Solving the Quantum Many-Body Problem with Neural Networks

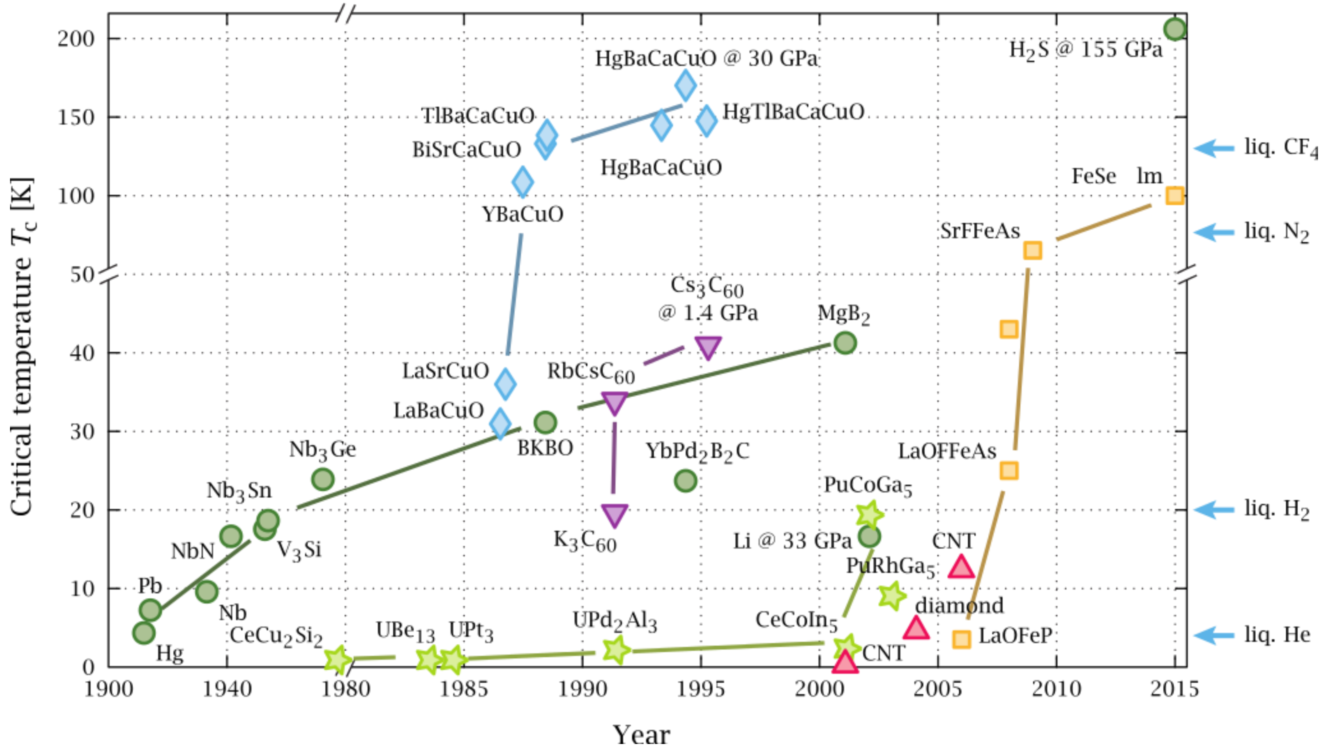
Alexander Wagner

May 28th, 2019

Contents

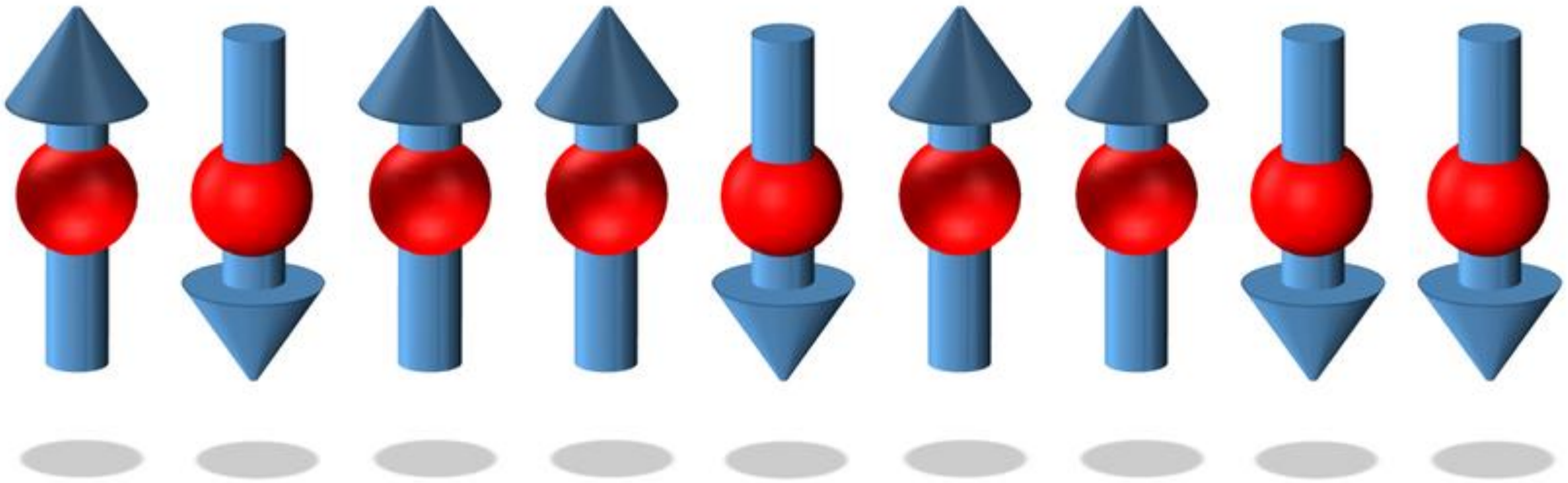
Many-Body Quantum Physics

- Open Questions:
  - Dynamical properties of high dimensional systems
  - Ground state of strongly interacting fermions



Many-Body Quantum Physics

$\dim(H) = 2^N$



Hilbert space is a big place!

Many-Body Quantum Physics

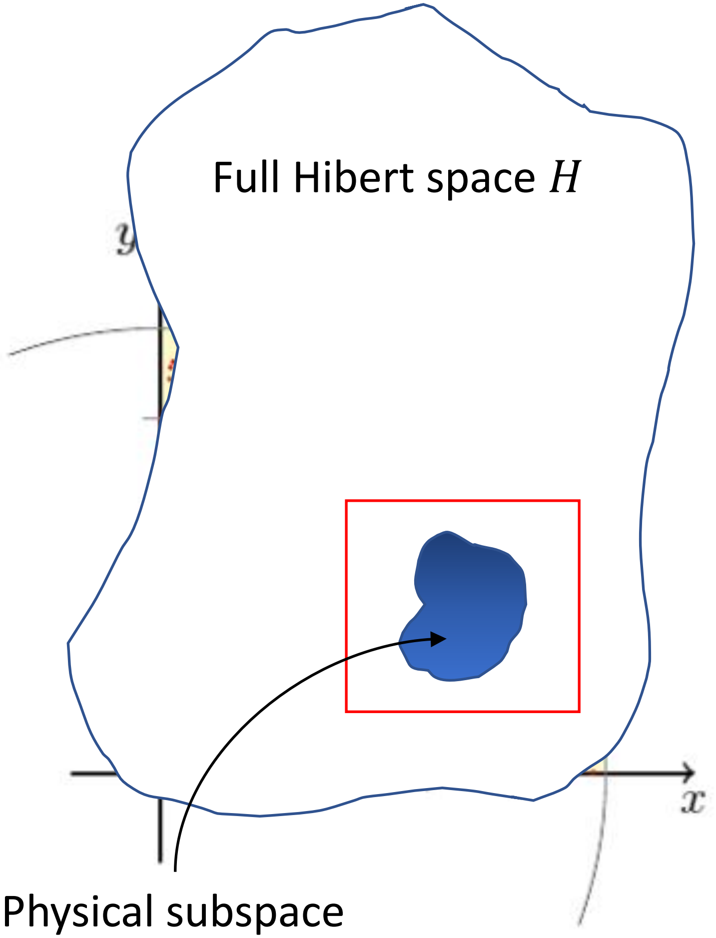
- Popular numerical approaches:
  1. Sample configurations (Variational Monte Carlo)
  2. Compression (Matrix Product States)

$$E(p) = \frac{\langle \psi(p)|\,\hat{H}\,|\psi(p)\rangle}{\langle \psi(p)|\psi(p)\rangle} = \frac{\int |\psi(X,p)|^2 \, \frac{H\psi(X,p)}{\psi(X,p)} \, dX}{\int |\psi(X,p)|^2 \, dX}$$

$$|\psi\rangle = \sum_{\{S\}} Tr\left[A_1^{(s_1)} A_2^{(s_2)} \ldots A_N^{(s_N)}\right] |s_1 s_2 \ldots s_N\rangle$$

$$\dim(A_i^{(s_i)}) \leq N^\alpha$$

$$\dim(H) = 2^N$$



Many-Body Quantum Physics

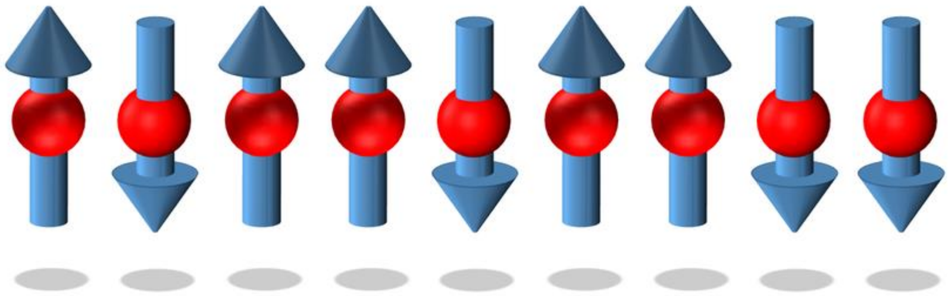
- Limiting factors:
  1. Variational Monte Carlo can fail due to sign problem
  2. Matrix product states become intractable for high dimensional systems

➢ 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



$$\mathbf{S} = \{S_1, S_2 \ldots S_N\} \quad \rightarrow \quad \psi(\mathbf{S})$$

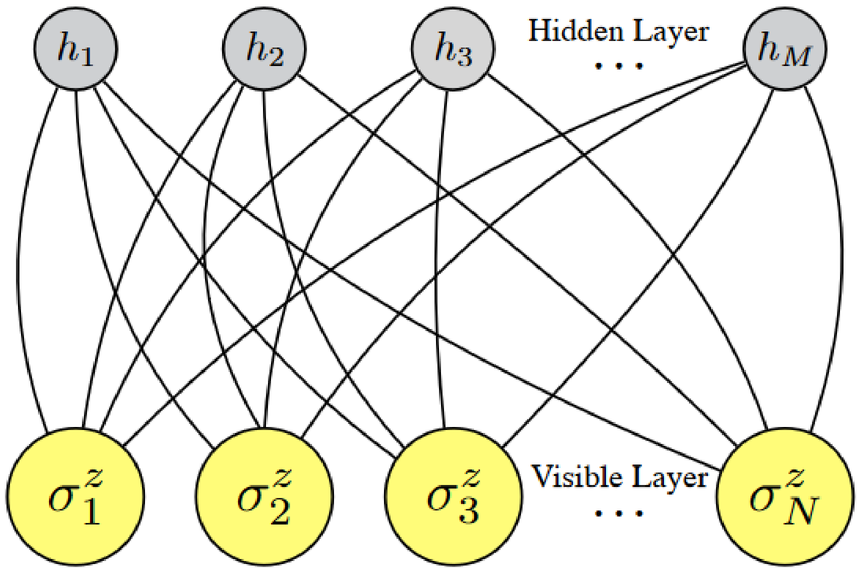
Restricted Boltzmann Machine

Representation of the wave function:

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

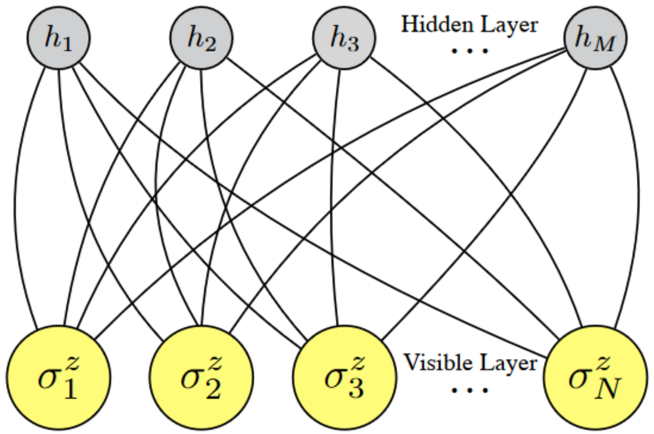
Complex valued parameters!

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



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

Restricted Boltzmann Machine



Simplifying the wave function representation:

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

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

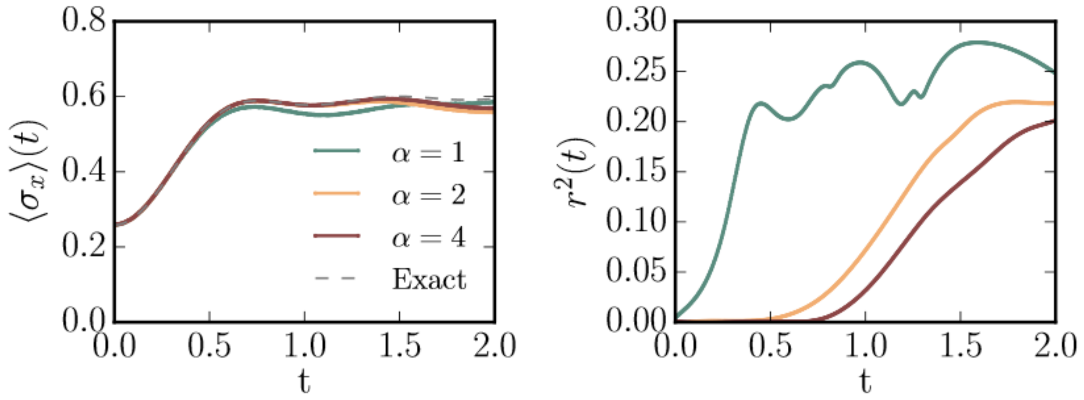
$$= \exp\left(\sum_j^N a_j S_j\right) \cdot \prod_i^M \exp\left(\sum_j^N b_i + W_{ij} S_j\right) + \exp\left(\sum_j^N -b_i - W_{ij} S_j\right)$$

$$= \exp\left(\sum_j^N a_j S_j\right) \cdot \prod_i^M 2\cosh\left(\sum_j^N b_i + W_{ij} S_j\right)$$

Restricted Boltzmann Machine

- Why are Restricted Boltzmann Machines a good choice?

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



Giuseppe Carleo, Matthias Troyer arXiv:1606.02318

Ground State

$$|\psi_M\rangle = \sum_{S'} \psi(S')\,|S'\rangle$$

- Solve Schrödinger equation through variational principles:

$$H\,|\psi\rangle = E\,|\psi\rangle \quad \rightarrow \quad \min_{\mathbf{W}}\left(E\left(\mathbf{W}\right)\right) = \min_{\mathbf{W}}\left(\frac{\langle\psi_M|\,H\,|\psi_M\rangle}{\langle\psi_M|\psi_M\rangle}\right)$$

- Two examples:

1. Transverse Field Ising:

$$H_{TFI} = -h\sum_i \sigma_i^x - \sum_{<i,j>} \sigma_i^z\sigma_j^z$$

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

procedure VARIATIONAL METHOD
    $\mathbf{W}_0 \leftarrow$ initialize weights randomly

Ground State

Algorithm 1 Learn ground state

```
procedure VARIATIONAL METHOD
    W_0 ← initialize randomly
    while E(W_k) not minimized do
        Ŝ = {S^(1), S^(2), ..., S^(L)} ← MetropolisHastings(|ψ_k|^2, L)
        O_i(S) = (1/ψ_k(S)) ∂_{W_i} ψ_k(S)    ∀S ∈ Ŝ
        E_loc(S) = ⟨S|H|ψ_k⟩ / ψ_k(S)    ∀S ∈ Ŝ
        F_i = ∂_{W_i} E(W) ≈ ⟨E_loc O_i^*⟩ − ⟨E_loc⟩⟨O_i^*⟩
        R_{ii'} = ⟨O_i^* O_{i'}⟩ − ⟨O_i^*⟩⟨O_{i'}⟩
        W_{k+1} = W_k − γ R^{-1} F
        F = 0 ?
```

- 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 fluctuactions/noise enables high learning rate

Ground State

- Stochastic reconfiguration matrix:

Algorithm 1 Learn ground state

```
procedure VARIATIONAL METHOD
    W_0 ← initialize randomly
    while E(W_k) not minimized do
        Ŝ = {S^(1), S^(2), ..., S^(L)} ← MetropolisHastings(|ψ_k|^2, L)
        O_i(S) = (1/ψ_k(S)) ∂_{W_i} ψ_k(S)    ∀S ∈ Ŝ
        E_loc(S) = ⟨S|H|ψ_k⟩ / ψ_k(S)    ∀S ∈ Ŝ
        F_i = ∂_{W_i} E(W) ≈ ⟨E_loc O_i^*⟩ − ⟨E_loc⟩⟨O_i^*⟩
        R_{ii'} = ⟨O_i^* O_{i'}⟩ − ⟨O_i^*⟩⟨O_{i'}⟩
        W_{k+1} = W_k − γ R^{-1} F
        F = 0 ?
```

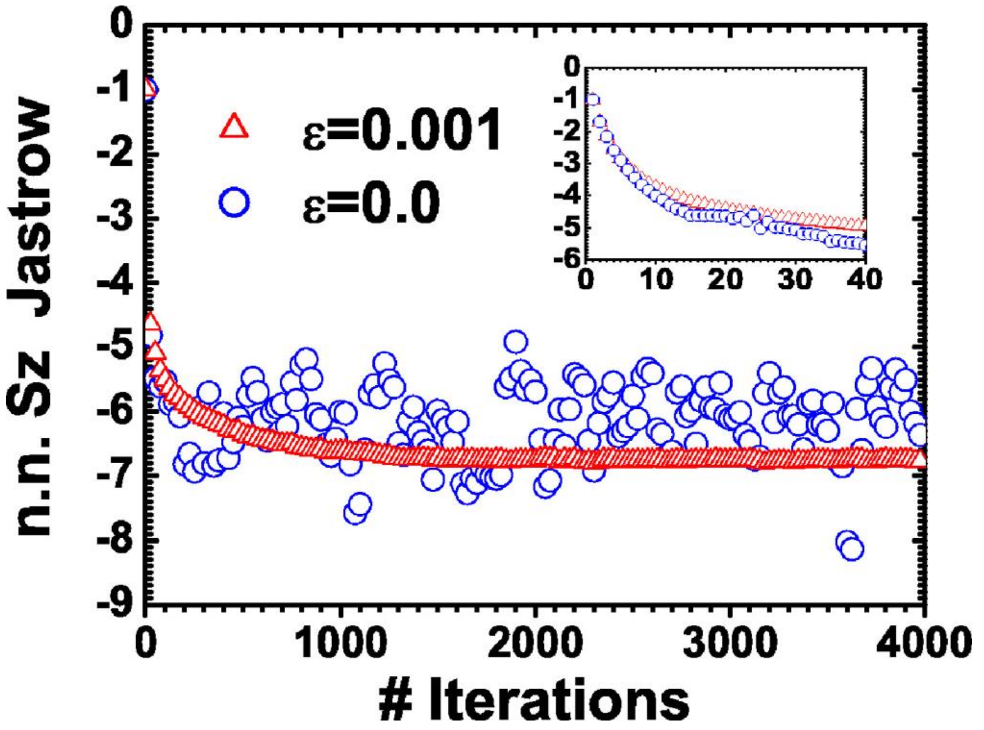
$$R_{ii'} = \langle \mathcal{O}_i^* \mathcal{O}_{i'} \rangle - \langle \mathcal{O}_i^* \rangle \langle \mathcal{O}_{i'} \rangle$$

$$R_{ii'}^{reg} = R_{ii'} + \epsilon\, \delta_{ii'} R_{ii'}$$

$$\epsilon(k) = \max(\epsilon_0 \cdot b^k, \epsilon_{min})$$

$$\epsilon_0 = 100; \quad b = 0,9; \quad \epsilon_{min} = 10^{-4}$$

$$\mathbf{W}_{k+1} = \mathbf{W}_k - \gamma \mathbf{R}^{-1} \mathbf{F}$$



Sandro Sorella et al. https://doi.org/10.1063/1.2746035

Ground State

- Self-suppression of statistical fluctuactions/noise:

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

Algorithm 1 Learn ground state

procedure VARIATIONAL METHOD

  $\mathbf{W}_0 \leftarrow$ initialize randomly

  while $E(\mathbf{W}_k)$ not minimized **do**

    $\hat{\mathbf{S}} = \{\mathbf{S}^{(1)}, \mathbf{S}^{(2)}, \ldots, \mathbf{S}^{(L)}\} \leftarrow MetropolisHastings(|\psi_k|^2, L)$

    $\mathcal{O}_i(\mathbf{S}) = \frac{1}{\psi_k(\mathbf{S})} \partial_{W_i} \psi_k(\mathbf{S}) \quad \forall \mathbf{S} \in \hat{\mathbf{S}}$

    $E_{loc}(\mathbf{S}) = \frac{\langle \mathbf{S} | H | \psi_k \rangle}{\psi_k(\mathbf{S})} \quad \forall \mathbf{S} \in \hat{\mathbf{S}}$

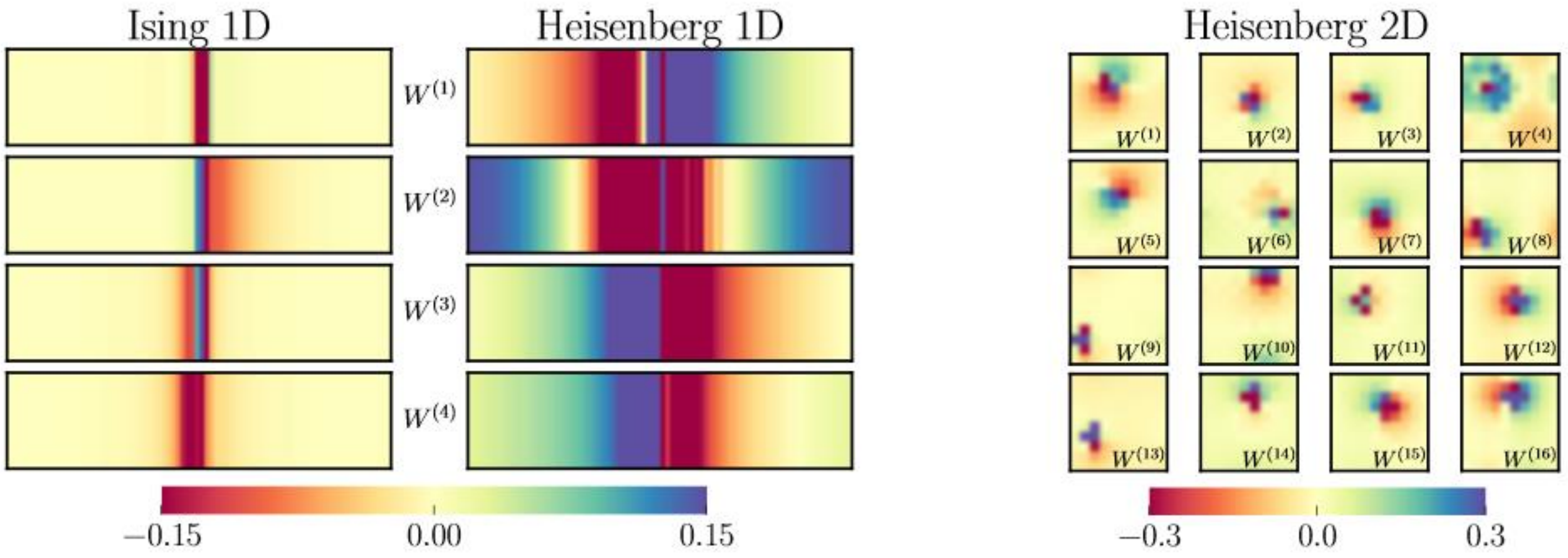
    $F_i = \partial_{W_i} E(\mathbf{W}) \approx \langle E_{loc} \mathcal{O}_i^* \rangle - \langle E_{loc} \rangle \langle \mathcal{O}_i^* \rangle$

    $R_{ii'} = \langle \mathcal{O}_i^* \mathcal{O}_{i'} \rangle - \langle \mathcal{O}_i^* \rangle \langle \mathcal{O}_{i'} \rangle$

    $\mathbf{W}_{k+1} = \mathbf{W}_k - \gamma \mathbf{R}^{-1} \mathbf{F}$

    $\mathbf{F} = 0$?

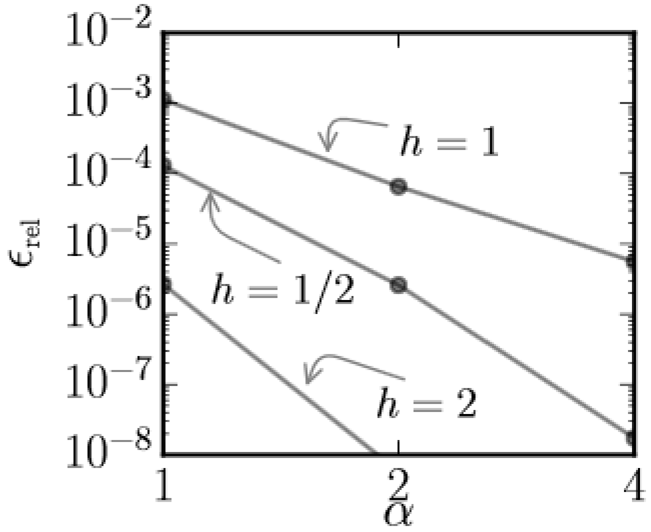
Ground State



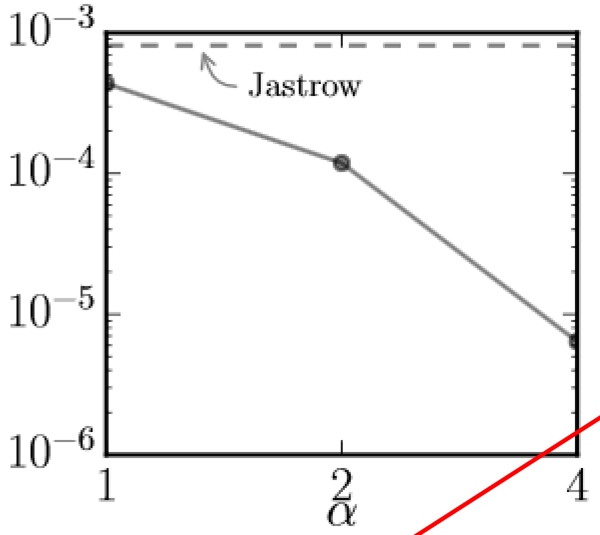
Giuseppe Carleo, Matthias Troyer arXiv:1606.02318

Ground State

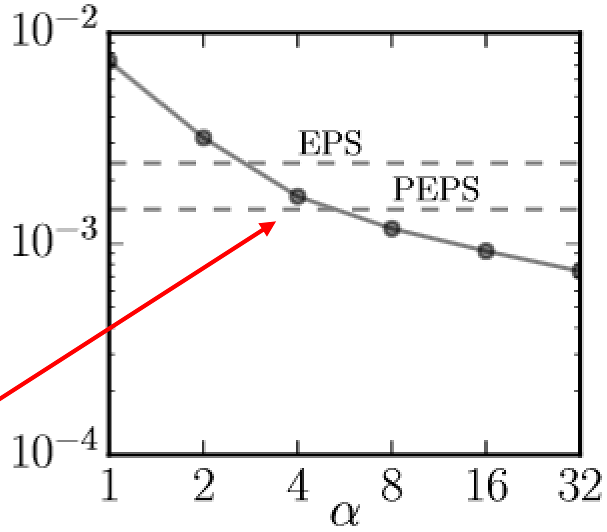
TFI model for 80 spins



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

$$iH\left|\psi\right\rangle = \frac{d}{dt}\left|\psi\right\rangle \quad \rightarrow \quad \min(R^2(t;\ \dot{\mathbf{W}}(t))) = \min(dist^2(\partial_t\psi((\mathbf{t})),\ -iH\psi))$$

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

- Two examples:

1. Transverse Filed Ising:

$$H_{TFI} = -h\sum_i \sigma_i^x - \sum_{<i,j>} \sigma_i^z\sigma_j^z$$

2. Antiferromagnetic Heisenberg:

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

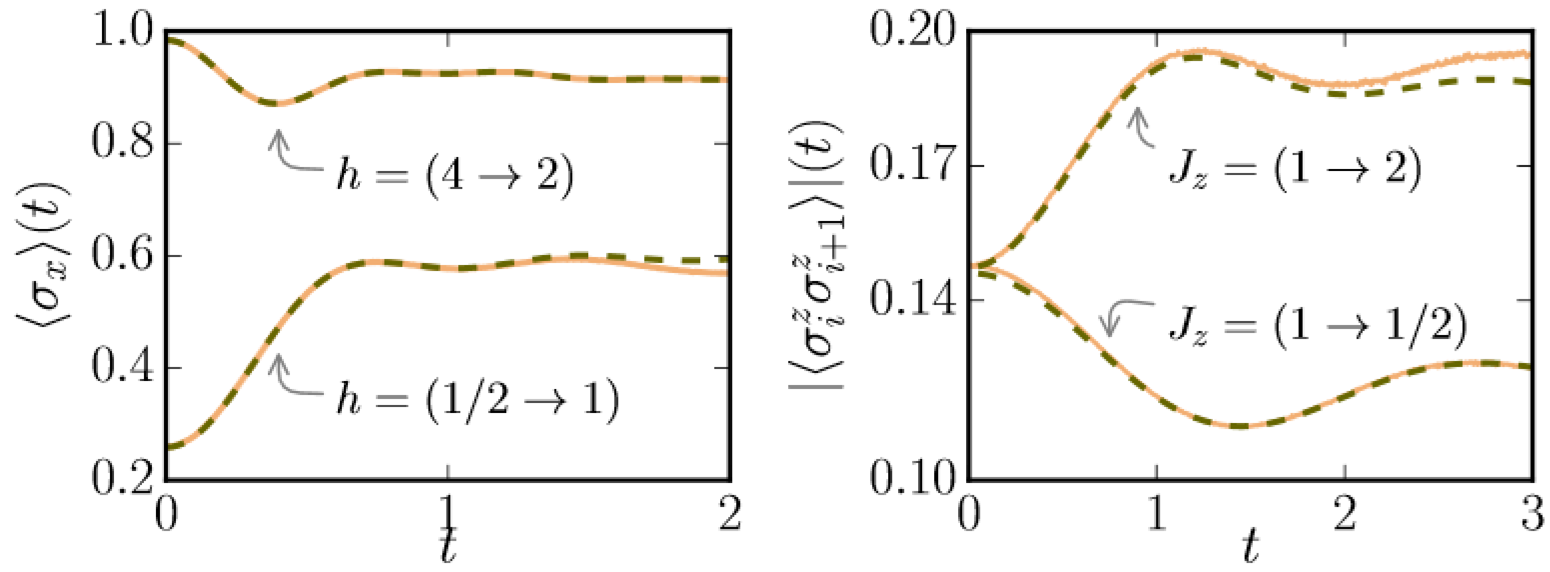
Unitary Dynamics

Algorithm 2 Learn unitary dynamics

procedure TIME DEPENDENT VARIATIONAL METHOD
  for every time step t **do**
    $\mathbf{W}_0(t) \leftarrow$ initialize weights randomly
    while $R^2(t;\ \dot{\mathbf{W}}_k(t))$ not minimized **do** ▷ k-th iteration
      $\hat{\mathbf{S}} = \{\mathbf{S}^{(1)}, \mathbf{S}^{(2)}, \ldots, \mathbf{S}^{(L)}\} \leftarrow \mathit{MetropolisHastings}(|\psi_k(t)|^2, L)$
      $F_i = \partial_{W_i} \left[R^2(t;\ \dot{\mathbf{W}}_k(t)) \right]$ ▷ generalised forces
      $\mathbf{W}_{k+1}(t) = \mathbf{W}_k(t) - i\gamma \mathbf{R}^{-1}(t)\mathbf{F}(t)$
      $\mathbf{F} = 0$?

Unitary Dynamics

Spin polarisation in 1D TFI model

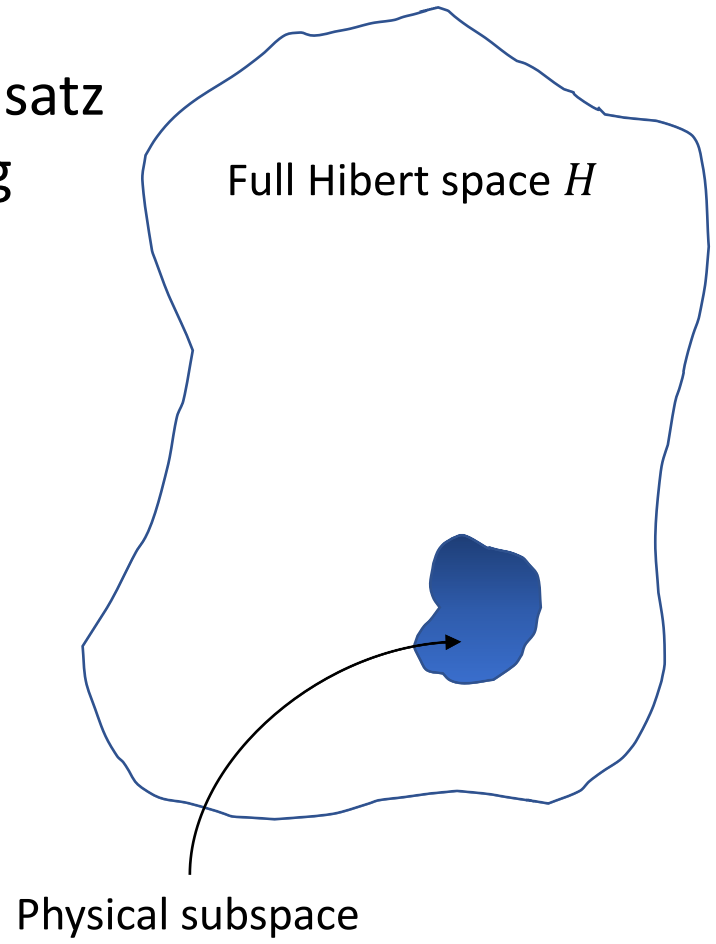


Spin correlations in 1D AFH model

Giuseppe Carleo, Matthias Troyer [arXiv:1606.02318](https://arxiv.org/abs/1606.02318)

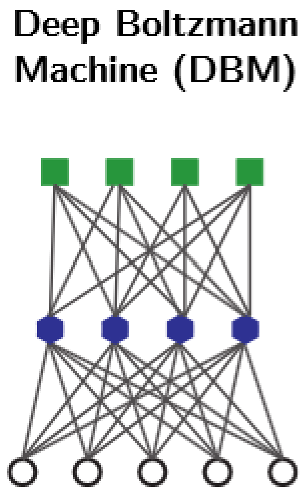
Summary

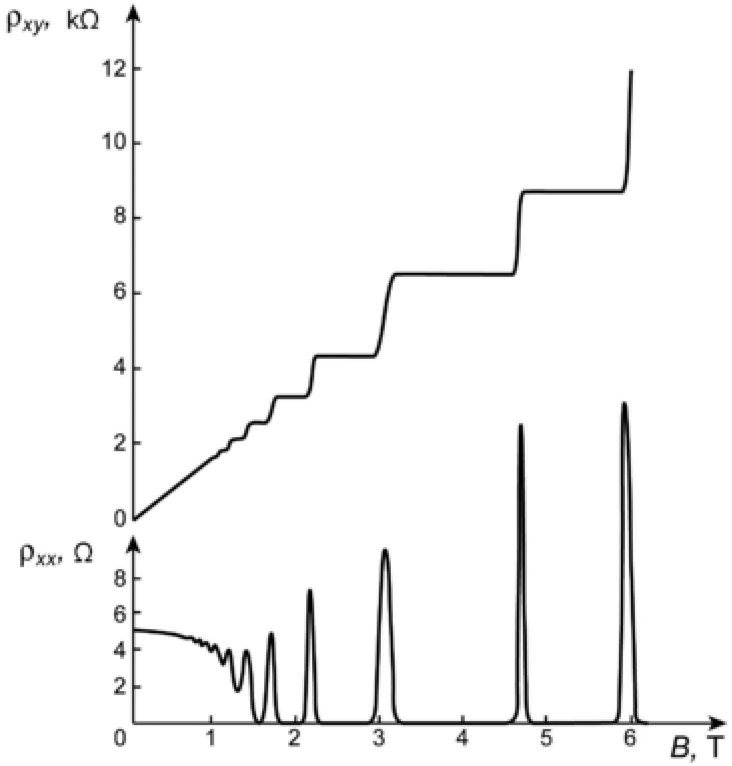
- Goal:
  - Dimensional Reduction
  - Feature Extraction
- Tools:
  - RBM as variational Ansatz
  - Monte Carlo Sampling
- Tricks:
  - Stochastic Reconfiguration
  - Zero variance property
  - Utilise symmetry
  - RBM satisfies volume law for entanglement



Summary

- Future Applications:
  - Deep network architecture
  - Topological Phases of Matter (e.g. Kitaev Model)





Sources (Images in order of appearance)

- https://en.wikipedia.org/wiki/High-temperature_superconductivity#/media/File:Timeline_of_Superconductivity_from_1900_to_2015.svg
- https://esc.fnwi.uva.nl/thesis/centraal/files/f1660852177.pdf
- https://de.wikipedia.org/wiki/Monte-Carlo-Simulation#/media/File:Pi_statistisch.png
- https://arxiv.org/pdf/1606.02318.pdf
- https://aip.scitation.org/doi/full/10.1063/1.2746035
- https://arxiv.org/pdf/1803.08823.pdf
- https://de.wikipedia.org/wiki/Quanten-Hall-Effekt#/media/File:Quantum_Hall_effect.png