Solving the Quantum Many-Body Problem with Neural Networks

Alexander Wagner

May 28th, 2019
Contents

• Many-Body Quantum Physics
• Restricted Boltzmann Machine
• Training Examples:
  • Ground State
  • Unitary Dynamics
• Summary:
  • Key Concepts
  • Future Applications
Many-Body Quantum Physics

- Open Questions:
  - Dynamical properties of high dimensional systems
  - Ground state of strongly interacting fermions
Many-Body Quantum Physics

Hilbert space is a big place!

$\dim(H) = 2^N$
Many-Body Quantum Physics

- Popular numerical approaches:
  1. Sample configurations (Variational Monte Carlo)
  2. Compression (Matrix Product States)

\[
E(\psi) = \sum_{\{S\}} \frac{\langle \psi(p)| \hat{H}| \psi(p) \rangle_{s_2}}{\langle \psi(p)| \psi(p) \rangle} = \frac{\int \langle \psi(X,p)|\hat{H}|\psi(X,p) \rangle^2 \frac{H\psi(X,p)}{\langle \psi(X,p)|\psi(X,p) \rangle} dX}{\int |\psi(X,p)|^2 dX}
\]

\[
\dim(A^{(s_i)}_i) \leq \dim(H) = 2^N
\]
Many-Body Quantum Physics

• Limiting factors:
  1. Variational Monte Carlo can fail due to sign problem
  2. Matrix product states become intractable for high dimensional systems

➤ Dimensional reduction and feature extraction with Neural Networks
  • Encode wave function in neural network
  • Neural network returns amplitude and phase upon input
Restricted Boltzmann Machine

• No knowledge of exact samples
  ➢ Use unsupervised learning

\[ S = \{S_1, S_2 \ldots S_N\} \rightarrow \psi(S) \]
Restricted Boltzmann Machine

Representation of the wave function:

$$\psi_M (\mathbf{S}, \mathbf{W}) = \sum_{\{h_i\}} \exp \left( \sum_j^N a_j S_j + \sum_i^M b_i h_i + \sum_{ij} W_{ij} h_i S_j \right)$$

Complex valued parameters!

$$F_{RB} (\mathbf{S}, \mathbf{W}) = \psi (\mathbf{S})$$

$$\mathbf{S} = \{S_1, S_2 \ldots S_N\}$$
Restricted Boltzmann Machine

Simplifying the wave function representation:

\[
\psi_M (S, W) = \sum_{\{h_i\}} \exp \left( \sum_j^N a_j S_j + \sum_i^M b_i h_i + \sum_{ij} W_{ij} h_i S_j \right)
\]

\[
= \exp \left( \sum_j^N a_j S_j \right) \cdot \sum_{\{h_i\}} \prod_i^M \exp \left( \sum_j^N b_i h_i + W_{ij} h_i S_j \right)
\]

\[
= \exp \left( \sum_j^N a_j S_j \right) \cdot \prod_i^M \exp \left( \sum_j^N b_i + W_{ij} S_j \right) + \exp \left( \sum_j^N -b_i - W_{ij} S_j \right)
\]

\[
= \exp \left( \sum_j^N a_j S_j \right) \cdot \prod_i^M 2 \cosh \left( \sum_j^N b_i + W_{ij} S_j \right)
\]
Restricted Boltzmann Machine

• Why are Restricted Boltzmann Machines a good choice?

1. Representability theorems
2. Systematic improvement through $\alpha = \frac{M}{N}$
3. Trainability
   • Stochastic reconfiguration
   • utilise symmetry
   • Self-suppression of fluctuations/noise

Giuseppe Carleo, Matthias Troyer arXiv:1606.02318
Ground State

- Solve Schrödinger equation through variational principles:

\[ H \psi = E \psi \quad \rightarrow \quad \min_W (E(W)) = \min_W \left( \frac{\langle \psi_M | H | \psi_M \rangle}{\langle \psi_M | \psi_M \rangle} \right) \]

- Two examples:

1. Transverse Field Ising:

\[ H_{TFI} = -\hbar \sum_i \sigma_i^x - \sum_{<i,j>} \sigma_i^z \sigma_j^z \]

2. Antiferromagnetic Heisenberg:

\[ H_{AFH} = \sum_{<i,j>} \sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y + \sigma_i^z \sigma_j^z \]
Ground State

Algorithm 1 Learn ground state

1: procedure VARIATIONAL METHOD
2: \[ W_0 \leftarrow \text{initialize weights randomly} \]
Ground State

• Three difficulties in training:
  1. Many local minima
  2. Strong dependence between variational parameters, instabilities
  3. Inefficient iterative methods, slow convergence

• Solutions:
  1. Random initial weights and stochastic fluctuations
  2. Stochastic reconfiguration matrix
  3. Self-suppression of statistical fluctuations/noise enables high learning rate

Algorithm 1 Learn ground state

1: procedure VARIATIONAL METHOD
2: \( W_0 \leftarrow \text{initialize randomly} \)
3: while \( E(W_k) \) not minimized do
4: \( S = \{ S^{(1)}, S^{(2)}, \ldots, S^{(L)} \} \leftarrow \text{MetropolisHastings}(|\psi_k|^2, L) \)
5: \( \mathcal{O}_i(S) = \frac{1}{\text{w}_k(S)} \partial_{W_i} \psi_k(S) \quad \forall S \in \hat{S} \)
6: \( E_{\text{loc}}(S) = \frac{\langle S | \hat{H} | S \rangle}{\langle S | S \rangle} \quad \forall S \in \hat{S} \)
7: \( F_i = \partial_{W_i} E(W) \approx \langle E_{\text{loc}} \mathcal{O}_i \rangle - \langle E_{\text{loc}} \rangle \langle \mathcal{O}_i \rangle \)
8: \( \mathbf{R} \mathbf{w}_{i+1} = \mathbf{w}_i - \gamma \mathbf{R}^{-1} F \)
9: \( \mathbf{F} = 0 \)
Ground State

- Stochastic reconfiguration matrix:

\[ R_{ii'} = \langle O_i^* O_{i'} \rangle - \langle O_i^* \rangle \langle O_{i'} \rangle \]
\[ R_{ii'}^{reg} = R_{ii'} + \epsilon \delta_{ii'} R_{ii'} \]
\[ \epsilon(k) = \max(\epsilon_0 \cdot b^k, \epsilon_{min}) \]
\[ \epsilon_0 = 100; \quad b = 0, 9; \quad \epsilon_{min} = 10^{-4} \]
\[ W_{k+1} = W_k - \gamma R^{-1} F \]
Ground State

• Self-suppression of statistical fluctuaactions/noise:

$$\text{Var}(E_{loc}) = \langle E_{loc}^2 \rangle - \langle E_{loc} \rangle^2$$

**Algorithm 1** Learn ground state

1: **procedure** VARIATIONAL METHOD
2: \( W_0 \leftarrow \text{initialize randomly} \)
3: **while** \( E(W_k) \) not minimized **do**
4: \( \hat{S} = \{S^{(1)}, S^{(2)}, \ldots, S^{(L)}\} \leftarrow \text{MetropolisHastings}(|\psi_k|^2, L) \)
5: \( \mathcal{O}_i(S) = \frac{1}{\psi_k(S)} \partial_{W_i} \psi_k(S) \quad \forall S \in \hat{S} \)
6: \( E_{loc}(S) = \frac{\langle S|H|\psi_k \rangle}{\psi_k(S)} \quad \forall S \in \hat{S} \)
7: \( F_i = \partial_{W_i} E(W) \approx \langle E_{loc}\mathcal{O}_i^* \rangle - \langle E_{loc} \rangle \langle \mathcal{O}_i^* \rangle \)
8: \( R_{ii'} = \langle \mathcal{O}_i^* \mathcal{O}_{i'} \rangle - \langle \mathcal{O}_i^* \rangle \langle \mathcal{O}_{i'} \rangle \)
9: \( W_{k+1} = W_k - \gamma R^{-1} F \)
10: \( F = 0 \) ?
Ground State

Giuseppe Carleo, Matthias Troyer arXiv:1606.02318
Ground State

TFI model for 80 spins  AFH model for 80 spins  AFH model for 10x10 spins

Three orders of magnitude less variational parameters!

Giuseppe Carleo, Matthias Troyer  arXiv:1606.02318
Unitary Dynamics

• Solve Schrödinger equation

\[ iH |\psi\rangle = \frac{d}{dt} |\psi\rangle \rightarrow \min(R^2(t; \dot{W}(t))) = \min(dist^2(\partial_t \psi((t)), -iH\psi)) \]

\[ dist(\psi, \psi') = \arccos \sqrt{\frac{\langle \psi' | \psi \rangle \langle \psi | \psi' \rangle}{\langle \psi' | \psi' \rangle \langle \psi | \psi \rangle}} \]

• Two examples:

1. Transverse Filed Ising:

\[ H_{TFI} = -\hbar \sum_i \sigma_i^x - \sum_{<i,j>} \sigma_i^z \sigma_j^z \]

2. Antiferromagnetic Heisenberg:

\[ H_{AFH} = \sum_{<i,j>} \sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y + J_z \sigma_i^z \sigma_j^z \]
Unitary Dynamics

Algorithm 2 Learn unitary dynamics

1: procedure TIME DEPENDENT VARIATIONAL METHOD
2: for every time step $t$ do
3:   $W_0(t) \leftarrow$ initialize weights randomly
4:   while $R^2(t; \hat{W}_k(t))$ not minimized do $\triangleright$ k-th iteration
5:     $\hat{S} = \{S^{(1)}, S^{(2)}, \ldots, S^{(L)}\} \leftarrow MetropolisHastings(|\psi_k(t)|^2, L)$
6:     $F_i = \partial_{W_i} \left[ R^2(t; \hat{W}_k(t)) \right] \triangleright$ generalised forces
7:     $W_{k+1}(t) = W_k(t) - i\gamma R^{-1}(t)F(t)$
8:     $F = 0$ ?
Unitary Dynamics

Spin polarisation in 1D TFI model

Spin correlations in 1D AFH model

Giuseppe Carleo, Matthias Troyer arXiv:1606.02318
Summary

• Goal:
  • Dimensional Reduction
  • Feature Extraction

• Tools:
  • RBM as variational Ansatz
  • Monte Carlo Sampling

• Tricks:
  • Stochastic Reconfiguration
  • Zero variance property
  • Utilise symmetry
  • RBM satisfies volume law for entanglement
Summary

• Future Applications:
  • Deep network architecture
  • Topological Phases of Matter (e.g. Kitaev Model)
Sources (Images in order of appearance)

- https://de.wikipedia.org/wiki/Monte-Carlo-Simulation#/media/File:Pi_statistisch.png