Modeling the Central Spin Problem with Restricted Boltzmann Machines

Alev Orfi

Motivation

July 2020

- Ground state determination and system dynamics
- Test on two systems:
  - Transverse Field Ising Model
  - Antiferromagnetic Heisenberg Model
- Accurate results for 80-spin chains as well as 2D lattices

RESEARCH ARTICLE

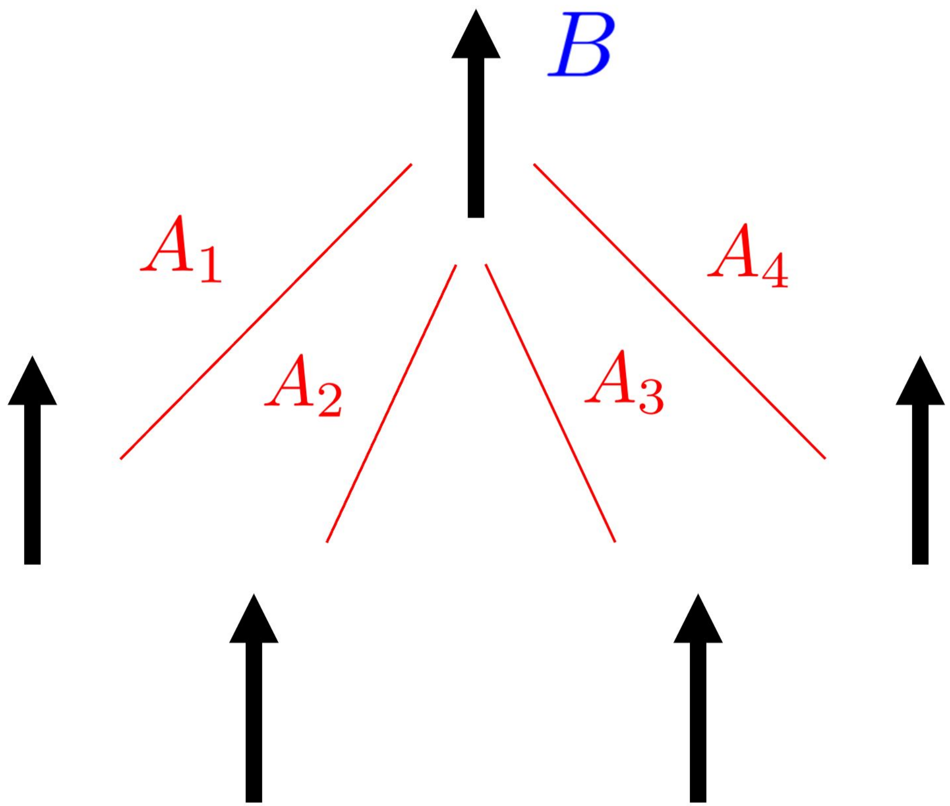
MANY-BODY PHYSICS

Solving the quantum many-body problem with artificial neural networks

Giuseppe Carleo[1*] and Matthias Troyer[1,2]

(Carleo and Troyer, 2017)

Central Spin Model



$$H = B S_0^z + \sum_{k=1}^{N-1} A_k \mathbf{S}_0 \cdot \mathbf{S}_k$$

- Decoherence of spin quantum computers
- BCS model of superconductivy

Model Validation

- Exact Diagonalization
- Block Diagonalization
- Verified through state evolution
- Computationally challenging to simulate due to exponential growth
- Hamiltonian of size $2^N \times 2^N$

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes ... \otimes \mathcal{H}_N$$

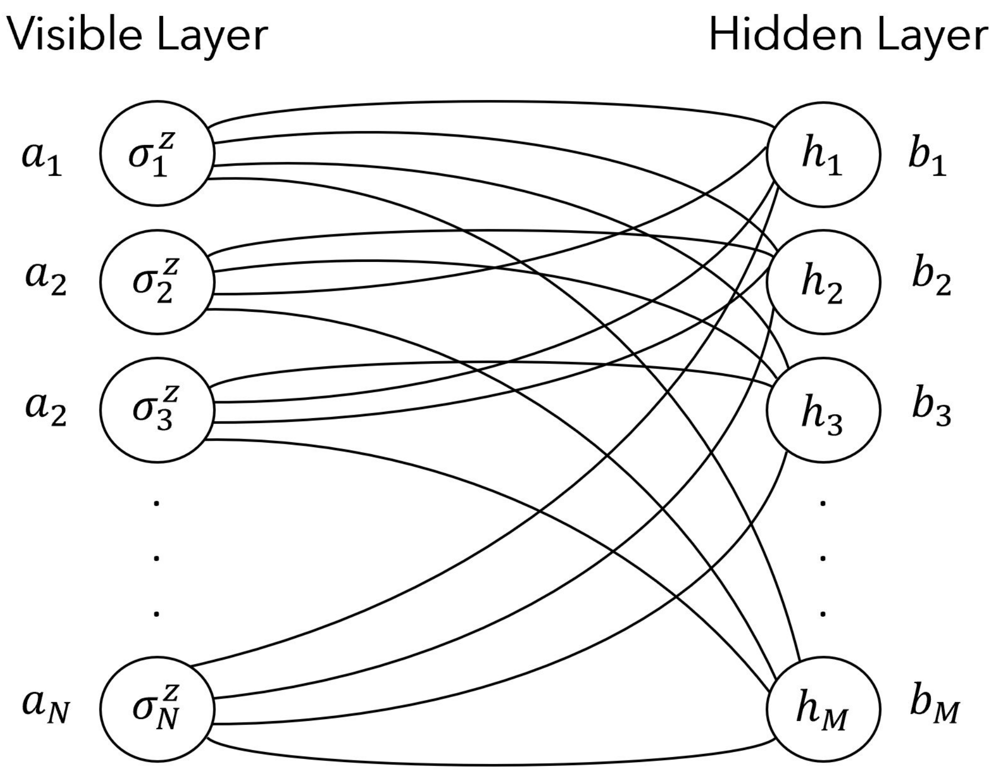
$$[H] = \begin{bmatrix} E_{m_j=s+\frac{1}{2}} & & & & \\ & \Big[2 \times 2\Big]_{m_j} & & & \\ & & \ddots & & \\ & & & \Big[2 \times 2\Big]_{m_j} & \\ & & & & E_{m_j=-s-\frac{1}{2}} \end{bmatrix}$$

Restricted Boltzmann Machine

$$\Psi(S; \mathbf{a}, \mathbf{b}, \mathbf{W}) = \sum_{\{h_i\}} e^{\sum_j a_j \sigma_j^z + \sum_i b_i h_i + \sum_{ij} W_{ij} h_i \sigma_j^z}$$

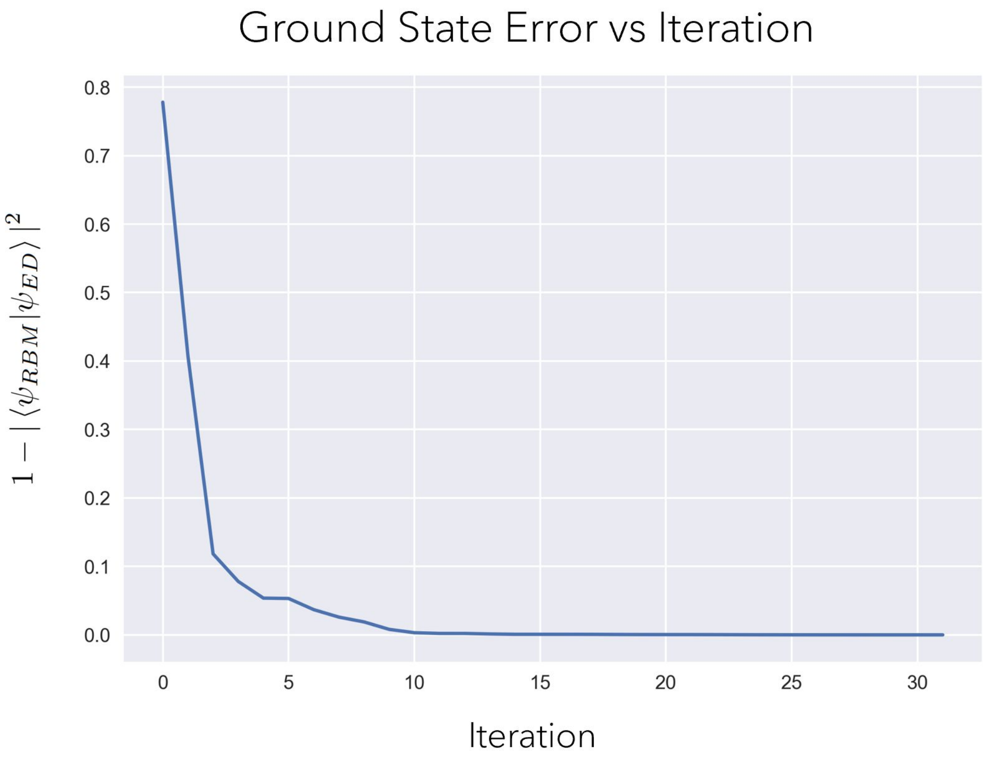
Model Parameters:

a_j	(N elements)
b_i	(M elements)
W_{ij}	(N×M elements)



Ground State Determination

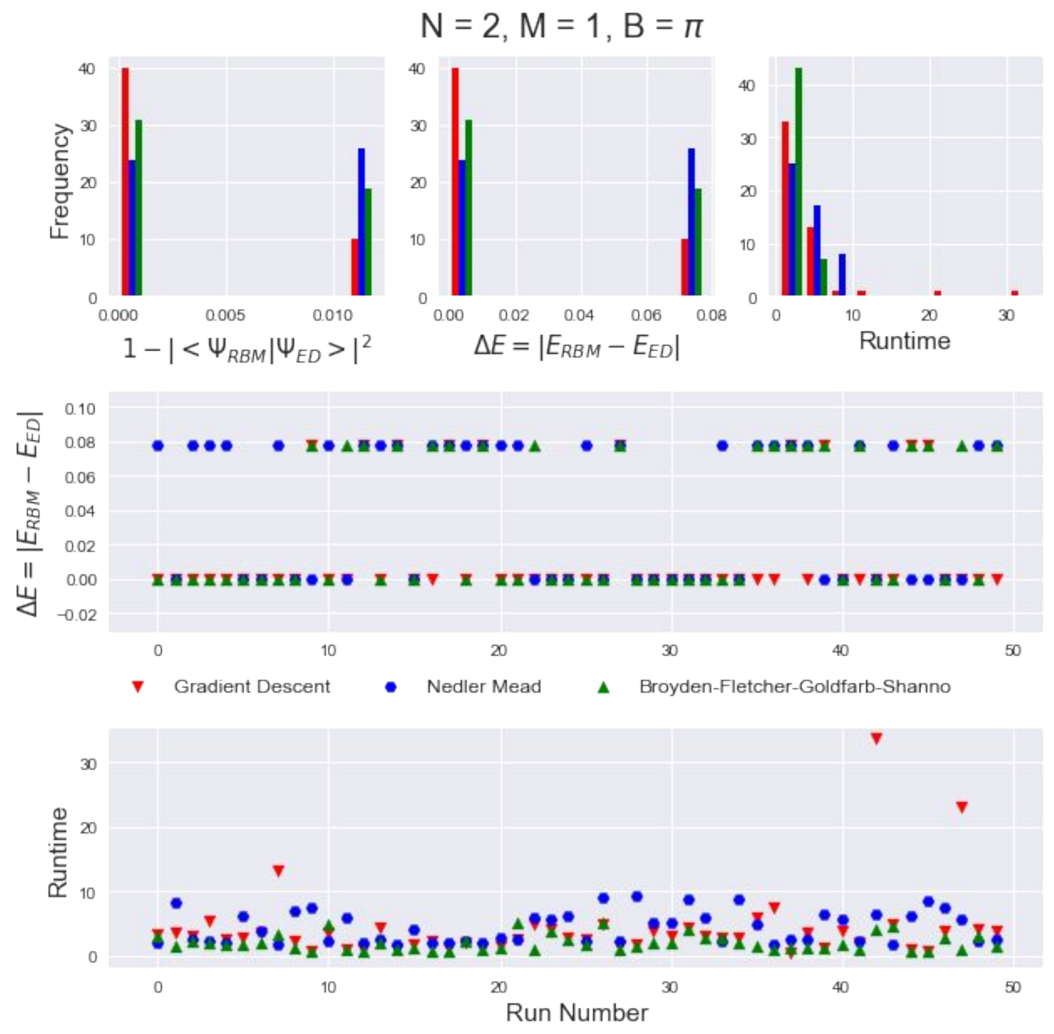
$$E(\mathbf{a}, \mathbf{b}, \mathbf{W}) = \frac{\langle \psi_{RBM} | H | \psi_{RBM} \rangle}{\langle \psi_{RBM} | \psi_{RBM} \rangle}$$



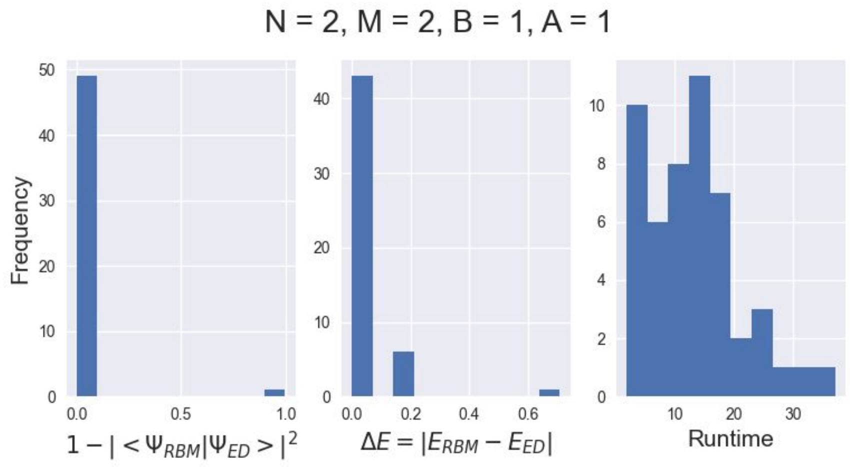
- The variational energy is minimal when ψ_{RBM} accurately models the ground state
- Reinforcement learning is achieved through this minimization
- RBM is sufficiently expressive to model the ground state

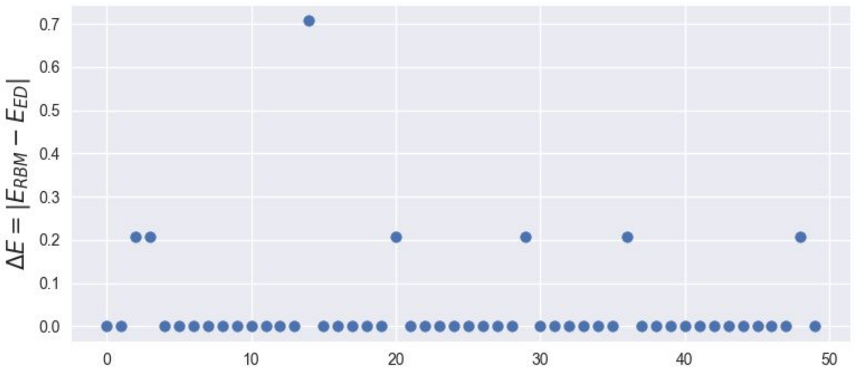
Descent Methods

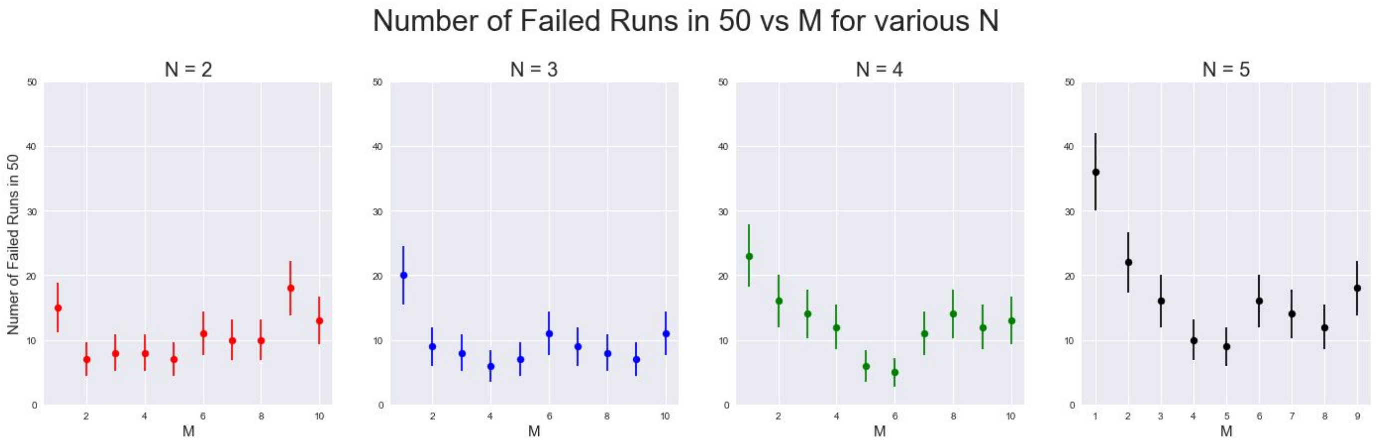
- Various descent methods were tested
- scipy.optimize conjugate gradient
- All had issues with local minimum



System Benchmarking



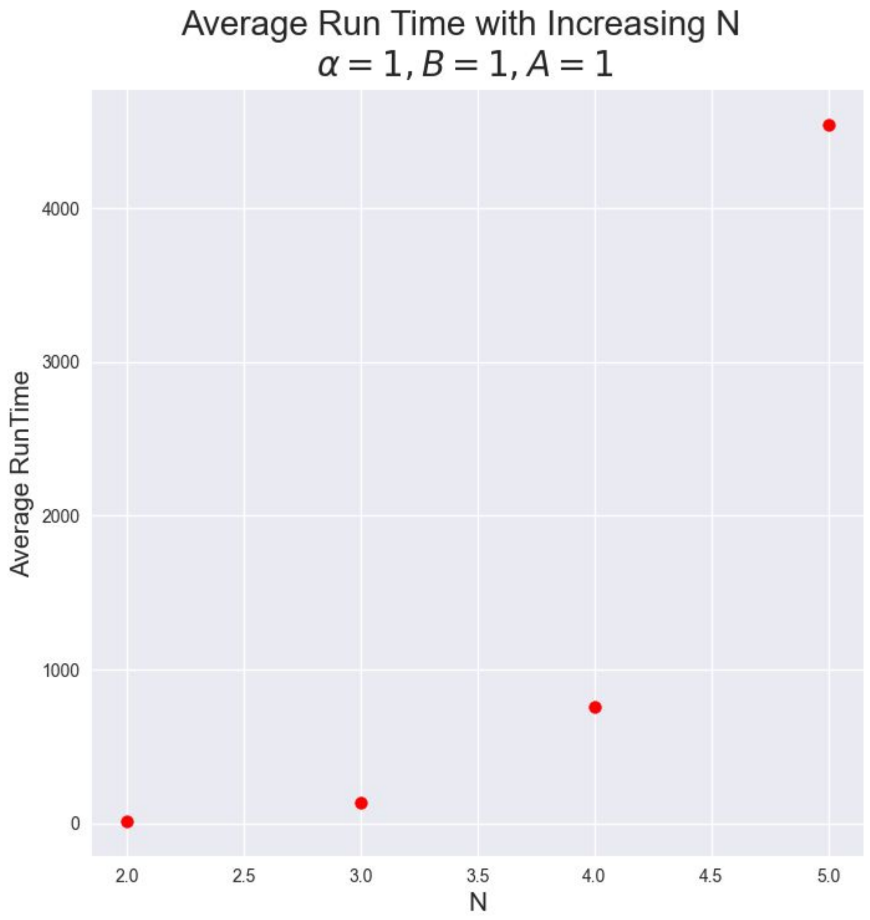




Failed runs correspond to energy error larger than 5×10^{-4}

- Frequency of local minima
- For large systems sufficient hidden nodes (M) are needed

Runtime Scaling



Here α = M/N

- Runtime increases drastically with system size
- Only small systems benchmarked
- Profiling shows majority of time is used calculating the variational energy at each step of the minimizer

NetKet Package

- C backend and parallelization
- Variational Monte Carlo
- Multiple descent methods
- Highly optimized

Carleo, Giuseppe et al. "NetKet: A Machine Learning Toolkit for Many-Body Quantum Systems"

Variational Monte Carlo

- Estimate of variational energy
- The last expectation value is over a sample of configurations
- This sampling is specifically constructed so that it is drawn from the probability distribution

$$\frac{|\Psi(\boldsymbol{\sigma})|^2}{\sum_{\boldsymbol{\sigma}'} |\Psi(\boldsymbol{\sigma}')|^2}$$

- This sampling is achieved through the Metropolis algorithm

$$\begin{aligned}
\langle \hat{H} \rangle &= \frac{\sum_{\boldsymbol{\sigma},\boldsymbol{\sigma}'} \Psi^*(\boldsymbol{\sigma}) \langle \boldsymbol{\sigma} | \hat{H} | \boldsymbol{\sigma}' \rangle \Psi(\boldsymbol{\sigma}')}{\sum_{\boldsymbol{\sigma}} |\Psi(\boldsymbol{\sigma})|^2} \\
&= \sum_{\boldsymbol{\sigma}} \left(\sum_{\boldsymbol{\sigma}'} \langle \boldsymbol{\sigma} | \hat{H} | \boldsymbol{\sigma}' \rangle \frac{\Psi(\boldsymbol{\sigma}')}{\Psi(\boldsymbol{\sigma})} \right) \frac{|\Psi(\boldsymbol{\sigma})|^2}{\sum_{\boldsymbol{\sigma}'} |\Psi(\boldsymbol{\sigma}')|^2} \\
&\approx \left\langle \sum_{\boldsymbol{\sigma}'} \langle \boldsymbol{\sigma} | \hat{H} | \boldsymbol{\sigma}' \rangle \frac{\Psi(\boldsymbol{\sigma}')}{\Psi(\boldsymbol{\sigma})} \right\rangle_{\boldsymbol{\sigma}}
\end{aligned}$$

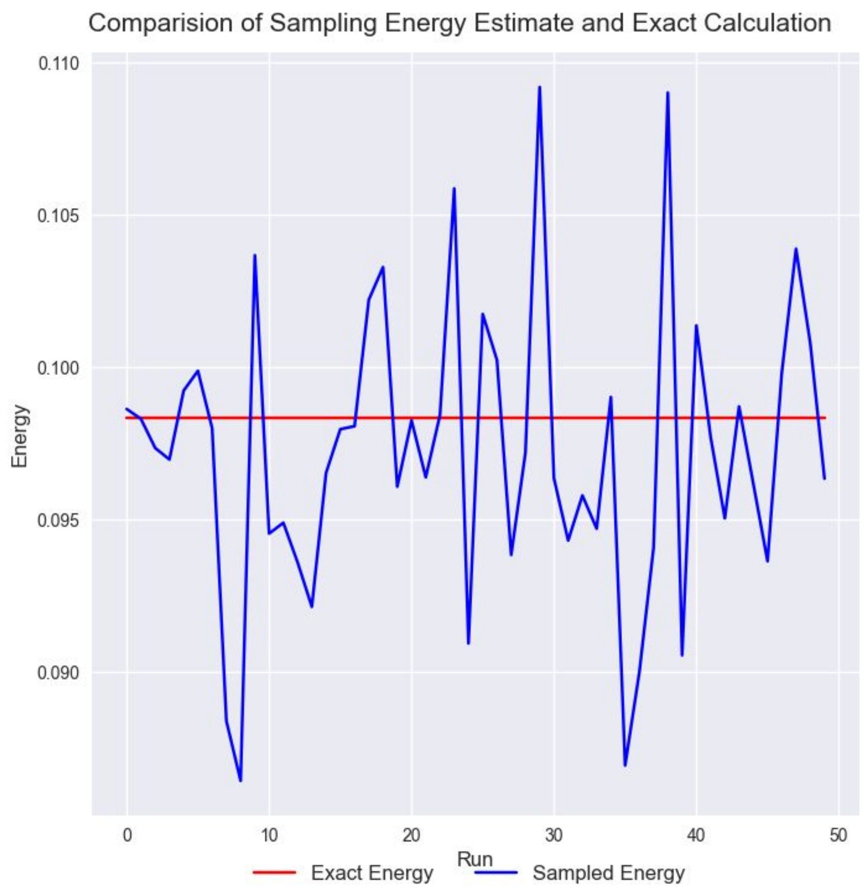
Metropolis Algorithm

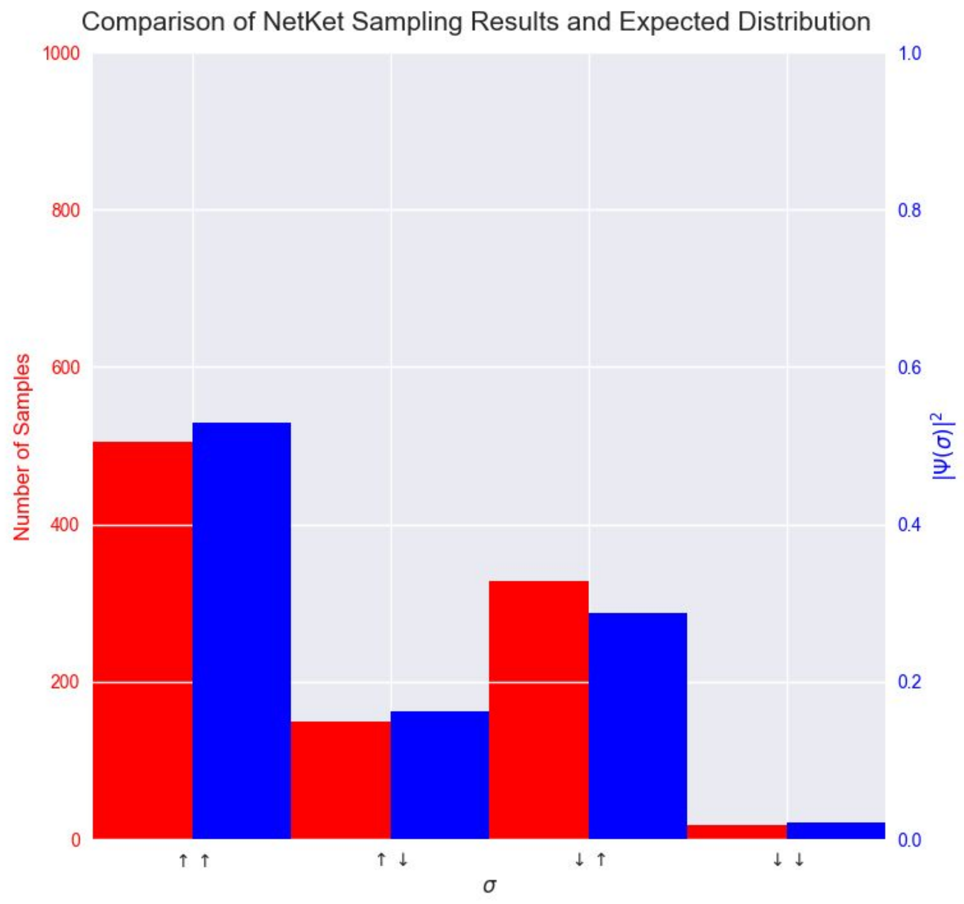
The goal is to sample configurations of spins from the distribution $\frac{|\Psi(\boldsymbol{\sigma})|^2}{\sum_{\boldsymbol{\sigma}'} |\Psi(\boldsymbol{\sigma}')|^2}$

Pseudo Code:

- Propose a new spin configuration (s → s')
  - Flip one spin at random
- Define the acceptance probability as $\alpha(s, s') = min\{\frac{|\Psi(s')|^2}{|\Psi(s)|^2}, 1\}$
- Take β from an uniform distribution between 0 and 1
- If $\beta \leq \alpha$ then accept s' as the new configuration, if not keep s

Verification of Energy Estimation

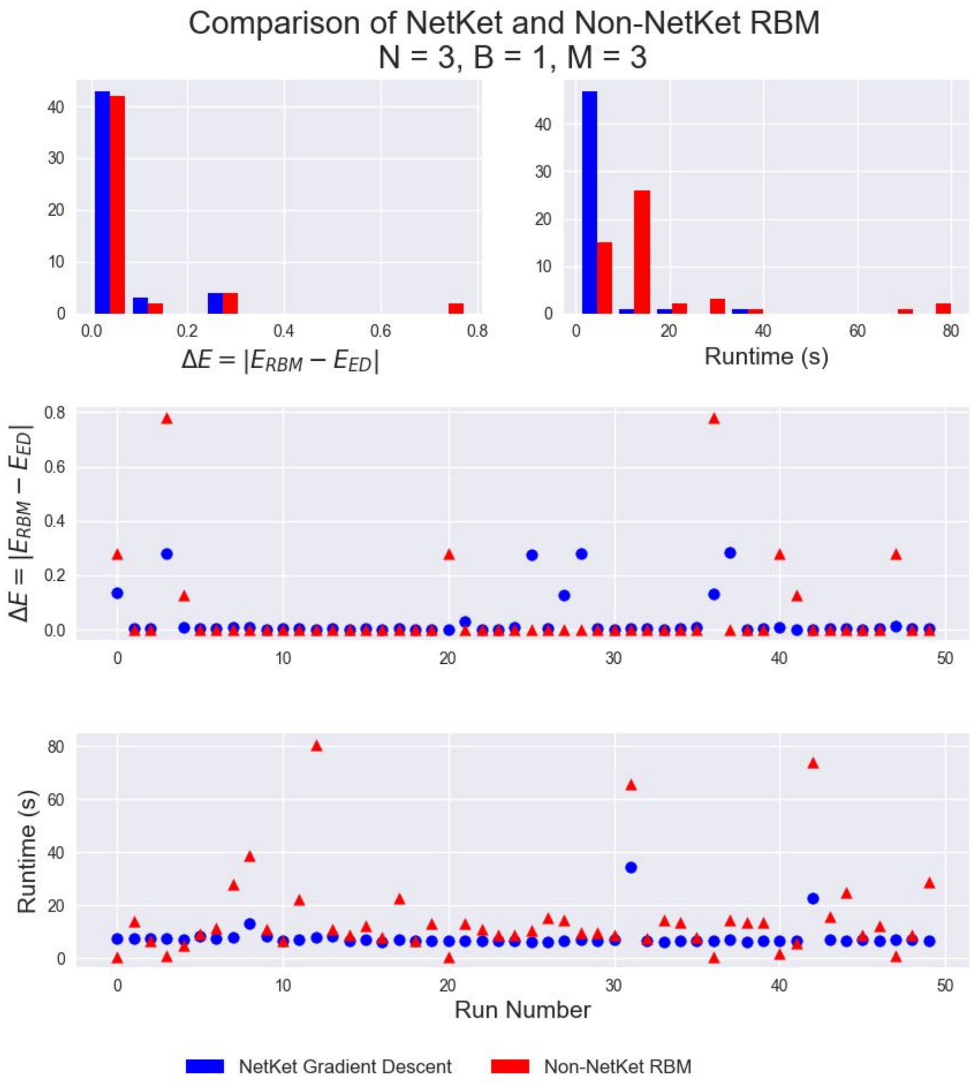


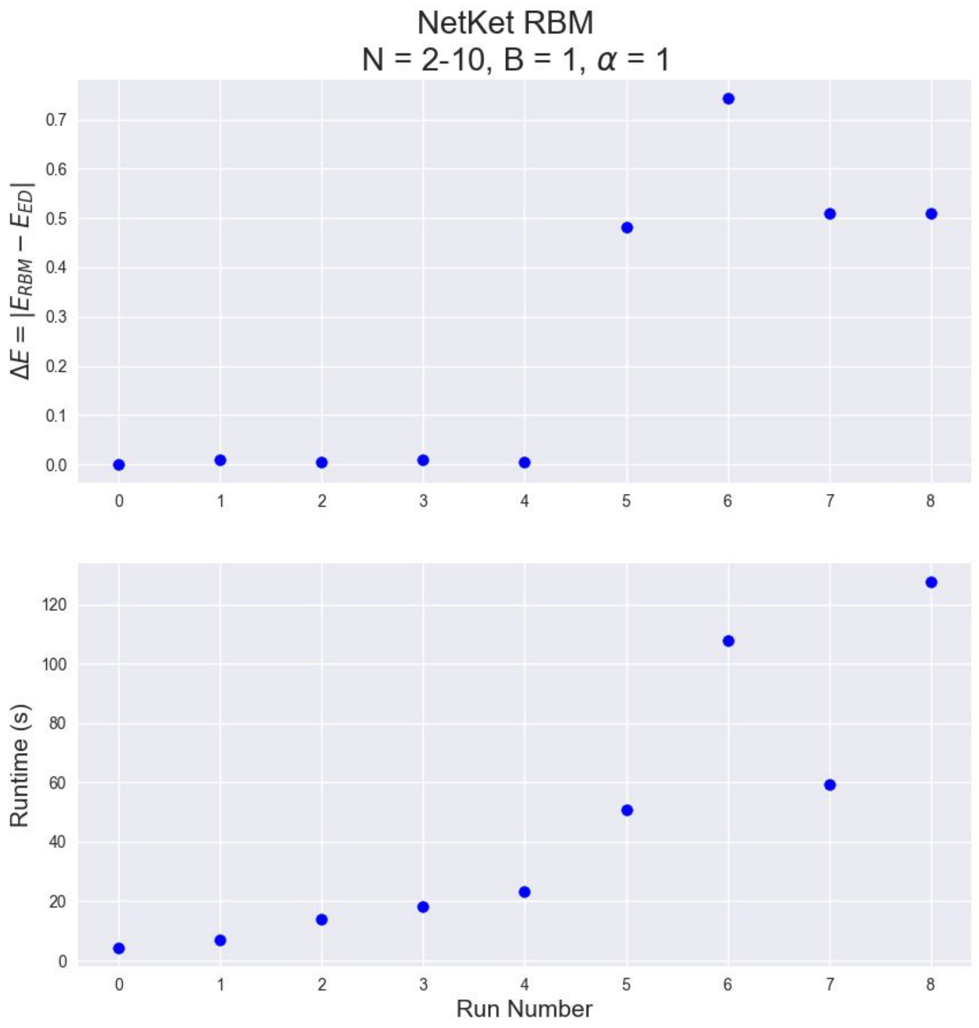


N = 2, Possible configurations: $\uparrow\uparrow, \uparrow\downarrow, \downarrow\uparrow, \downarrow\downarrow$

NetKet Comparison

July 2020





Further Goals

- Fully benchmark the advantages of the stochastic sampling (NetKet)
- Understand why there are high errors in large systems
- Test out other minimization techniques supported by NetKet (stochastic reconfiguration)
- Try a varying values of A_k
- Dynamics