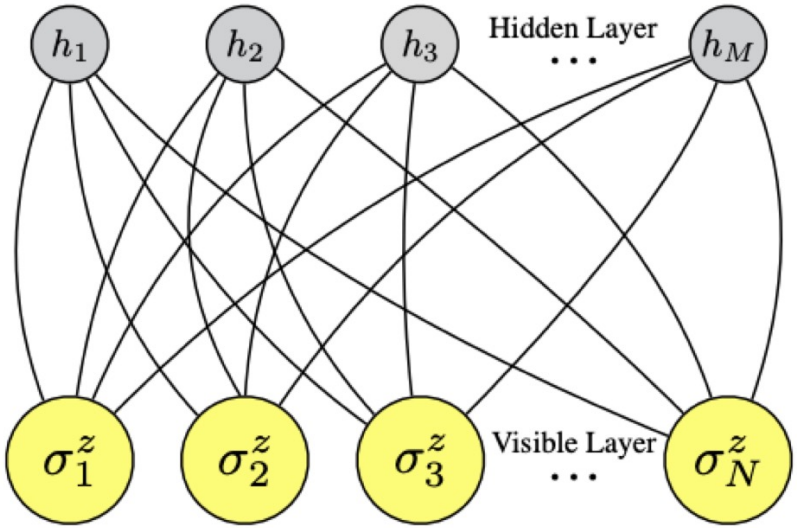
THE QUANTUM MANY-BODY PROBLEM IN THE AGE OF MACHINE LEARNING

Adrian Feiguin

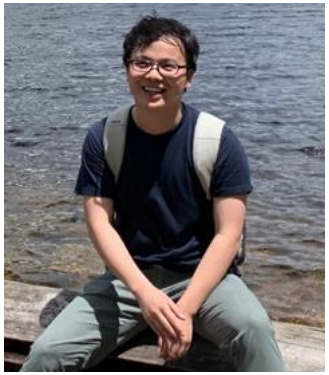
Northeastern University (Boston)



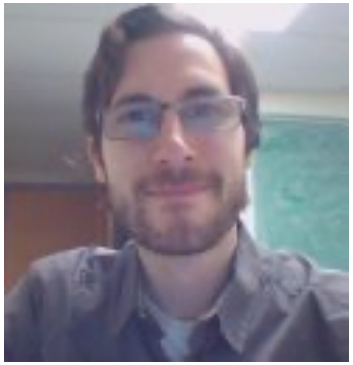
THE QUANTUM MANY-BODY PROBLEM IN THE AGE OF MACHINE LEARNING

Collaborators:

Doug Hendry



Hongwei Chen
(NU and SLAC)



Phillip Weinberg
(now at QuERA)

References:

1. "Chebyshev expansion of spectral functions using restricted Boltzmann machines"; D. Hendry, Hongwei Chen, Phillip Weinberg, A. E. Feiguin ; Phys. Rev. B **104**, 205130 (2021)
2. "A machine learning approach to dynamical properties of quantum many-body systems"; Douglas Hendry, Adrian E. Feiguin; Phys. Rev. B **100**, 245123 (2019).
3. "Systematic improvement of neural network quantum states using a Lanczos recursion"; Hongwei Chen, Douglas Hendry, Phillip Weinberg, Adrian E. Feiguin; NeurIPS 2022 (Accepted). arXiv: 2206.14307
4. "Neural network representation for minimally entangled typical thermal states"; Douglas Hendry, Hongwei Chen, Adrian Feiguin. Phys. Rev. B **106**, 165111 (2022).

The representation problem

$$|\psi\rangle = \sum_{\{s\}} \psi(s_1, s_2, s_3 \dots s_L)|s_1, s_2, s_3 \dots s_L\rangle$$

The # of configurations/coefficients $\psi(s_1, s_2, s_3 \dots s_L)$ grows exponentially and cannot be stored.

Variational solution:

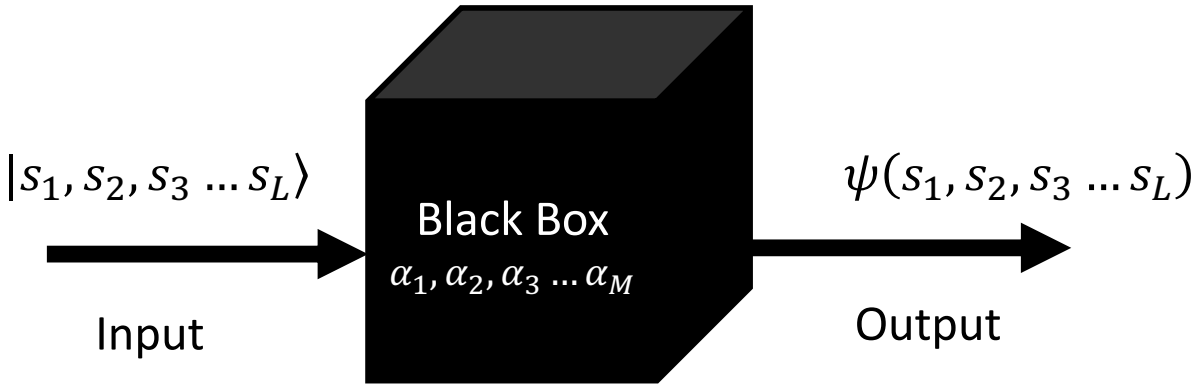
Propose a functional form/analytical expression for the coefficients in terms of a reduced set of *M* free parameters.

$$|\psi_{\alpha_1,\alpha_2,\alpha_3\dots\alpha_M}\rangle = \sum_{\{s\}} \psi_{\alpha_1,\alpha_2,\alpha_3\dots\alpha_M}(s_1, s_2, s_3 \dots s_L)|s_1, s_2, s_3 \dots s_L\rangle$$

More concretely:

$$|\psi(\alpha)\rangle = \sum_{s} \psi_s(\alpha)|s\rangle; \alpha \in \mathbb{C}^M \quad \text{with} \quad \psi_s(\alpha) = f(x(s); \alpha)$$

What's a variational wave function?

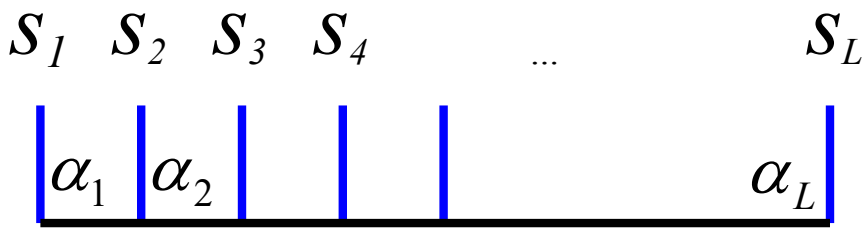


Example: Matrix product states (MPS)

$$|\psi\rangle = \sum_{\{s\}} M[s_1]_{\alpha_1} M[s_2]_{\alpha_1,\alpha_2} \ldots M[s_2]_{\alpha_{L-1},\alpha_L} M[s_L]_{\alpha_L} \, | s_1 \ldots s_L \rangle$$

$$= \sum_{\{s\}} M[s_1] M[s_2] \ldots M[s_L] \, | s_1 \ldots s_L \rangle$$

$$= \sum_{\{s\}} \prod_{l=1}^{L} M[s_l] \, | s_1 \ldots s_L \rangle$$



Quantum spin systems

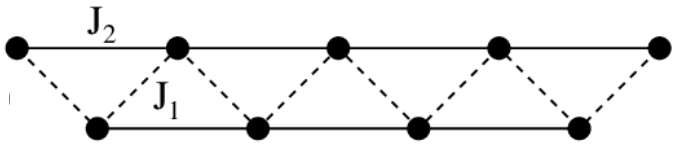
Each site is occupied by a bound electron with two degrees of freedom (spin up and spin down,↓).

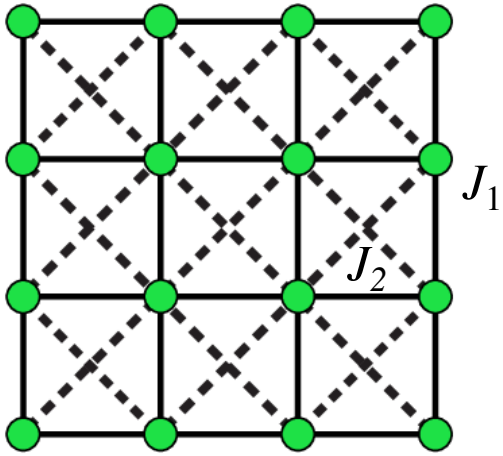
A typical spin configuration is represented as $|s_1, s_2, s_3 \ldots s_L\rangle = |\uparrow\uparrow\downarrow\cdots\uparrow\rangle$

We have 2^L of such configurations (less if considering symmetries)

We sill consider the problem in the context of the Heisenberg model

$$\hat{H} = J_1 \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j + J_2 \sum_{\langle\langle ij \rangle\rangle} \vec{S}_i \cdot \vec{S}_j$$





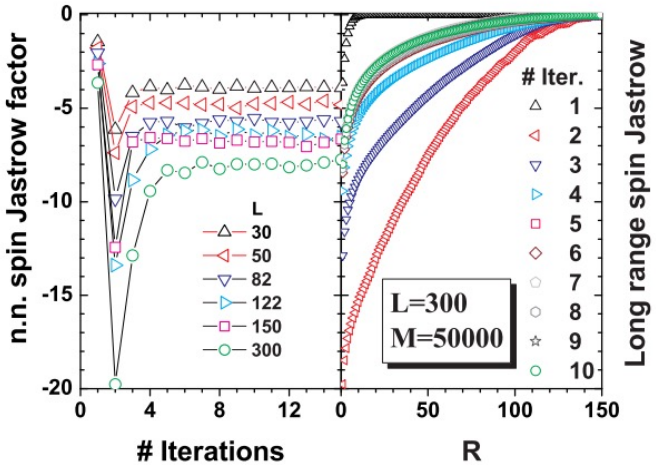
Example: Jastrow wave-function

$$|\psi_{\alpha_1,\alpha_2,\alpha_3\ldots\alpha_M}\rangle = \sum_{\{s\}} \psi_{\alpha_1,\alpha_2,\alpha_3\ldots\alpha_M}(s_1, s_2, s_3 \ldots s_L)|s_1, s_2, s_3 \ldots s_L\rangle$$

$$\psi_\alpha(s_1, s_2, s_3 \ldots s_L) = \langle s_1, s_2, s_3 \ldots s_L|e^{\sum_{ij}\alpha_{ij}s_i^z s_j^z} \quad |s_1, s_2, s_3 \ldots s_L\rangle = e^{\sum_{ij}\alpha_{ij}s_i s_j}$$

Its simple form is derived from spin-wave theory.
Breaks SU(2) symmetry, but gives the correct physics in the thermodynamic limit.
Good for ordered phases.
Further simplification, in terms of a single parameter

$$\alpha_{ij} = \frac{\alpha}{|i-j|}$$



L variational parameters with translational symmetry

S. Sorella, PRL 2005

Horsch and von der Linden (1988), Huse and Elser (1988), Liu and Manousakis (1989), and Manousakis (1989) (See Manousakis RMP (1991))

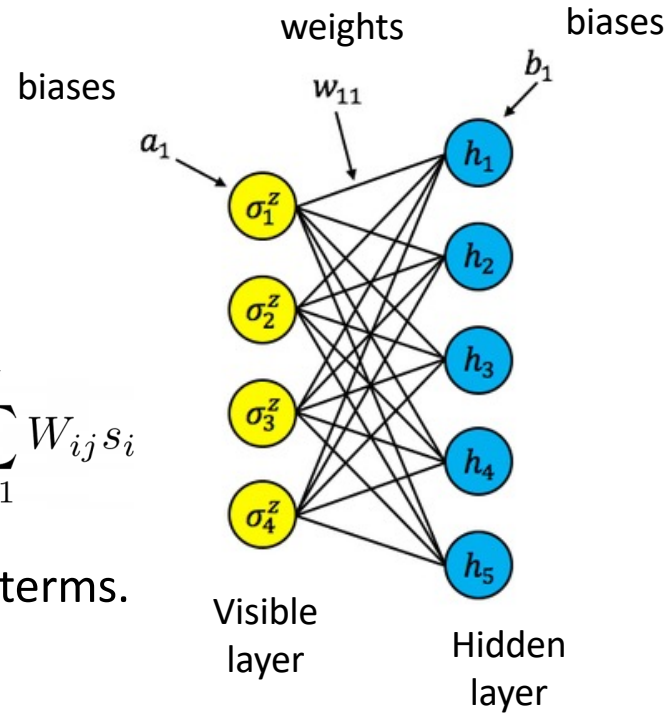
Restricted Boltzmann Machines

- In machine learning they are used to represent probability densities of binary vectors (spins). A joint probability between the real/visible spins and latent hidden spins is given by a Boltzmann distribution with interaction between the visible and hidden.
- For VWF , the weights are made complex allowing all possible phases.
- Representational power for VWF is surprisingly good for their relative simplicity.

$$\psi_s(a,b,W) = \sum_{h\in\{-1,+1\}^M} e^{-E(s,h)}$$

$$-E(s,h) = \sum_{i=1}^{N} a_i s_i + \sum_{i=1}^{M} b_i h_i + \sum_{i=1}^{N}\sum_{j=1}^{M} W_{ij} s_i h_j$$

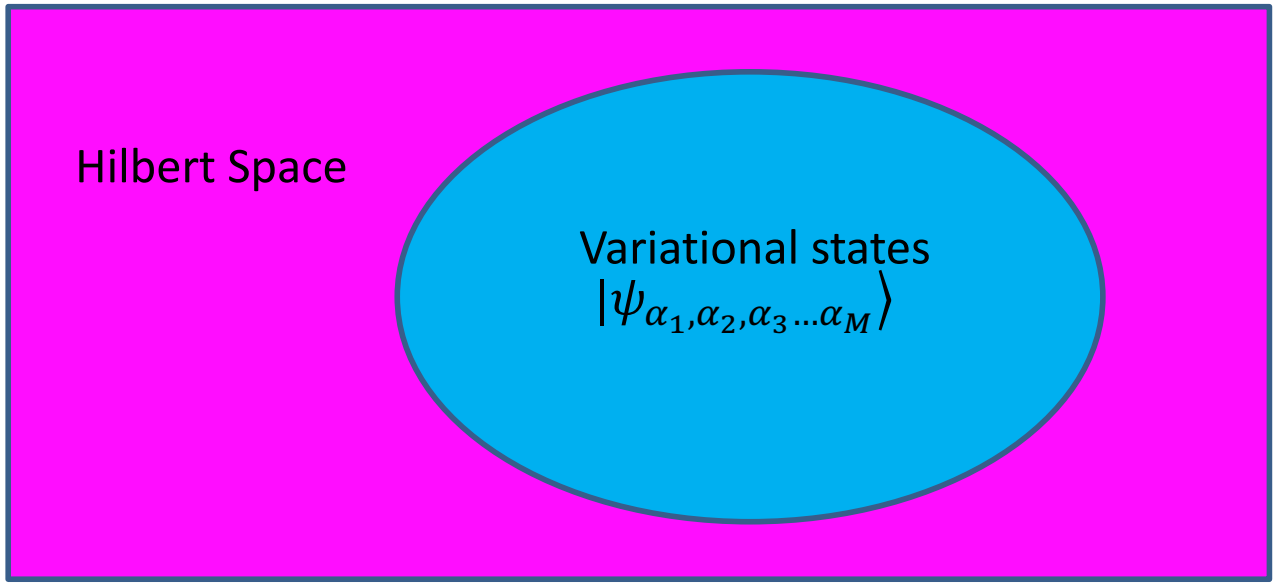
$$\psi_s(a,b,W) = e^{\sum_{i=1}^{N} a_i s_i} \prod_{j=1}^{M} 2\cosh(\theta_j); \quad \theta_j = b_j + \sum_{i=1}^{N} W_{ij} s_i$$



It goes beyond Jastrow by introducing many-body terms.

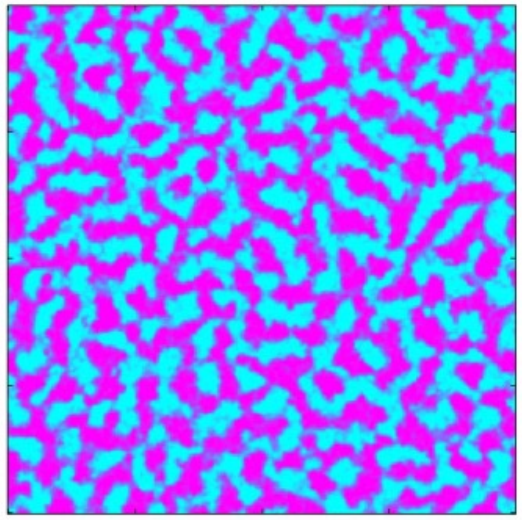
G. Carleo, M. Troyer, Science (2017)

Representability



In reality, the picture is more like this…

By adding more variational parameters or changing the "geometry" of the wave function, we can improve its "expressivity"



Variational solution

The variational coefficients are obtained by minimizing a loss function, in this case, the variational energy:

$$E_\alpha = \frac{\langle \psi_\alpha | H | \psi_\alpha \rangle}{\langle \psi_\alpha | \psi_\alpha \rangle}$$

It can be easily re-expressed as

$$E_\alpha = \sum_{\{s\}} P_s \, E_s \qquad P_s = \frac{|\langle s | \psi_\alpha \rangle|^2}{\sum_{\{s\}} |\langle s | \psi_\alpha \rangle|^2}$$

It clearly looks like a Boltzmann distribution, where the Boltzmann weight is replaced by the wave-function coefficient squared.
Still, sampling over all the configurations is a formidable task, however…we know the analytical expression of these coefficients and we can carry out the equivalent to “classical Monte Carlo” : Variational MC

Sampling with Variational Monte Carlo

Variational energy $$E_\alpha = \frac{\langle \psi_\alpha | \hat{H} | \psi_\alpha \rangle}{\langle \psi_\alpha | \psi_\alpha \rangle} \geq E_{min}$$ Where $\langle \dots \rangle$ means statistical average.

The expectation value of an arbitrary operator O is

$$\begin{aligned}
\langle O \rangle_{Var} &= \frac{\langle \psi_T \mid O \mid \psi_T \rangle}{\langle \psi_T \mid \psi_T \rangle} \\
&= \frac{\sum_x \langle \psi_T \mid x \rangle \langle x \mid O \mid \psi_T \rangle}{\sum_x \langle \psi_T \mid x \rangle \langle x \mid \psi_T \rangle} \\
&= \frac{\sum_x |\langle \psi_T \mid x \rangle|^2 \sum_{x'} \langle x \mid O \mid \psi_T \rangle / \langle x \mid \psi_T \rangle}{\sum_x |\langle \psi_T \mid x \rangle|^2} \\
&= \sum_x P_x O_x,
\end{aligned}$$

with

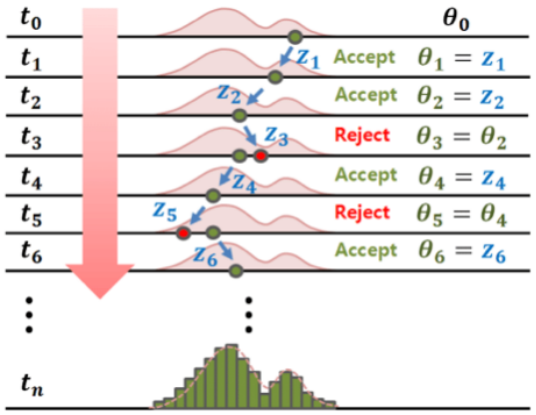
$$P_x = \frac{|\langle \psi_T \mid x \rangle|^2}{\sum_x |\langle \psi_T \mid x \rangle|^2},$$

and

$$O_x = \frac{\langle x \mid O \mid \psi_T \rangle}{\langle x \mid \psi_T \rangle}.$$

Algorithms:

Markov Chain Monte Carlo
Importance sampling
Metropolis algorithm



Variational optimization (I)

Gradient descent $\alpha_k = \alpha_k - \eta \Delta\alpha_k$

$$\Delta\alpha_k = \frac{\partial\left(\frac{\langle\psi_\alpha|\hat{H}|\psi_\alpha\rangle}{\langle\psi_\alpha|\psi_\alpha\rangle}\right)}{\partial\alpha_k^*}$$

$$= \frac{(\partial_{\alpha_k^*}\langle\psi_\alpha|)\hat{H}|\psi_\alpha\rangle}{\langle\psi_\alpha|\psi_\alpha\rangle} - \frac{[(\partial_{\alpha_k^*}\langle\psi_\alpha|)|\psi_\alpha\rangle]\langle\psi_\alpha|\hat{H}|\psi_\alpha\rangle}{\langle\psi_\alpha|\psi_\alpha\rangle^2}$$

Define: $E_{loc} = \frac{\langle s|H|\psi_\alpha\rangle}{\langle\psi_\alpha|\psi_\alpha\rangle}$ $\mathcal{O}_k(s) = \frac{1}{\psi_s}\frac{\partial\psi_s}{\partial\alpha_k}$

$$\Delta\alpha_k = \langle E_{loc}\mathcal{O}_k(s)\rangle - \langle E_{loc}\rangle\langle\mathcal{O}_k(s)\rangle$$

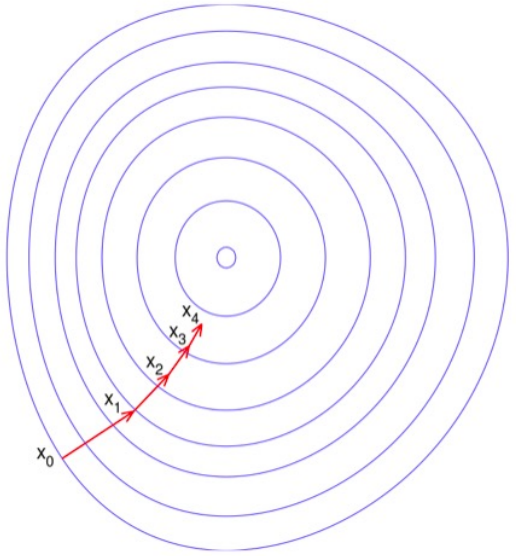
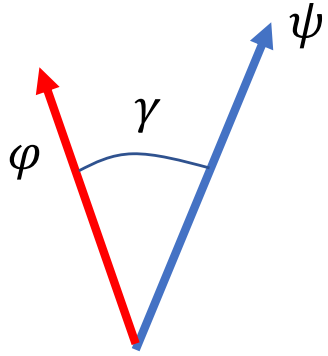


Illustration of gradient descent

Variational optimization (II)

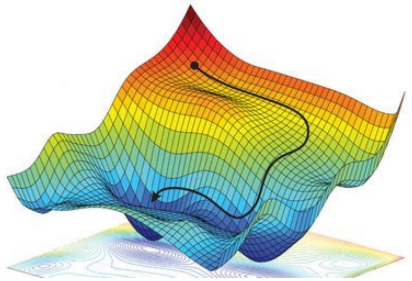
"Stochastic reconfiguration" (or "Natural gradient descent"): minimize the distance between quantum states following the exact and variational imaginary-time revolution



Optimal solution

$$\delta\widetilde{\theta} = \arg\min_{\delta\theta} \mathcal{F}[e^{-\delta_\tau \mathcal{H}} |\Psi_\theta\rangle, |\Psi_{\theta+\delta\theta}\rangle],$$

$$\mathcal{F}[|\psi\rangle, |\phi\rangle] := \arccos \sqrt{\frac{\langle\psi||\phi\rangle\langle\phi||\psi\rangle}{\langle\psi||\psi\rangle\langle\phi||\phi\rangle}}.$$

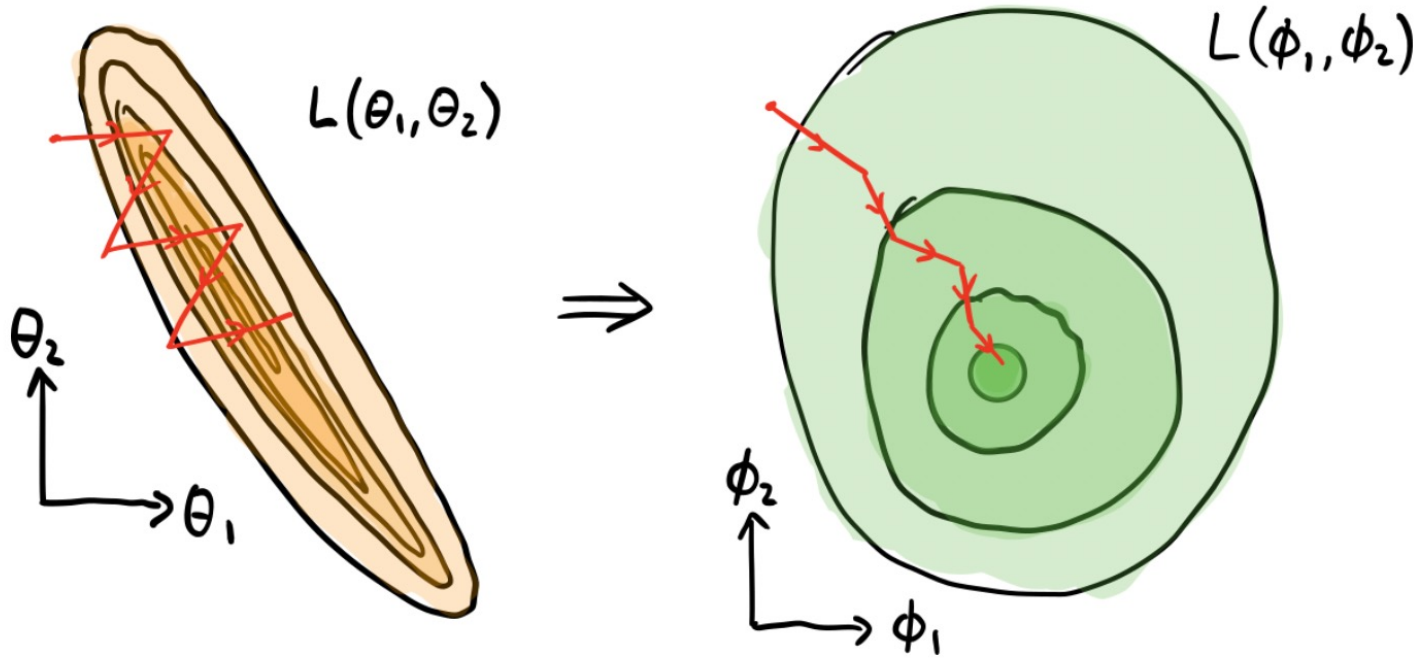


$$\delta\widetilde{\theta}_k = -\delta_\tau \sum_l S_{kl}^{-1} f_l.$$

$$S_{kl} = \frac{\langle\partial_k\psi|\partial_l\psi\rangle}{\langle\psi|\psi\rangle} - \frac{\langle\partial_k\psi|\psi\rangle}{\langle\psi|\psi\rangle}\frac{\langle\psi|\partial_l\psi\rangle}{\langle\psi|\psi\rangle} = \langle O_k(s)^* O_l(s)\rangle - \langle O_k(s)\rangle^* \langle O_l(s)\rangle \qquad O_k(s) = \frac{1}{\psi}\frac{\partial\psi}{\partial\alpha_k}$$

$$f_k = \frac{\langle\partial_k\Psi_\theta|\mathcal{H}|\Psi_\theta\rangle}{\langle\Psi_\theta||\Psi_\theta\rangle} - \frac{\langle\partial_k\Psi_\theta||\Psi_\theta\rangle}{\langle\Psi_\theta||\Psi_\theta\rangle}\frac{\langle\Psi_\theta|\mathcal{H}|\Psi_\theta\rangle}{\langle\Psi_\theta||\Psi_\theta\rangle}$$
$$= \langle O_k^\dagger \mathcal{H}\rangle - \langle O_k^\dagger\rangle\langle\mathcal{H}\rangle,$$

Natural Gradient descent vs gradient descent

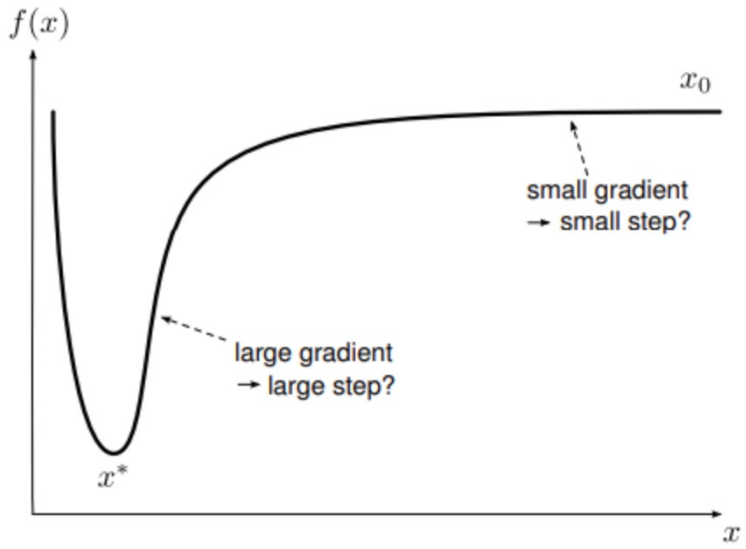


Euclidian metric

"Natural" metric

https://medium.com/xanaduai/optimizing-quantum-computations-with-the-quantum-natural-gradient-ba0636ebdb86

Natural Gradient descent vs gradient descent



Fisher information matrix

$$x = x - \alpha \frac{\partial f(x)}{\partial x} \quad \Rightarrow \quad x_{t+1} = x_t - \eta_t \cdot F(\theta_t)^{-1} \cdot \frac{\partial f(x)}{\partial x}$$

1.α, the learning rate, is replaced with η_t to make it clear that the step size may change in each iteration

2.An additional term $F(\theta_t)_{-1}$ has been added to the normal gradient.

https://towardsdatascience.com/natural-gradient-ce454b3dcdfa

Natural gradient descent as imaginary time evolution

The main idea to obtain an approximation to the ground-state consists of evolving the wave-function in imaginary time:

$$\frac{d}{d\tau}|\psi(\tau)\rangle = -(\hat{H} - E(\tau))\,|\psi(\tau)\rangle$$

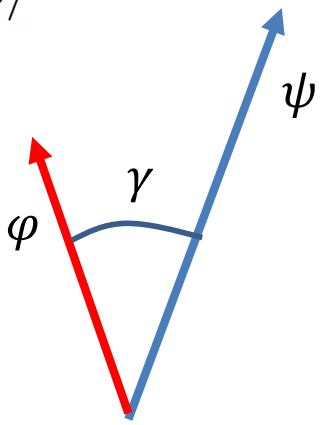
We need to minimize the following projection error:

$$\left\| \tfrac{d}{d\tau}|\psi(\tau)\rangle - (E - \hat{H})\,|\psi(\tau)\rangle \right\|$$

In general, when we have a variational wave-function ψ and a target wave function φ, we use the loss function

$$L(\psi) = \gamma(\psi,\phi)^2 \qquad \gamma(\psi,\phi) = \arccos\sqrt{\frac{\langle\psi|\phi\rangle\langle\phi|\psi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}}$$

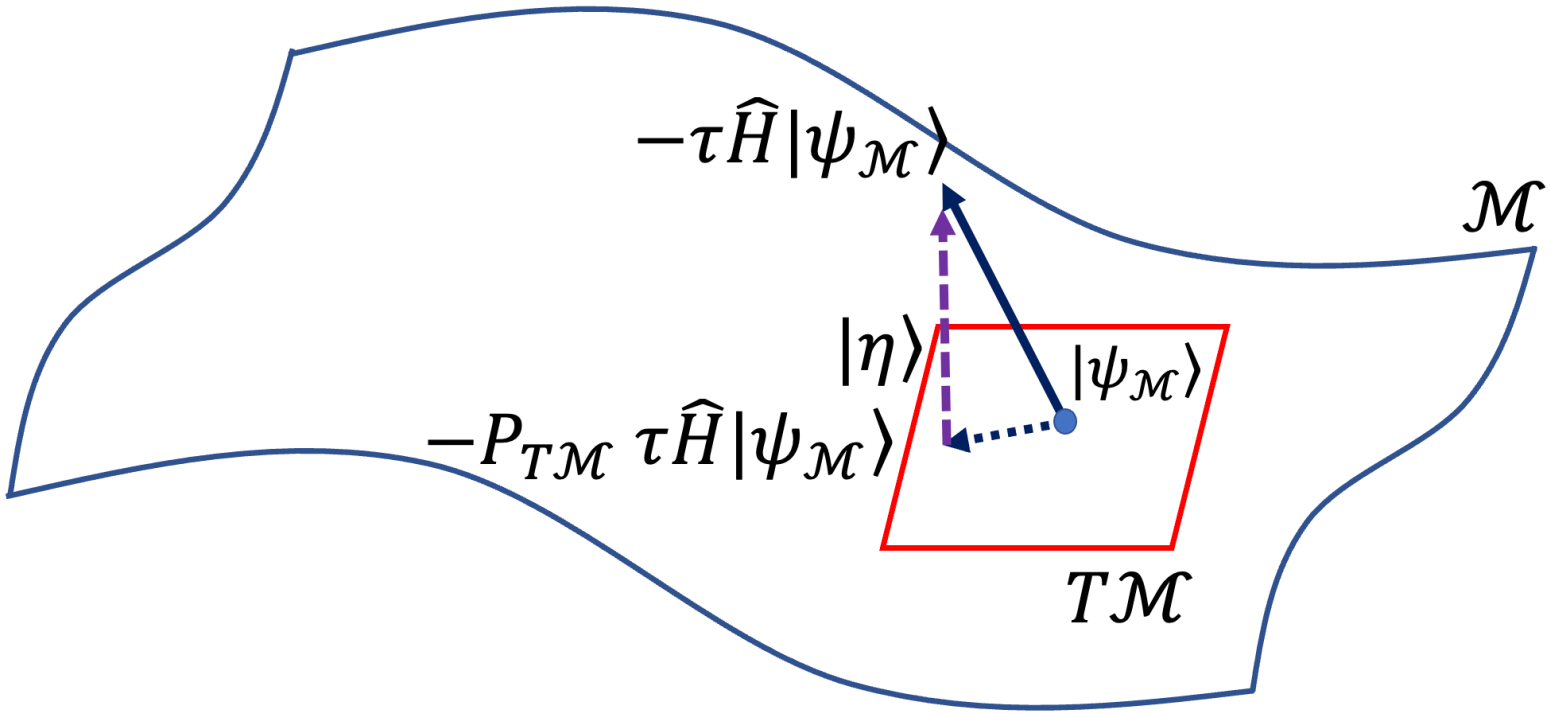
Where γ is the so-called Fubini-Study metric.
This minimization leads to a set of equations to update the variational parameters α in terms of the derivatives $\frac{d}{d\alpha_k}|\psi_\alpha\rangle$
(Notice lack of normalization)



Gradient descent: Euclidian geometry of the parameter space
Natural GD: "information geometry" of the variational manifold

For a review: Lucas Hackl, et al, SciPost (2020). J. Martens cs/1412.1193. See also Sorella's "Stochastic reconfiguration".

Geometric interpretation NGD as a time-evolution process



Restoring symmetry

Original RBM wave function

$$\psi(\overrightarrow{\sigma^z}, \vec{a}, \vec{b}, W) = e^{\sum_{i=1}^{N} a_i \sigma_i^z} \prod_{i=1}^{M} 2\cosh\left(\sum_{j=1}^{N} W_{ij}\sigma_j^z + b_i\right)$$

Spin flip symmetry:

$$\psi_s(\overrightarrow{\sigma^z}, W) = \prod_{i=1}^{M} 2\cosh\left(\sum_{j=1}^{N} W_{ij}\sigma_j^z\right). \qquad \sigma_i^z \rightarrow -\sigma_i^z$$

Translational symmetry:

$$\psi_{\mathbf{K}} = \sum_{\mathbf{R}} e^{-i\mathbf{K}\cdot\mathbf{R}} \psi_s(T_{\mathbf{R}}\overrightarrow{\sigma^z}, W)$$

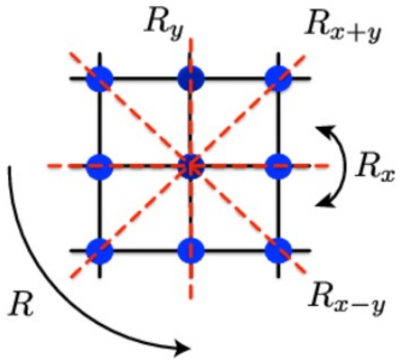
The translation operator T_R shift all particles by a distance $\mathbf{R} = m\hat{\mathbf{x}} + n\hat{\mathbf{y}}$

Lattice point symmetry:

$$\psi_{\mathbf{K}\mathscr{L}} = \sum_{\mathbf{R},\mathscr{L}} e^{-i\mathbf{K}\cdot\mathbf{R}} \chi(\mathscr{L}) \psi_s(T_{\mathbf{R}}\mathscr{L}\overrightarrow{\sigma^z}, W)$$

Table 1: Character table of the C_{4v} point group for square lattice.

	E	$2C_4$	C_2	$2\sigma_v$	$2\sigma_d$
A_1	1	1	1	1	1
A_2	1	1	1	-1	-1
B_1	1	-1	1	1	-1
B_2	1	-1	1	-1	1
E	2	0	-2	0	0



Lanczos recursion

With initial trial wave function ψ_0, define a new vector by applying Hamiltonian

$$\psi_1 = \frac{H\psi_0 - \langle H\rangle\psi_0}{(\langle H^2\rangle - \langle H\rangle^2)^{1/2}}$$

Using ψ_0 and ψ_1 as basis, the Hamiltonian can be written as a 2x2 matrix

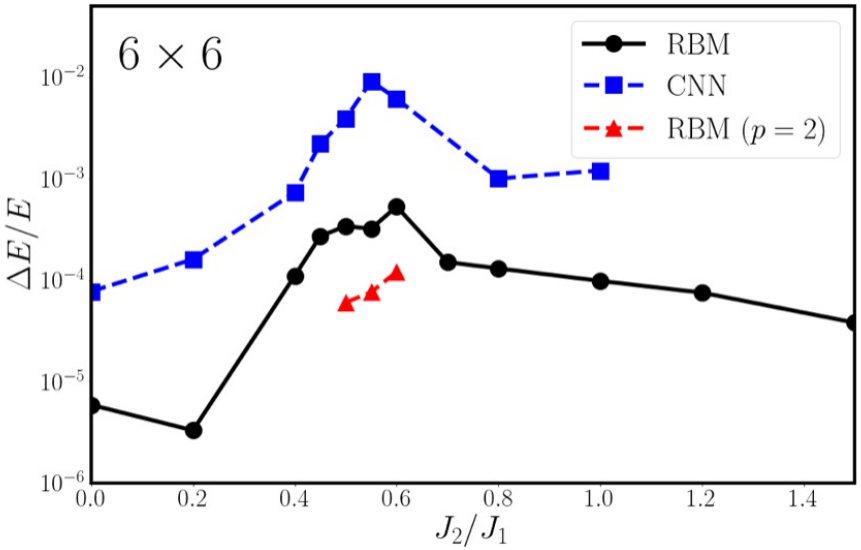
$$\tilde{E}_0 = \langle H\rangle + v\alpha$$

$$\tilde{\psi}_0 = \frac{1}{(1+\alpha^2)^{1/2}}\psi_0 + \frac{\alpha}{(1+\alpha^2)^{1/2}}\psi_1$$

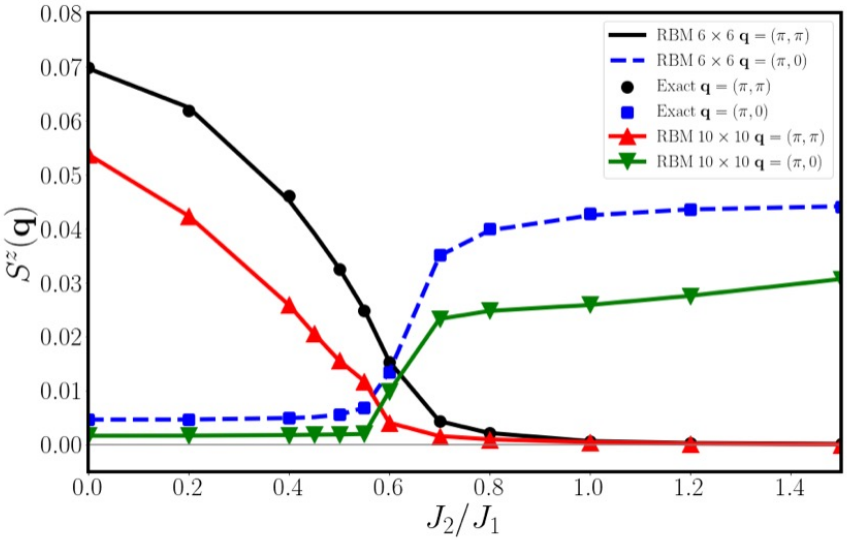
Where $\quad v = (\langle H^2\rangle - \langle H\rangle^2)^{1/2} \qquad \alpha = r - (r^2+1)^{1/2}$

$$r = \frac{\langle H^3\rangle - 3\langle H^2\rangle\langle H\rangle + 2\langle H\rangle^3}{2(\langle H^2\rangle - \langle H\rangle^2)^{3/2}}$$

Results for the ground state



Relative error in the ground state energy obtained with variational Monte Carlo using symmetrized RBM wave functions (this work) and convolutional neural network (CNN) wave functions.



Static spin structure factor for ordering wave vectors and obtained with VMC using symmetrized RBM wave functions, compared to numerically exact results on a 6x6 lattice as a function of j2/j1. We also include VMC results for 10x10.

Results (cont'd)

Spin structure factor: $S(\mathbf{q}) = \frac{1}{N^2} \sum_{i,j} \langle \sigma_i^z \sigma_j^z \rangle e^{i\mathbf{q}\cdot(\mathbf{r_i}-\mathbf{r_j})}$

6×6	$J_2 = 0.5$	$J_2 = 0.55$	$J_2 = 0.6$
Energy(Exact)	-0.503810	-0.495178	-0.493239
Energy(CNN)	-0.50185(1)	-0.49067(2)	-0.49023(1))
Energy(RBM+PP)	-0.503765(1)	-0.495075(1)	-
Energy(RBM)	-0.50364(2)	-0.49501(1)	-0.49298(5)
Energy(RBM) $p = 1$	-0.50376(3)	-0.49512(4)	-0.49313(5)
Energy(RBM) $p = 2$	-0.50378(4)	-0.49514(4)	-0.49318(5)
S(π, π)(Exact)	1.16989	0.89452	0.5545
S(π, π)(RBM)	1.177(8)	0.902(6)	0.555(4)
S$(\pi, 0)$(Exact)	0.201907	0.2489	0.48412
S$(\pi, 0)$(RBM)	0.200(2)	0.246(2)	0.486(6)

Ground state energy per site and the spin structure factor obtained by our RBM wave function, CNN, RBM+PP, and exact diagonalization for the $J_1 - J_2$ model on 6x6 square lattice. **p** represents the number of Lanczos steps applied.

Conclusions

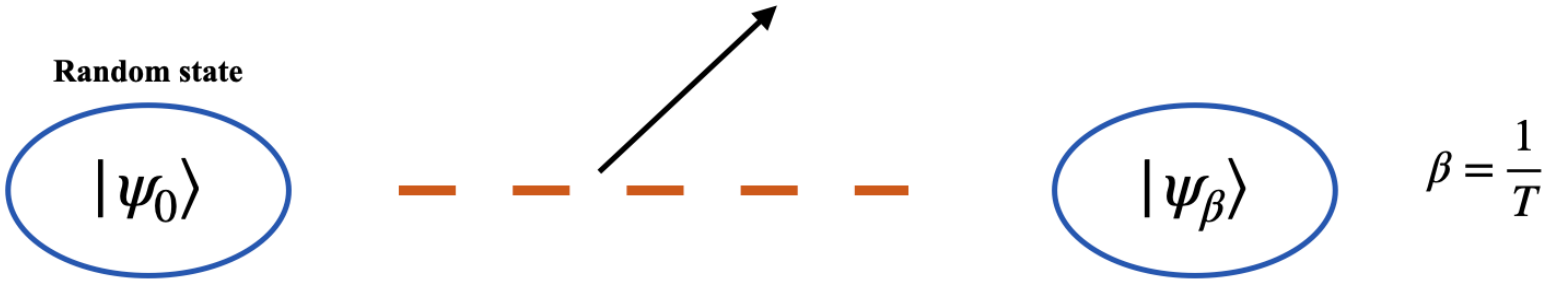
- RBM wave functions, initially deemed too simple, can be used as building blocks for complicated wave functions with much more representation power by enforcing the internal symmetries of the model and the point group symmetries of the lattice.
- The Lanczos method offers an effective solution to further improve the variational energy. With symmetry averaged RBM wave functions, it can produce the state-of-the-art accuracy for the ground state calculation of the model on square lattice.

Thermodynamics using VMC

Current approaches:

1. Minimally entangled typical thermal states (METTS)
2. Canonical thermal pure quantum states (CTPQS)

Evolve it with operator: $|\psi(\beta)\rangle = \exp(-\beta H/2)\,|\psi_0\rangle$



$$\langle \hat{A} \rangle_T = \frac{\langle \psi(\beta)\,|\,\hat{A}\,|\,\psi(\beta)\rangle}{\langle \psi(\beta)\,|\,\psi(\beta)\rangle}$$

Typical thermal states

A set of initial states: $\{|\phi_0(\xi)\}$ Each ξ is drawn from a probability distribution function $P_0(\xi)$

$$\int d\xi P_0(\xi)\,|\phi_0(\xi)\rangle\langle\phi_0(\xi)| = 1.$$

Imaginary time evolution

$$|\phi(\beta;\xi)\rangle = e^{-\frac{1}{2}\beta\hat{H}}\,|\phi_0(\xi)\rangle \qquad \beta = \frac{1}{T}$$

Partition function

$$\mathcal{Z}(\beta) = < Z(\beta,\xi) >_{P_0} \qquad Z(\beta;\xi) = \langle\phi(\beta;\xi)\,|\,\phi(\beta;\xi)\rangle$$

Any observable given by operator $\hat{A}$

$$\mathcal{A}(\beta) = \frac{< Z(\beta;\xi)A(\beta;\xi) >_{P_0}}{< Z(\beta;\xi) >_{P_0}} \qquad A(\beta;\xi) = \left.\frac{\langle\phi\,|\,\hat{A}\,|\,\phi\rangle}{\langle\phi\,|\,\phi\rangle}\right|_{(\beta;\xi)}$$

Importance weighted probability distribution function for ξ: $\quad P_\beta(\xi) = (Z(\beta,\xi)/\mathcal{Z}(\beta)) * P_0(\xi)$

Initialization

Random product states: Consider spin 1/2 on a lattice $\mathscr{L}$ with N sites

$$|\phi_0(\xi)\rangle = \otimes_{l\in\mathscr{L}} \left(\frac{\xi_{l\uparrow}|\uparrow\rangle_l + \xi_{l\downarrow}|\downarrow\rangle_l}{\sqrt{|\xi_{l\uparrow}|^2 + |\xi_{l\downarrow}|^2}} \right) \qquad P_0(\xi) = (\frac{1}{\pi})^{2N} \exp\left(-\sum_{l\in\mathscr{L}} \sum_{\sigma=\uparrow,\downarrow} |\xi_{l\sigma}|^2 \right)$$

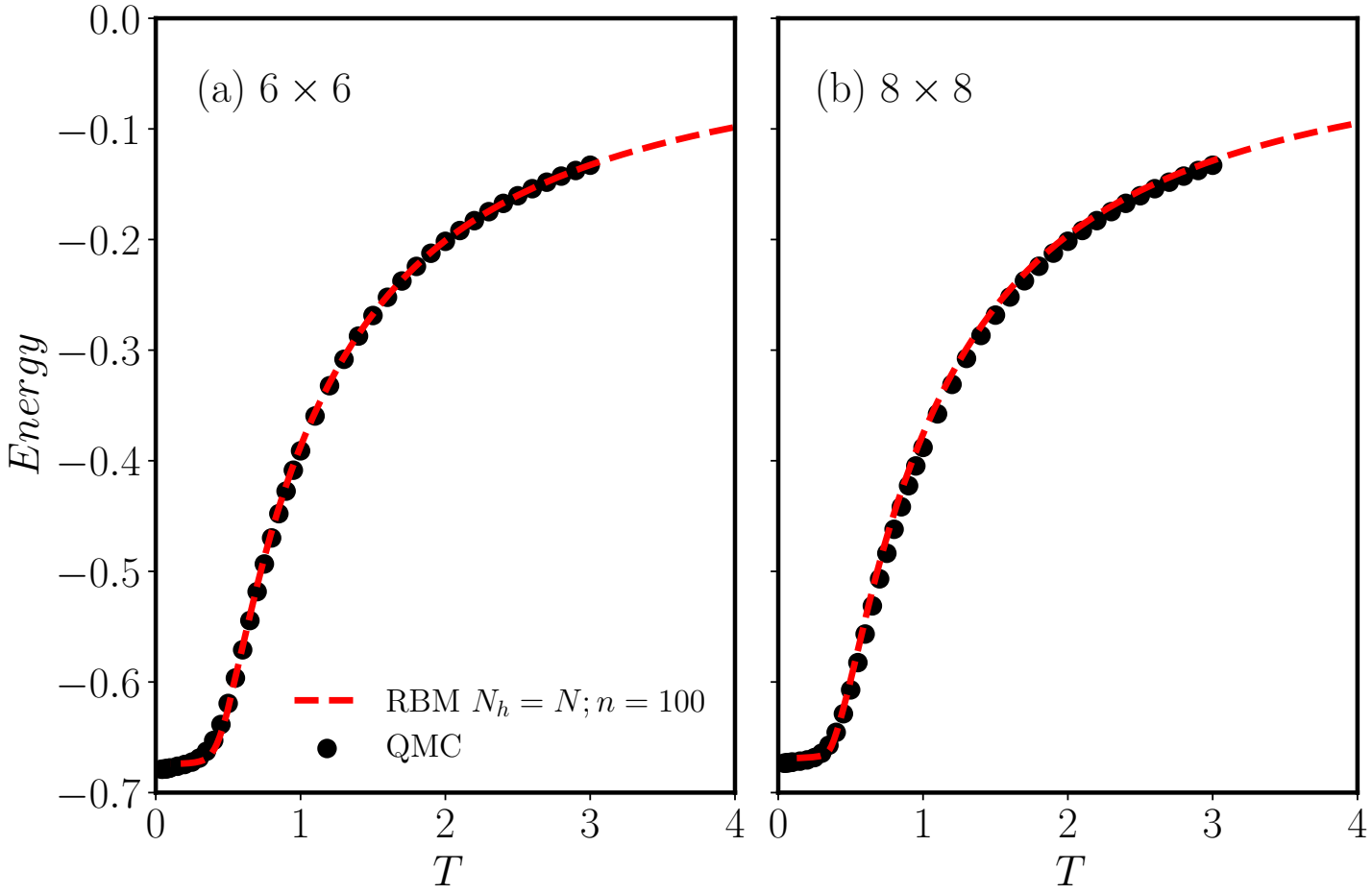
RBM initialization

$$\psi(\vec{s};\theta) = e^{\sum_{l\in\mathscr{L}} a_l s_l} \prod_{i=1}^{M} \cosh\left(b_i + \sum_{l\in\mathscr{L}} W_{il} \cdot s_l \right)$$

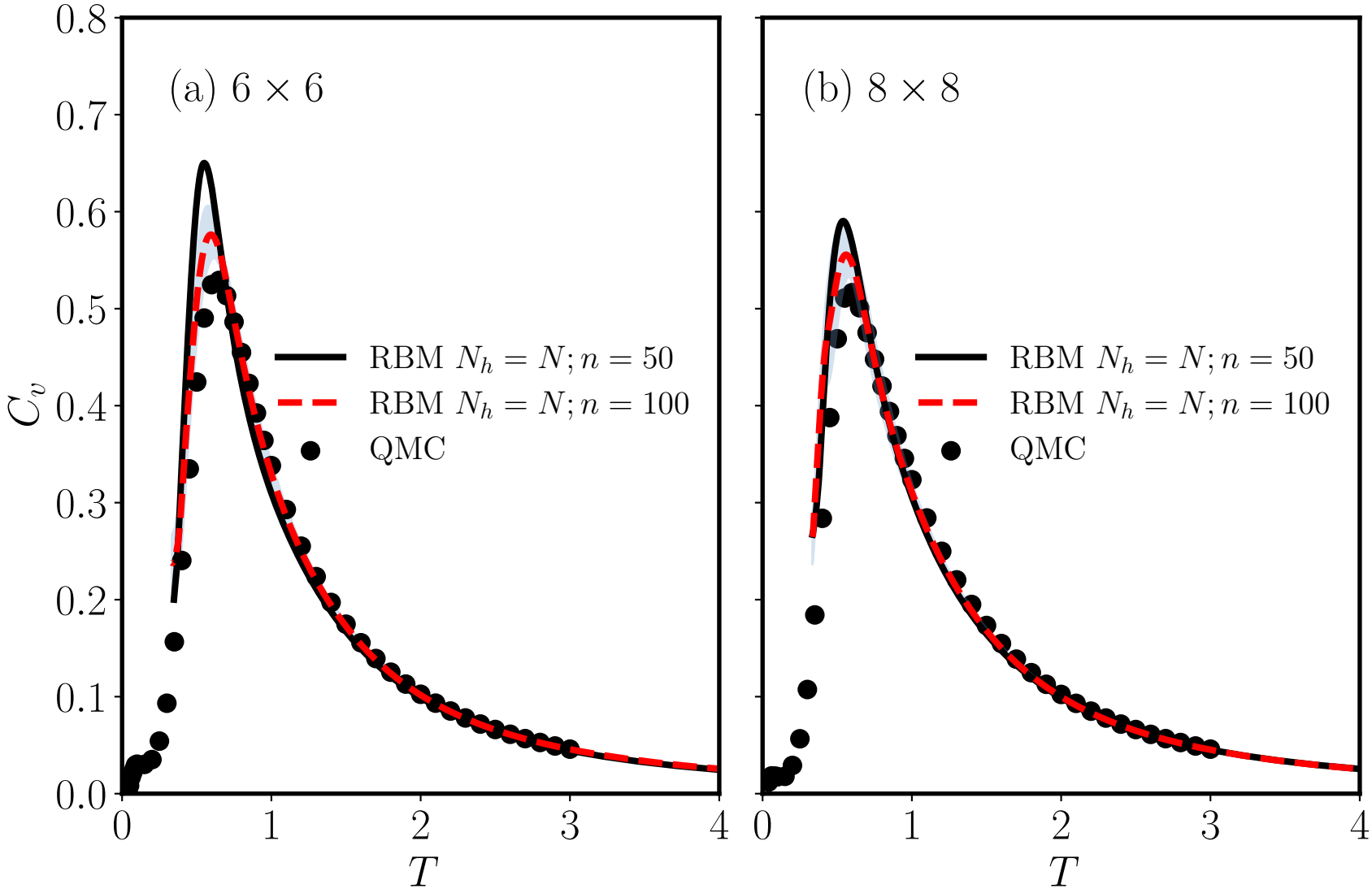
Let RBM exactly represent random product states

$$\textbf{Setting } W, b = 0 \qquad a_l = \frac{1}{2}\ln(\xi_{l+}/\xi_{l-})$$

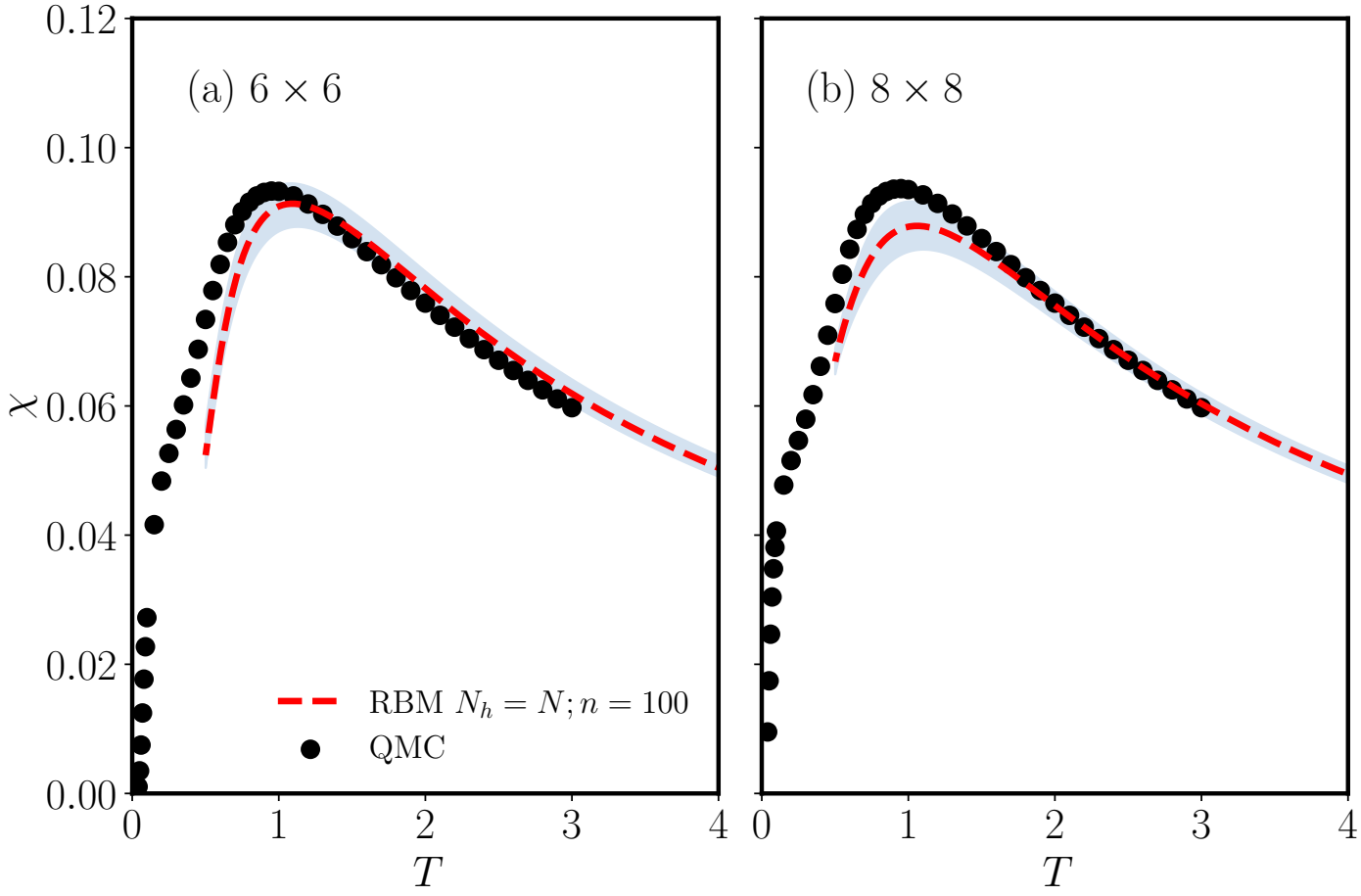
Results (Energy)



Results (Specific heat)



Results (Magnetic Susc.)



Conclusions

- We propose a new method to carry out thermodynamic simulations of quantum many-body models using a neural network representation of the wave functions.
- The method does not suffer from the sign problem and offers an alternative to matrix product states for studying two dimensional models.
- This method is general and can be extended to other variational forms.

Spectral functions with VMC

Calculate the spectral function for one and two dimensional quantum many body systems

Approximate zero temperature Green's function $G_{ij}(z) = \langle \psi | \hat{A}_j^\dagger \frac{1}{z - \hat{H}} \hat{A}_i | \psi \rangle$

Where $z = \omega + E_0 + i\eta$

Spectral function $A_{ij}(\omega) = -\frac{1}{\pi} G_{ij}(z)$

Hamiltonian $\hat{H} = J \sum_{\langle ij \rangle} \vec{S}_i \cdot \vec{S}_j$

Our approach: expanding the spectral function in terms of Chebyshev polynomials, and using restricted Boltzmann machines as variational wave function

Spectral functions: The correction vector

We first define the Green's function:

$$G_{ij}(z) = \langle\psi|A_i^\dagger \frac{1}{z-\hat{H}} A_j|\psi\rangle \quad \xrightarrow{\text{For spins}} \quad S^z(k,\omega) = -\frac{1}{L\pi}\mathrm{Im}\sum_n e^{ikn}\langle\psi|\hat{S}_0^z \frac{1}{z-\hat{H}}\hat{S}_n^z|\psi\rangle.$$

where A is some operator of interest and $z = E_0 + \omega + i\eta$

We introduce the two auxiliary states:

$$|A_i\rangle = \hat{A}_i|\psi\rangle$$
$$|\chi_j(z)\rangle = \frac{1}{z-\hat{H}}|A_j\rangle, \qquad \text{"Correction vector"}$$

$|\chi_j(z)\rangle$ can be obtained by solving the equation:

$$(z-\hat{H})|\chi_j(z)\rangle = \hat{A}_j|\psi\rangle = |A_j\rangle.$$

The spectral function is defined as the imaginary part of the Green's function, $A_{ij}(\omega) = -\frac{1}{\pi}\mathrm{Im}G_{ij}(z)$, or:

$$A_{ij}(\omega) = -\frac{1}{\pi}\mathrm{Im}\langle A_i|\chi_j(z)\rangle.$$

E. Jeckelmann, Phys. Rev. B 66, 045114 (2002).
T. D. Kühner and S. R. White, Phys. Rev. B 60, 335 (1999).

Variational optimization

The problem consists of solving the equation

$$(z - \hat{H})|\chi_j(z)\rangle = \hat{A}_j|\psi\rangle = |A_j\rangle.$$

Same as before, we use the Fubini-Study metric

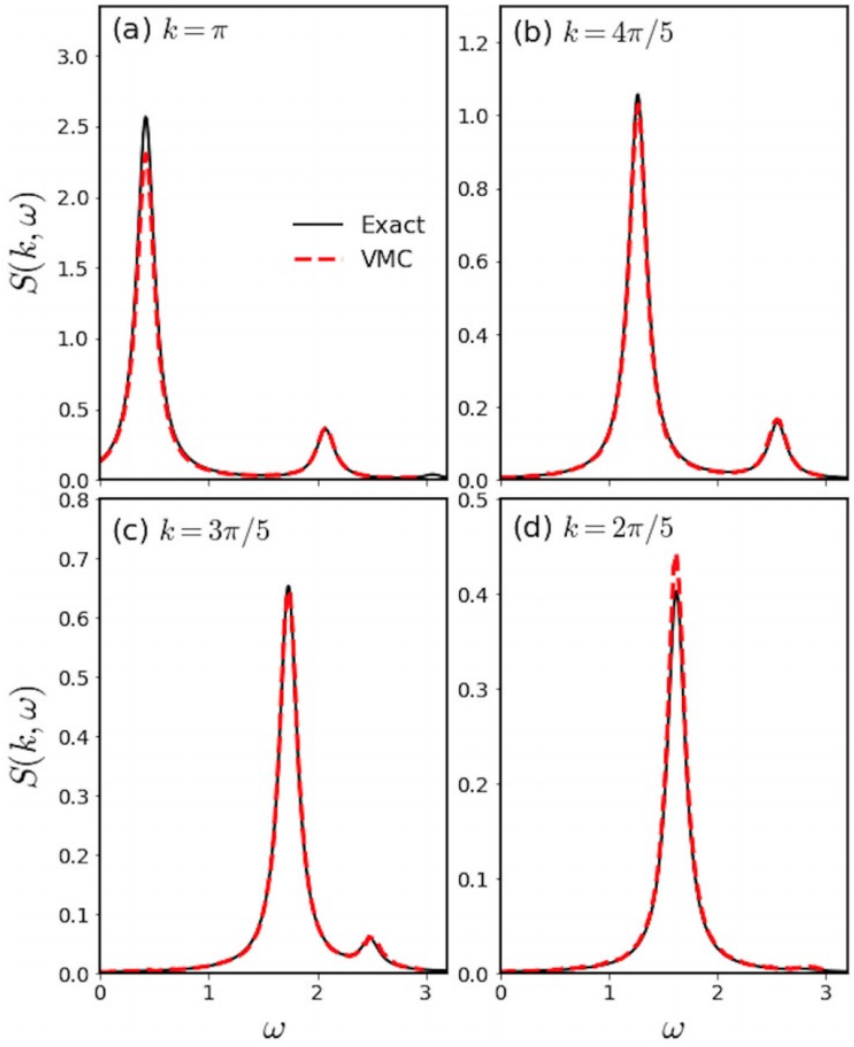
$$L(\psi) = \gamma(\psi, \phi)^2 \qquad \gamma(\psi, \phi) = \arccos \sqrt{\frac{\langle\psi|\phi\rangle\langle\phi|\psi\rangle}{\langle\psi|\psi\rangle\langle\phi|\phi\rangle}}$$

The target function is $|A\rangle$ and we need to solve for $|\chi\rangle$
So, we are solving a system of (complex) linear equations, where the solution is encoded in the form of an RBM.

Notice that the problem has to be solved for each value of z (or ω) -> hundreds of VMC runs in parallel.

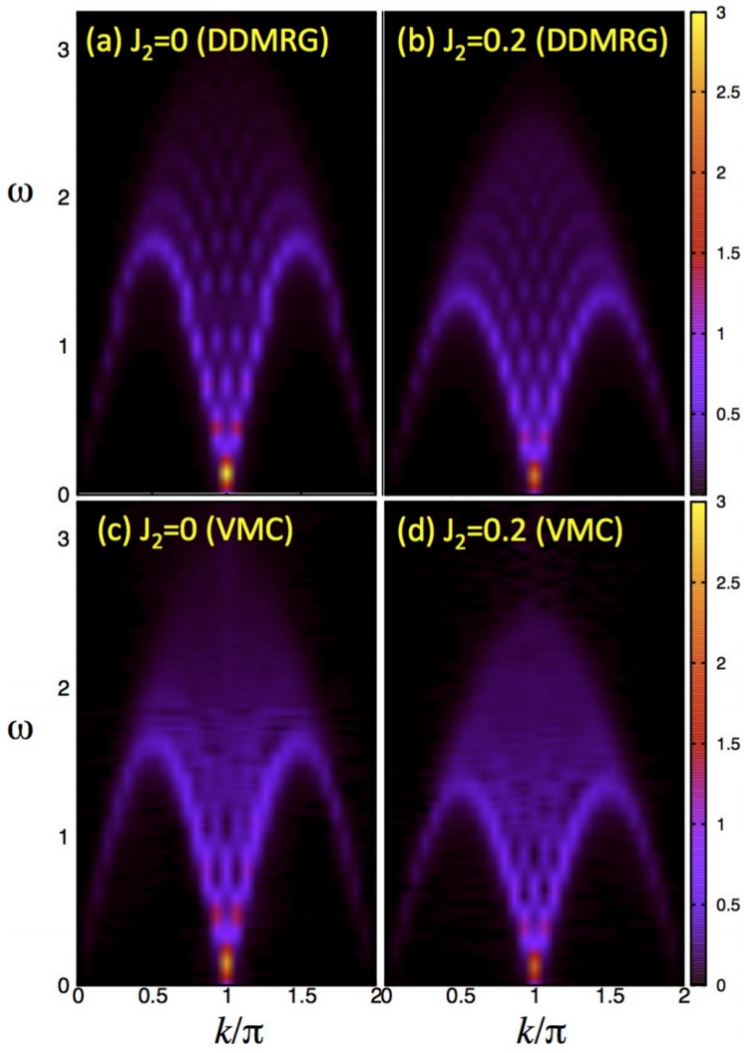
Results - Benchmark

Heisenberg chain L=10



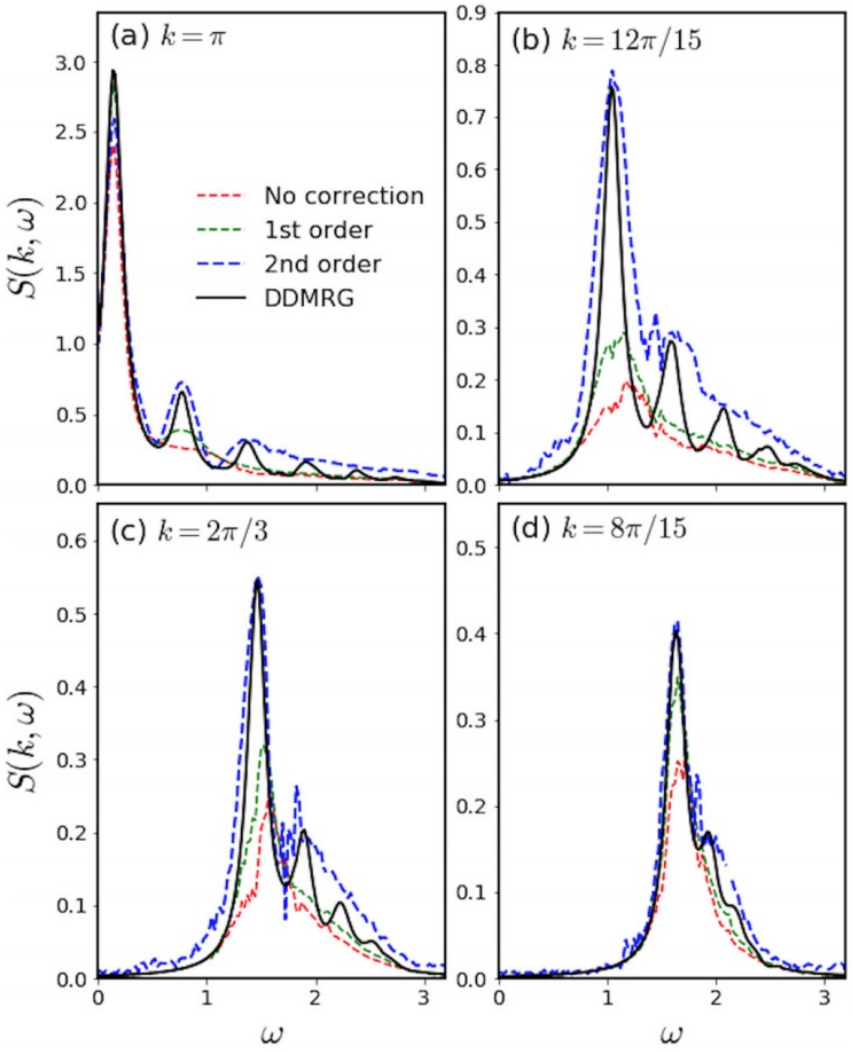
Results

Heisenberg chain L=30



Results

Heisenberg chain L=30



Chebyshev polynomials

They are defined by simple recursion relations

$$T_0 = 1; T_1 = x$$
$$T_{n+1} = 2xT_n - T_{n-1}$$

Each polynomial is bounded by 1 in the interval [-1,1]

They are orthogonal according to

$$\int_{-1}^{1} T_n(x)T_m(x)\frac{dx}{\sqrt{1-x^2}} = \begin{cases} 0 \; if \; n \neq m \\ \pi \; if \; n = m = 0 \\ \pi/2 \; if \; n = m \neq 0 \end{cases}$$

According to this, a function f:[-1,1]$\rightarrow \mathbb{R}$ can be expanded as

$$f(x) = \frac{1}{\pi\sqrt{1-x^2}}\left[\mu_0 + 2\sum_n \mu_n T_n(x)\right]$$

with $$\mu_n = \int_{-1}^{1} T_n(x)f(x)\frac{dx}{\sqrt{1-x^2}}$$ "Chebyshev moments"

Chebyshev expansion of spectral functions

If the energy spectrum of H is defined between (-1,1) one can show that the spectral function can be expanded as:

$$A(\omega) = \langle\psi|A_i^+ \delta(E_0 + \omega - H)A_j|\psi\rangle \approx \frac{1}{\pi\sqrt{1-x^2}}\left[g_0\mu_0 + 2\sum_n g_n\mu_n T_n(\omega)\right]$$

(Here, the g coefficient are "damping factors" that suppress spurious oscillations if we truncate the series to a finite number of moments).

The Chebyshev moments in this case are given by:

$$\mu_n = \langle\psi|A_i^+ T_n(H)A_j|\psi\rangle$$

We introduce the (non-orthogonal) Chebyshev vectors:

$$|t_0\rangle = A_i|\psi\rangle;\ |t_n\rangle = T_n(H)A_j|\psi\rangle \quad \text{"Chebyshev vectors"}$$

The moments are obtained as: $\mu_n = \langle t_o|t_n\rangle$

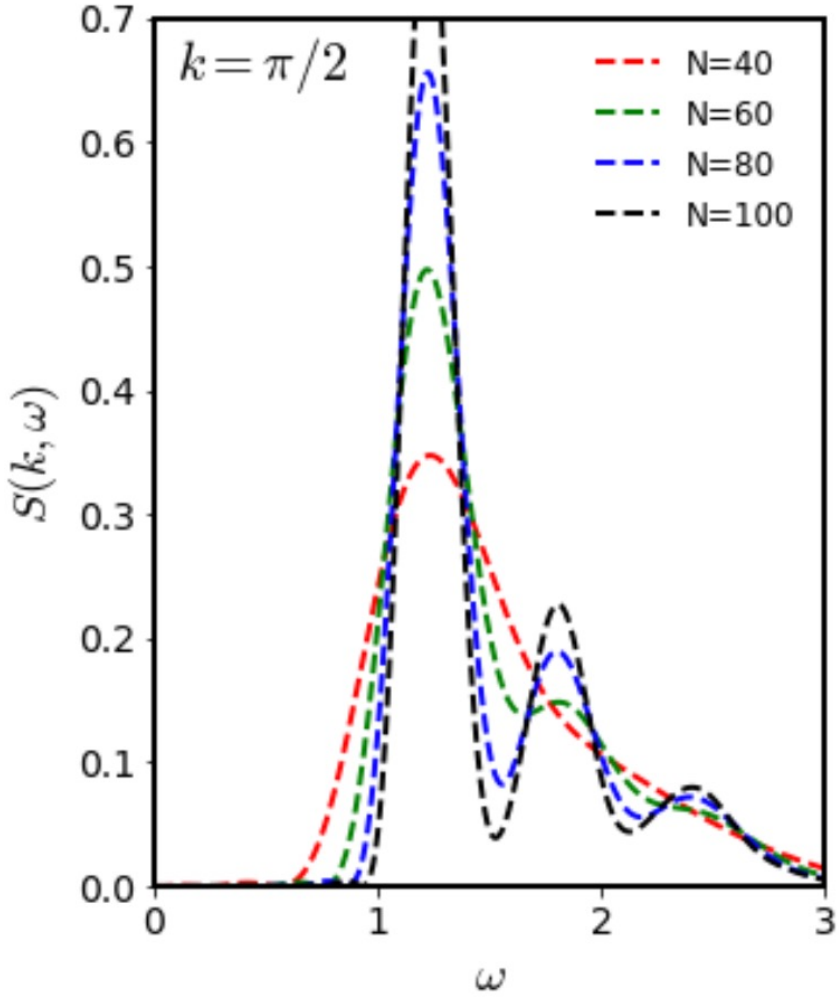
We can generate the Chebyshev vectors with the recursion

$$|t_1\rangle = H|t_0\rangle;\ |t_{n+1}\rangle = 2H\ |t_n\rangle - \ |t_{n-1}\rangle$$

A. Holzner, A. Weichselbaum, I. P. McCulloch, U. Schollwöck, and J. von Delft, PRB 83, 195115 (2011).
F. A. Wolf, J. A. Justiniano, I. P. McCulloch, and U. Schollwöck, Phys. Rev. B 91, 115144 (2015).

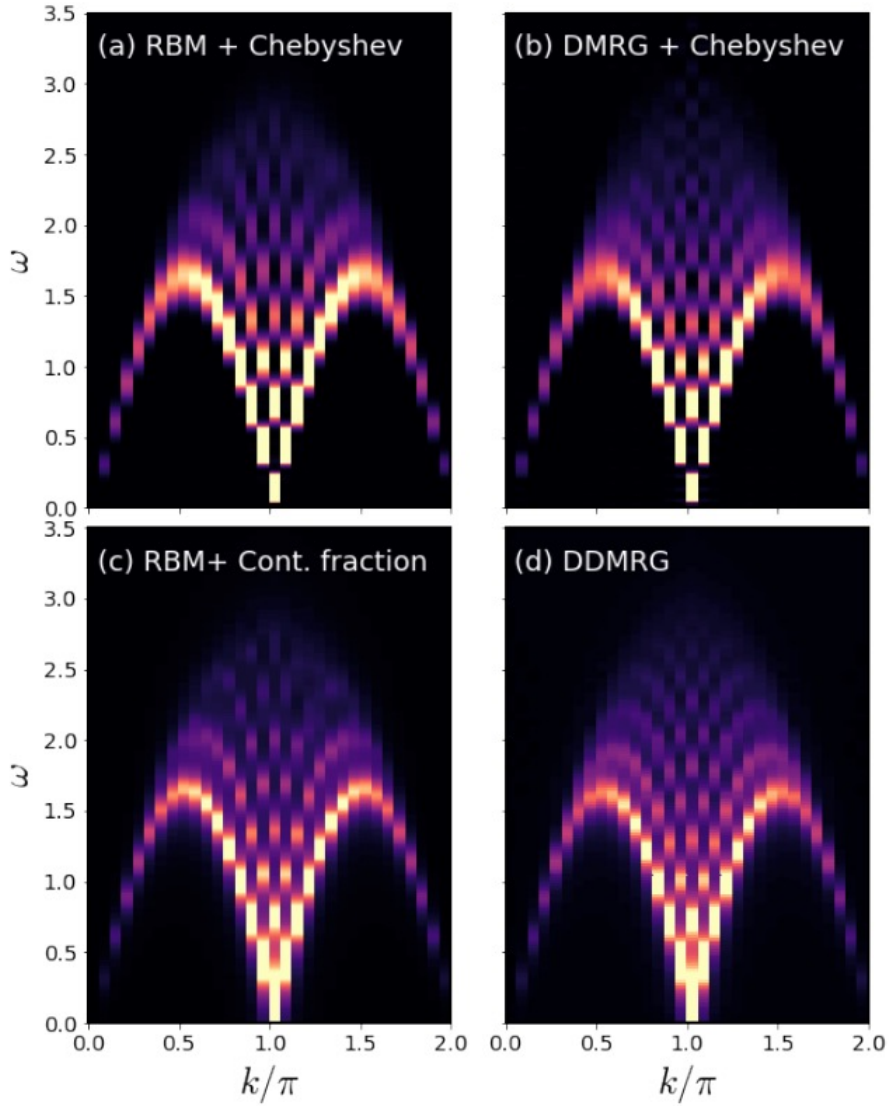
Results - Convergence

Heisenberg chain L=32



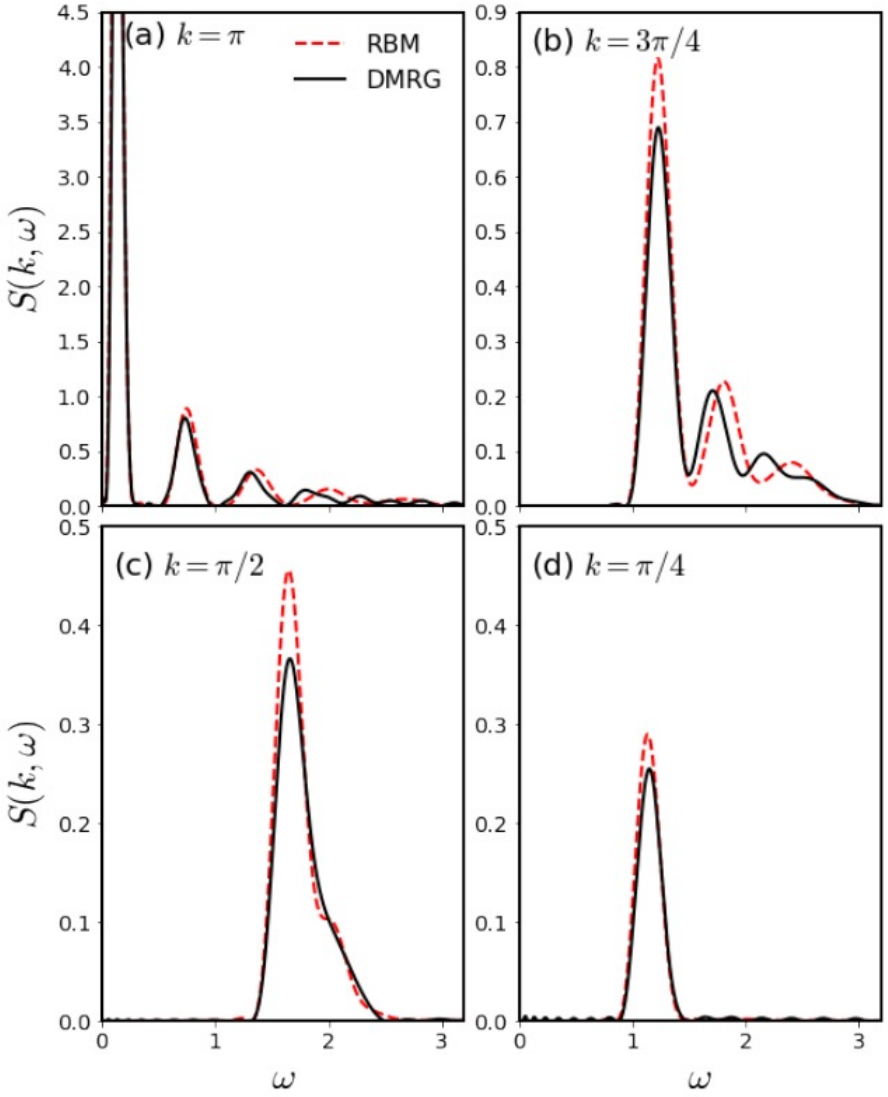
Results

Heisenberg chain L=32; 100 moments



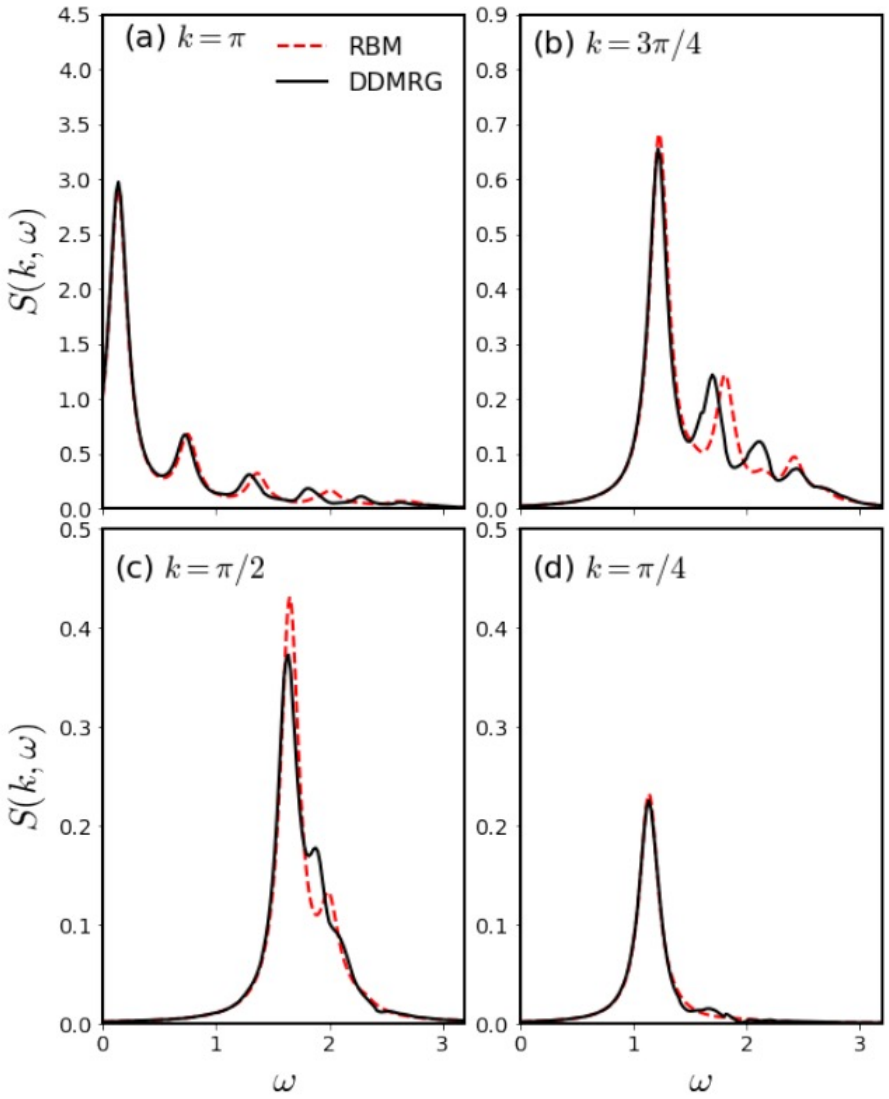
Results

Heisenberg chain L=32; 100 moments



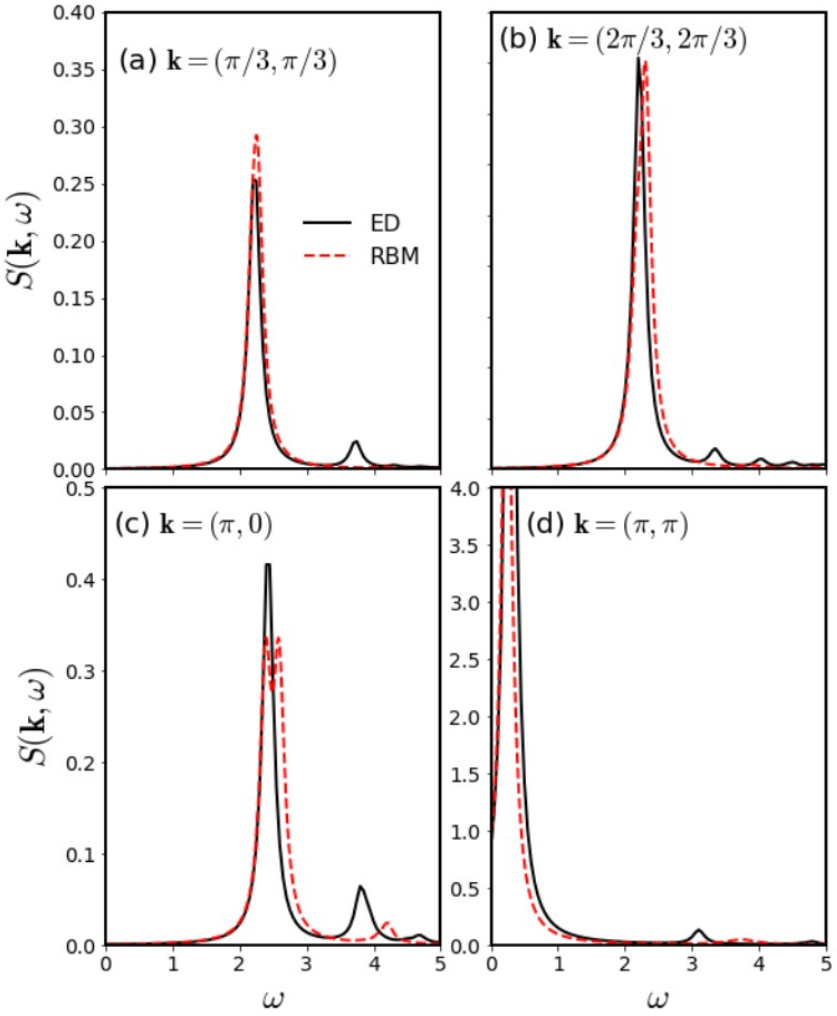
Results

Heisenberg chain L=32; 100 moments



Results in 2D

Heisenberg on the 6x6 square lattice



Conclusions

- We have developed two variational algorithms to calculate the spectral function of quantum many-body systems based on machine learning ideas. The information about the wave functions is encoded in the form of an RBM.
- Our work establishes new variational formulations to calculate Green's functions directly in the frequency domain using variational Monte Carlo.
- This formulation can be extended to other variational forms beyond neural networks and can help to overcome the curse of the area law that plagues methods such as tensor networks.
- Limitations stem from the expressivity of the RBMs and the computational cost of optimizing it.