\)](#) find a correlation energy – the difference between the best HF value and the numerically exact values – of 0.042, or 1.4% of the exact value.

We used the PAIRDRIFT network to model the Helium atom. The ground state energy was found to be within 0.2% of the exact value.

For both the Helium atom and the Hydrogen molecule, we used a skip connection taking advantage of Kato’s cusp conditions ([Kato \(1957\)](#)). These are conditions on the cusps that appear in the wavefunction whenever two particles with a Coulomb interaction are at the same position. Kato’s cusp conditions for wave-functions translate directly into cusp conditions for drifts: they should be proportional to \hat{r} in the limit of zero separation. For electron–electron interactions, the proportionality factor is $\frac{1}{2}$ and the drift is repulsive, while for electron–nucleus interactions, the proportionality factor is the charge of the nucleus and the drift is attractive. For our experiments on the Helium atom and Hydrogen molecule, we therefore use a skip connection that is a sum of terms \hat{r} , i.e. one term for each Coulomb interaction.

Although we could have incorporated other conditions arising from rotational symmetries, we chose not to do so because these symmetries do not generalize to other continuum systems.

8. The antisymmetry of the wavefunction demanded by the Pauli principle is achieved for the ground state by putting the spin states of the two electrons in a spin singlet. Spin triplet states – requiring an antisymmetric spatial wavefunction – are not accessible by our method.

	H atom	He atom	H ₂ molecule	H ₂ molecule ($R = 2.8$)
Numerically exact	-0.5	-2.903	-1.173	-1.071
Hartree-Fock	N/A	-2.862 (1.4%)	-1.129 (3.8%)	
Ours	-0.497 (0.6%)	-2.898 (0.2%)	-1.169 (0.3%)	-1.068 (0.3%)

Table 1: Comparison of our approach with Hartree-Fock and numerically exact values for the ground state energies. The percentages within parentheses are the relative errors of an energy estimate compared to the numerically exact value.

The Hydrogen molecule consists of two electrons orbiting a pair of bound protons at separation R , which repel each other. With the protons at positions $\pm R\hat{z}/2$, the Hamiltonian becomes

$$H = -\frac{1}{2}(\nabla_{\mathbf{r}_1}^2 + \nabla_{\mathbf{r}_2}^2) + \frac{1}{R} + \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} - \frac{1}{|\mathbf{r}_1 - R\hat{z}/2|} - \frac{1}{|\mathbf{r}_1 + R\hat{z}/2|} - \frac{1}{|\mathbf{r}_2 - R\hat{z}/2|} - \frac{1}{|\mathbf{r}_2 + R\hat{z}/2|}. \quad (41)$$

The correlation energy of the Hydrogen molecule is -0.044 Hartrees, corresponding to 3.8% of the numerically exact value.

To model the Hydrogen molecule, we also used the PAIRDIFT network with the skip connection describing the electron-electron and electron-nucleus cusps. We find a ground state energy within 0.3% of the exact value at the equilibrium separation of $R = 1.401$. For a stretched Hydrogen molecule at $R = 2.800$ (i.e. twice the equilibrium separation), our result is similarly 0.3% from the exact value.

All our results for the atomic and molecular systems are given in Table 1.

4.3. Bosons in a Harmonic Trap

As an example of a many-particle system we study a system of N identical bosons trapped in a two-dimensional isotropic harmonic potential, with the atom-atom interactions approximated by a finite-sized Gaussian contact potential (Mujal et al. (2017)). The relevant Hamiltonian reads

$$H = \sum_{i=1}^N \left[-\frac{1}{2}\nabla_i^2 + \frac{1}{2}\mathbf{r}_i^2 \right] + \frac{g}{\pi s^2} \sum_{i<j}^N \exp \left[-(\mathbf{r}_i - \mathbf{r}_j)^2 / s^2 \right] \quad (42)$$

where g and s are respectively the strength and range of the pairwise interaction.

We simulated the ground state for $N = 2, 3$, and 4 bosons for $s = 0.5$ and $g = 0 - 15$ using the PAIRDIFT neural architecture. The energies derived via the neural network are compared with those of Mujal et al. (2017) (obtained by exact diagonalization) in Section 4.3. An energy correspondence within 1% of the exact diagonalization results was observed for all parameter combinations.

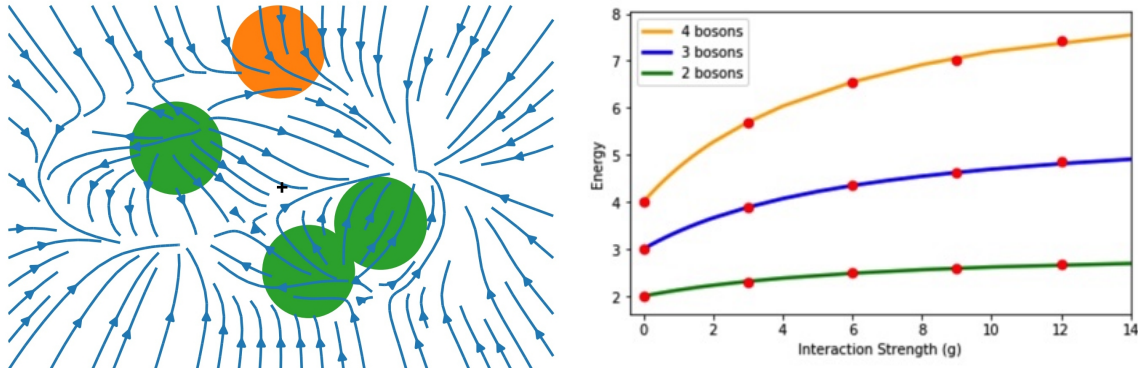


Figure 3: (Left) The learned drift field for interaction strength $g = 15$ of one boson (orange) in the presence of the others (green), together with the restoring drift arising from the harmonic potential pushing particles to the center (cross). Note the resultant drift field tends to keep the bosons apart. (Right) Comparison of ground state energies of $N = 2$, $N = 3$, and $N = 4$ bosons (blue, orange and green, respectively) with results of [Mujal et al. \(2017\)](#).

5. Conclusion

In this work we have presented a novel neural approach to many-body quantum mechanics that leverages the optimal control formulation of [Holland \(1977\)](#) to approximate the Feynman–Kac path measure. This contrasts with earlier deep learning approaches to the many-body problem that start from the Schrödinger picture. We have demonstrated the utility of our approach for some simple many-body problems.

Future work will extend this approach in the following directions:

1. Many models studied in condensed matter physics are defined on a *lattice* rather than in continuous space. In this case identical particles are handled by working with *occupation numbers* – the count of particles on each lattice site – rather than by introducing a fictitious index for each particle and then demanding invariance under labelling. The control theoretic perspective is well suited to such discrete state spaces as described in [Todorov \(2007\)](#).
2. Rotating systems or systems with finite angular momentum may be handled with a modified cost function that leads to a rotational component to the drift. This is how time reversal symmetry is broken in our formalism, corresponding to complex wavefunctions in the Schrödinger picture.
3. The memory cost of backpropagating through SDE trajectories is a major limitation of our approach that could be addressed using checkpointing [Martens and Sutskever \(2012\)](#); [Chen et al. \(2016\)](#); [Gruslys et al. \(2016\)](#). [Chen et al. \(2018\)](#) suggested that neural ODE models be trained at constant memory cost by solving the linear adjoint equation backwards in time. The difficulties associated with implementing this method in the SDE case are discussed in [Tzen and Raginsky \(2019\)](#), though very recent work has addressed this issue [Li et al. \(2020\)](#).

Acknowledgments

We thank Alex Gaunt for useful discussions. We acknowledge support from a National Science Foundation Graduate Research Fellowship under Grant No. DGE-174530 (ARB) and EPSRC Grant No. EP/P034616/1 (AL).

References

- Adam L Baskerville, Andrew W King, and Hazel Cox. Electron correlation in li+, he, h- and the critical nuclear charge system zc: energies, densities and coulomb holes. *Royal Society open science*, 6(1):181357, 2019.
- Jörg Behler. Perspective: Machine learning potentials for atomistic simulations. *The Journal of chemical physics*, 145(17):170901, 2016.
- Michelle Boué, Paul Dupuis, et al. A variational representation for certain functionals of brownian motion. *The Annals of Probability*, 26(4):1641–1659, 1998.
- Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017.
- D. M. Ceperley. Path integrals in the theory of condensed helium. *Rev. Mod. Phys.*, 67:279–355, Apr 1995. doi: 10.1103/RevModPhys.67.279. URL <https://link.aps.org/doi/10.1103/RevModPhys.67.279>.
- Tian Qi Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. In *Advances in neural information processing systems*, pages 6571–6583, 2018.
- Tianqi Chen, Bing Xu, Chiyuan Zhang, and Carlos Guestrin. Training deep nets with sublinear memory cost. *arXiv preprint arXiv:1604.06174*, 2016.
- Kenny Choo, Giuseppe Carleo, Nicolas Regnault, and Titus Neupert. Symmetries and many-body excitations with neural-network quantum states. *Physical review letters*, 121(16):167204, 2018.
- Paolo Dai Pra and Michele Pavon. On the markov processes of schrödinger, the feynman-kac formula and stochastic control. In *Realization and Modelling in System Theory*, pages 497–504. Springer, 1990.
- Freeman J Dyson. A brownian-motion model for the eigenvalues of a random matrix. *Journal of Mathematical Physics*, 3(6):1191–1198, 1962.
- RP Feynman. *Statistical mechanics: a set of lectures*. Westview Press, 1998.
- RP Feynman and AR Hibbs. Path integrals and quantum mechanics. *McGraw, New York*, 1965.
- Wendell H Fleming. Exit probabilities and optimal stochastic control. *Applied Mathematics and Optimization*, 4(1):329–346, 1977.

- W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal. Quantum monte carlo simulations of solids. *Rev. Mod. Phys.*, 73:33–83, Jan 2001. doi: 10.1103/RevModPhys.73.33. URL <https://link.aps.org/doi/10.1103/RevModPhys.73.33>.
- Audrunas Gruslys, Rémi Munos, Ivo Danihelka, Marc Lanctot, and Alex Graves. Memory-efficient backpropagation through time. In *Advances in Neural Information Processing Systems*, pages 4125–4133, 2016.
- Nicholas Guttenberg, Nathaniel Virgo, Olaf Witkowski, Hidetoshi Aoki, and Ryota Kanai. Permutation-equivariant neural networks applied to dynamics prediction. *arXiv preprint arXiv:1612.04530*, 2016.
- Jiequn Han, Arnulf Jentzen, and E Weinan. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34):8505–8510, 2018.
- Jiequn Han, Linfeng Zhang, and E Weinan. Solving many-electron schrödinger equation using deep neural networks. *Journal of Computational Physics*, 399:108929, 2019.
- Jan Hermann, Zeno Schätzle, and Frank Noé. Deep neural network solution of the electronic schrödinger equation. *arXiv preprint arXiv:1909.08423*, 2019.
- Charles J Holland. A new energy characterization of the smallest eigenvalue of the schrödinger equation. *Communications in Pure Applied Mathematics*, 30:755–765, 1977.
- Benton Jamison. Reciprocal processes. *Probability Theory and Related Fields*, 30(1):65–86, 1974.
- Mark Kac. On distributions of certain wiener functionals. *Transactions of the American Mathematical Society*, 65(1):1–13, 1949.
- Hilbert J Kappen. An introduction to stochastic control theory, path integrals and reinforcement learning. In *AIP conference proceedings*, volume 887, pages 149–181. AIP, 2007.
- Tosio Kato. On the eigenfunctions of many-particle systems in quantum mechanics. *Communications on Pure and Applied Mathematics*, 10(2):151–177, 1957. ISSN 1097-0312. doi: 10.1002/cpa.3160100201.
- Jan Kessler, Francesco Calcavecchia, and Thomas D Kühne. Artificial neural networks as trial wave functions for quantum monte carlo. *arXiv preprint arXiv:1904.10251*, 2019.
- Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015. URL <http://arxiv.org/abs/1412.6980>.
- Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.
- Peter E Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations*, volume 23. Springer Science & Business Media, 2013.

- IE Lagaris, A Likas, and DI Fotiadis. Artificial neural network methods in quantum mechanics. *Computer Physics Communications*, 104(1-3):1–14, 1997.
- Christian Léonard. A survey of the schrödinger problem and some of its connections with optimal transport. *Discrete & Continuous Dynamical Systems-A*, 34(4):1533–1574, 2014.
- Sergey Levine. Reinforcement learning and control as probabilistic inference: Tutorial and review. *arXiv preprint arXiv:1805.00909*, 2018.
- Xuechen Li, Ting-Kam Leonard Wong, Ricky TQ Chen, and David Duvenaud. Scalable gradients for stochastic differential equations. *arXiv preprint arXiv:2001.01328*, 2020.
- James Martens and Ilya Sutskever. Training deep and recurrent networks with hessian-free optimization. In *Neural networks: Tricks of the trade*, pages 479–535. Springer, 2012.
- Pere Mujal, Enric Sarlé, Artur Polls, and Bruno Juliá-Díaz. Quantum correlations and degeneracy of identical bosons in a two-dimensional harmonic trap. *Phys. Rev. A*, 96:043614, Oct 2017. doi: 10.1103/PhysRevA.96.043614. URL <https://link.aps.org/doi/10.1103/PhysRevA.96.043614>.
- Yusuke Nomura, Andrew S. Darmawan, Youhei Yamaji, and Masatoshi Imada. Restricted boltzmann machine learning for solving strongly correlated quantum systems. *Phys. Rev. B*, 96:205152, Nov 2017. doi: 10.1103/PhysRevB.96.205152. URL <https://link.aps.org/doi/10.1103/PhysRevB.96.205152>.
- Grigorios A Pavliotis. *Stochastic processes and applications: diffusion processes, the Fokker-Planck and Langevin equations*, volume 60. Springer, 2014.
- David Pfau, James S. Spencer, Alexander G. de G. Matthews, and W. M. C. Foulkes. Ab-Initio Solution of the Many-Electron Schrödinger Equation with Deep Neural Networks. *arXiv e-prints*, art. arXiv:1909.02487, Sep 2019.
- Christopher Rackauckas and Qing Nie. Stability-optimized high order methods and stiffness detection for pathwise stiff stochastic differential equations. *arXiv preprint arXiv:1804.04344*, 2018.
- Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic backpropagation and approximate inference in deep generative models. *arXiv preprint arXiv:1401.4082*, 2014.
- Hannes Risken. Fokker-planck equation. In *The Fokker-Planck Equation*, pages 63–95. Springer, 1996.
- Michele Ruggeri, Saverio Moroni, and Markus Holzmann. Nonlinear network description for many-body quantum systems in continuous space. *Phys. Rev. Lett.*, 120:205302, May 2018. doi: 10.1103/PhysRevLett.120.205302. URL <https://link.aps.org/doi/10.1103/PhysRevLett.120.205302>.
- Thomas Ryder, Andrew Golightly, A Stephen McGough, and Dennis Prangle. Black-box variational inference for stochastic differential equations. *arXiv preprint arXiv:1802.03335*, 2018.
- Hiroki Saito. Solving the bose–hubbard model with machine learning. *Journal of the Physical Society of Japan*, 86(9):093001, 2017.

- Erwin Schrödinger. *Über die umkehrung der naturgesetze*. Verlag der Akademie der Wissenschaften in Kommission bei Walter De Gruyter u . . . , 1931.
- Erwin Schrödinger. Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique. In *Annales de l'institut Henri Poincaré*, volume 2, pages 269–310, 1932.
- John Shawe-Taylor. Building symmetries into feedforward networks. In *1989 First IEE International Conference on Artificial Neural Networks,(Conf. Publ. No. 313)*, pages 158–162. IET, 1989.
- Bill Sutherland. Exact results for a quantum many-body problem in one dimension. ii. *Physical Review A*, 5(3):1372, 1972.
- Evangelos Theodorou, Jonas Buchli, and Stefan Schaal. A generalized path integral control approach to reinforcement learning. *journal of machine learning research*, 11(Nov):3137–3181, 2010.
- Emanuel Todorov. Linearly-solvable markov decision problems. In *Advances in neural information processing systems*, pages 1369–1376, 2007.
- Emanuel Todorov. General duality between optimal control and estimation. In *2008 47th IEEE Conference on Decision and Control*, pages 4286–4292. IEEE, 2008.
- Belinda Tzen and Maxim Raginsky. Neural stochastic differential equations: Deep latent gaussian models in the diffusion limit. *arXiv preprint arXiv:1905.09883*, 2019.
- Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabas Poczos, Ruslan R Salakhutdinov, and Alexander J Smola. Deep sets. In *Advances in neural information processing systems*, pages 3391–3401, 2017.

Appendix A. Probabilistic interpretation of the cost function

We show that the KL divergence between the probability of paths under the stochastic process (20) and under the Feynman–Kac measure is, in the limit of long paths, given by (22).

A very general Gibbs variational principle for functionals of Brownian motion, including functionals of FK type, appears in Boué et al. (1998). Here we provide only a heuristic argument adapted to the $T \rightarrow \infty$ limit of interest for the ground state properties.

We first consider the Radon–Nikodym (RN) derivative between \mathbb{P}_v and the distribution of Brownian paths \mathbb{P}_0 . Girsanov’s theorem tells us that this is

$$\frac{d\mathbb{P}_v}{d\mathbb{P}_0} = \exp \left(\int \mathbf{v}(\mathbf{r}_t) \cdot d\mathbf{r}_t - \frac{1}{2} \int |\mathbf{v}(\mathbf{r}_t)|^2 dt \right). \quad (43)$$

Next, the Feynman–Kac formula can be expressed as the RN derivative (Dai Pra and Pavon (1990))

$$\frac{d\mathbb{P}_{\text{FK}}}{d\mathbb{P}_0} = \frac{\tilde{\psi}(\mathbf{r}_T, T)}{\tilde{\psi}(\mathbf{r}_0, 0)} \exp \left(- \int_0^T V(\mathbf{r}_t) dt \right), \quad (44)$$

where $\tilde{\psi}(\mathbf{r}_t, t)$ is the solution of the ‘backwards’ imaginary time Schrödinger equation

$$\frac{\partial \tilde{\psi}(\mathbf{r}, t)}{\partial t} = [H\tilde{\psi}](\mathbf{r}, t). \quad (45)$$

The RN derivative (44) defines a path measure that relates initial and final distributions

$$\begin{aligned} \pi(\mathbf{r}, 0) &= \psi(\mathbf{r}, 0)\tilde{\psi}(\mathbf{r}, 0) \\ \pi(\mathbf{r}, T) &= \psi(\mathbf{r}, T)\tilde{\psi}(\mathbf{r}, T). \end{aligned} \quad (46)$$

If we take both distributions to be the ground state distribution $\pi_0(\mathbf{r})$ then

$$\frac{\tilde{\psi}(\mathbf{r}_T, T)}{\tilde{\psi}(\mathbf{r}_0, 0)} = e^{E_0 T} \frac{\varphi_0(\mathbf{r}_T)}{\varphi_0(\mathbf{r}_0)}.$$

Using these results to evaluate the overall RN derivative gives

$$\begin{aligned} \log \left(\frac{d\mathbb{P}_v}{d\mathbb{P}_{\text{FK}}} \right) &= \log \left(\frac{d\mathbb{P}_v}{d\mathbb{P}_0} \frac{d\mathbb{P}_0}{d\mathbb{P}_{\text{FK}}} \right) \\ &= \int \mathbf{v}(\mathbf{r}_t) \cdot d\mathbf{r}_t + \int dt \left(-\frac{1}{2} |\mathbf{v}(\mathbf{r}_t)|^2 + V(\mathbf{r}_t) \right) - E_0 T + \log \left(\frac{\varphi_0(\mathbf{r}_0)}{\varphi_0(\mathbf{r}_T)} \right) \\ &= \int \mathbf{v}(\mathbf{r}_t) \cdot d\mathbf{B}_t + \int dt \left(\frac{1}{2} |\mathbf{v}(\mathbf{r}_t)|^2 + V(\mathbf{r}_t) \right) - E_0 T + \log \left(\frac{\varphi_0(\mathbf{r}_0)}{\varphi_0(\mathbf{r}_T)} \right) \\ &= \ell_T - E_0 T + \log \left(\frac{\varphi_0(\mathbf{r}_0)}{\varphi_0(\mathbf{r}_T)} \right) \end{aligned} \quad (47)$$

This yields the result (22).

It is instructive to show how the RN derivative vanishes when the drift is optimal. In this case $\mathbf{v}_U(\mathbf{r}) = -\nabla U(\mathbf{r})$ where $\varphi_0(\mathbf{r}) = e^{-U(\mathbf{r})}$. Itô calculus then gives

$$dU(\mathbf{r}_t) = -\mathbf{v}_U(\mathbf{r}_t) \cdot d\mathbf{r}_t - \frac{1}{2} \nabla \cdot \mathbf{v}_U(\mathbf{r}_t) dt, \quad (48)$$

so that

$$U(\mathbf{r}_T) - U(\mathbf{r}_0) = - \int \mathbf{v}_U(\mathbf{r}_t) \cdot d\mathbf{r}_t - \frac{1}{2} \nabla \cdot \mathbf{v}_U(\mathbf{r}_t) dt.$$

Thus the RHS of (47) is

$$\log \left(\frac{d\mathbb{P}_v}{d\mathbb{P}_{\text{FK}}} \right) = \int_0^T [V(\mathbf{r}_t) - V_U(\mathbf{r}_t)] dt - E_0 T, \quad (49)$$

where $V_U = \frac{1}{2} (\nabla \cdot \mathbf{v}_U + \mathbf{v}_U^2)$ as before. This vanishes exactly when $V = V_U + E_0$ for any T . This is the analogue of the ‘zero variance property’ in variational quantum Monte Carlo (Foulkes et al. (2001)) that greatly enhances the efficiency of that algorithm.

Appendix B. Hyperparameters

We used an exponential decay of the learning rate: the learning rate was multiplied by a factor 0.95 every 10 training steps. For the bosons in a harmonic trap, we used the same hyperparameters for all simulations. The initial learning rates, as well as other training hyperparameters can be found in Table 2.

	H atom	He atom	H ₂ molecule	H ₂ molecule ($R = 2.8$)	Bosons
Batch size	2^{10}	2^{10}	2^{10}	2^{10}	2^9
Initial learning rate	10^{-2}	10^{-3}	5×10^{-4}	10^{-2}	10^{-2}
Number of time steps	2^{10}	2^{10}	2^{10}	2^{10}	2^6
Time step	0.01	0.01	0.01	0.01	0.01
Width of hidden layer	256	64	64	64	64

Table 2: Hyperparameters used in training.