Variational wave functions for frustrated spin models: from traditional methods to neural networks... and back

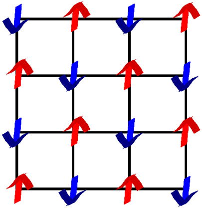
Federico Becca

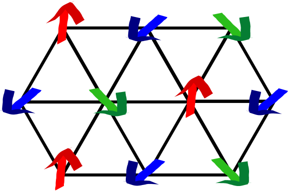
Machine Learning for Quantum Many-Body Physics, KITP 2019

UNIVERSITÀ
DEGLI STUDI DI TRIESTE

F. Ferrari, W.-J. Hu (SISSA → Rice), and S. Sorella (SISSA, Trieste),
F. Ferrari and J. Carrasquilla (Vector Institute)

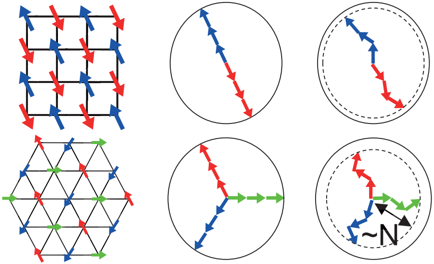
$$\mathcal{H} = J \sum_{\langle R,R' \rangle} \mathbf{S}_R \cdot \mathbf{S}_{R'}$$





- Classical limit ($S \to \infty$): **broken O(3) symmetry**
  (magnetization can be collinear, coplanar, or non-coplanar)
- Semi-classical corrections (linear spin waves): **gapless excitations**
  Magnons carrying $S = 1$ quantum number (Goldstone modes)
  Holstein and Primakoff, Phys. Rev. **58**, 1098 (1940)

The classical ground state is "dressed" by quantum fluctuations



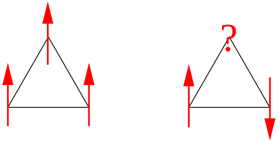
- The lattice breaks up into sublattices
- Each sublattice keeps an extensive magnetization

$$S(q) = \frac{1}{N} \langle \Psi_0 | \left| \sum_R \mathbf{S}_R e^{iqR} \right|^2 | \Psi_0 \rangle = \frac{1}{N} \sum_{R,R'} \langle \Psi_0 | \mathbf{S}_R \cdot \mathbf{S}_{R'} | \Psi_0 \rangle e^{iq(R-R')}$$

$$S(q) = \begin{cases} O(1) & \text{for all q's} \quad \rightarrow \text{short-range correlations} \\ S(q_0) \propto N & \text{for} q = q_0 \quad \rightarrow \text{long-range order} \end{cases}$$

We have to stay away from the classical limit

- Small value of the spin S, e.g., $S = 1/2$ or $S = 1$
- Frustration of the super-exchange interactions
  (not all terms of the energy can be optimized simultaneously)



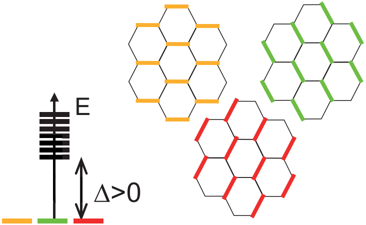
- Low spatial dimensionality: $D = 2$ is the "best" choice
  In $D = 1$ there is no magnetic order, given the Mermin-Wagner theorem
  (not possible to break a continuous symmetry in D=1, even at $T = 0$)

  Pitaevskii and Stringari, J. Low Temp. Phys. **85**, 377 (1991)

- [Large continuous rotation symmetry group, e.g., SU(2), SU(N) or Sp(2N)]

  Arovas and Auerbach, Phys. Rev. B **38**, 316 (1988); Arovas and Auerbach, Phys. Rev. Lett. **61**, 617 (1988)

  Read and Sachdev, Phys. Rev. Lett. **66**, 1773 (1991); Read and Sachdev, Nucl. Phys. **B316**, 609 (1989)

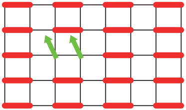


$J_1 - J_2$ Heisenberg model on the hexagonal lattice

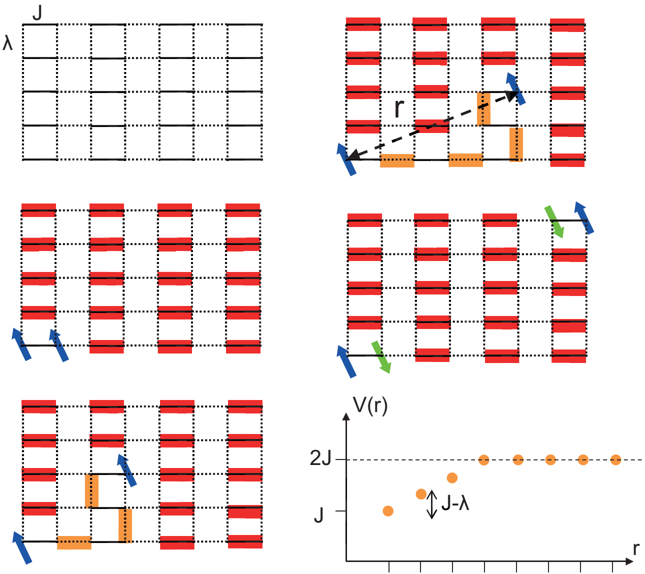
Fouet, Sindzingre, and Lhuillier, Eur. Phys. J. B **20**, 241 (2001)

Properties:

- Short-range spin-spin correlations
- Spontaneous breakdown of some lattice symmetries $\rightarrow$ ground-state degeneracy
- **Gapped $S = 1$ excitations** (triplons)



Valence-bond solids have conventional excitations

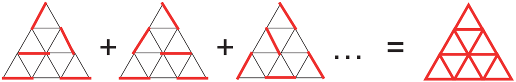


What's happening when destroying magnetic order: spin liquids

- Anderson's idea: the short-range resonating-valence bond (RVB) state:

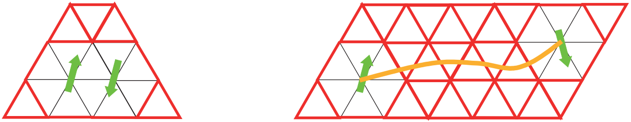
  Anderson, Mater. Res. Bull. **8**, 153 (1973)

  Linear superposition of many (an exponential number) of valence-bond configurations

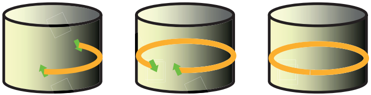


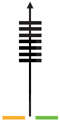
Spatially uniform state

- Spin excitations? No dimer order → we may have deconfined spinons



- Spinon fractionalization and topological degeneracy





Distinct ground states that are not connected by any local operator

Wen, Phys. Rev. B **44**, 2664 (1991); Oshikawa and Senthil, Phys. Rev. Lett. **96**, 060601 (2006)

Spin liquids are "highly-entangled" states

$\rho_A = Tr_B |\Psi\rangle\langle\Psi|$

$S(A) = -Tr_A \rho_A \log \rho_A$

$S(A) \approx c \times L - \gamma$

(L is the length of the boundary)

$\gamma > 0 \Longrightarrow$ NO product state

[This highly-entangled state has been introduced by Chernyshev (HFM 2018, unpublished)]

Some general features of highly-entangled phases are:

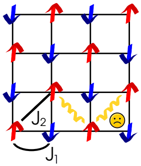
- The ground state cannot be smoothly deformed into a product state
- The entanglement entropy shows deviations from the strict area law
- Some elementary excitations are *non-local*
  (they cannot be created individually by any set of local operators)
- These quasiparticles exhibit some form of long-range interactions
  (anyonic mutual statistics)

Savary and Balents, Rep. Prog. Phys. **80**, 016502 (2017)

The frustrated Heisenberg model in two dimensions

- The simplest model on the square lattice

$$\mathcal{H} = J_1 \sum_{\langle R,R' \rangle} \mathbf{S}_R \cdot \mathbf{S}_{R'} + J_2 \sum_{\langle\langle R,R' \rangle\rangle} \mathbf{S}_R \cdot \mathbf{S}_{R'}$$



- Infinitely many papers with partially contradictory results

Gong *et al.*, Phys. Rev. Lett. **113**, 027201 (2014)

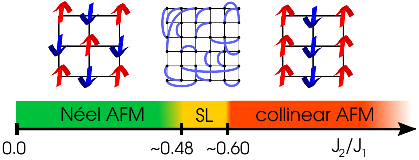
Wang *et al.*, Phys. Rev. B **94**, 075143 (2016)

Poilblanc and Mambrini, Phys. Rev. B **96**, 014414 (2017)

Haghshenas and Sheng, Phys. Rev. B **97**, 174408 (2018)

Wang and Sandvik, Phys. Rev. Lett. **121**, 107202 (2018)

- Possibly, a gapless spin liquid (SL) emerges between two AF phases



Hu *et al.*, Phys. Rev. B **88**, 060402 (2013)

- Start from a (classical) ordered state in the XY plane

$$|\Phi_{\rm cl}\rangle = \prod_R \left(|\uparrow\rangle_R + e^{iQR} |\downarrow\rangle_R \right)$$

The weight of every spin configuration (along z) is 1

Relative phases are determined by Q

- Include a two-body Jastrow factor to modify the weights

$$|\Psi\rangle = \exp\left[-\frac{1}{2} \sum_{R,R'} v_{R,R'} S^z_R S^z_{R'} \right] |\Phi_{\rm cl}\rangle$$

$v_{R,R'}$ is a pseudo-potential that can be optimized

The Jastrow factor creates entanglement (typically area law)

This wave function corresponds to the one of the spin-wave approximation

Manousakis, Rev. Mod. Phys. **63**, 1 (1991)

Franjic and Sorella, Prog. Theor. Phys. **97**, 399 (1997)

Accuracy of Jastrow wave function

- Size consistent wave function

  $O(N)$ variational parameters (with translational invariance)

  $O(N^2)$ scaling for sampling: easy calculations up to $N \approx 500 \div 1000$ (on a desktop)

- The accuracy depends upon the lattice

  Rather good variational energy for unfrustrated lattices: $\Delta E/E_{\rm ex} \approx 1\%$

  Accuracy on observables follows (ϵ on $E \to \sqrt{\epsilon}$ on O): $\Delta M/M_{\rm ex} \approx 10\%$

- It breaks spin SU(2) symmetry

  Bad for finite lattices (the ground state is fully symmetric)

  Good for the thermodynamic limit (if the ground state breaks the symmetry)

- The Jastrow factor gives the correct physics

  For small momenta: $S^z(q) \propto q$: Goldstone modes from the Feynman construction

  $|\Psi_q\rangle = S^z_q|\Psi\rangle$ gives $E_q - E \propto \frac{q^2}{S^z_q}$

Standard mean-field approach

Consider the spin-1/2 Heisenberg model on a generic lattice

$$\mathcal{H} = \sum_{R,R'} J_{R,R'} \mathbf{S}_R \cdot \mathbf{S}_{R'}$$

In a standard mean-field approach, each spin couples to an effective field generated by the surrounding spins:

$$\mathcal{H}_{\rm MF} = \sum_{R,R'} J_{R,R'} \left\{ \langle \mathbf{S}_R \rangle \cdot \mathbf{S}_{R'} + \mathbf{S}_i \cdot \langle \mathbf{S}_{R'} \rangle - \langle \mathbf{S}_R \rangle \cdot \langle \mathbf{S}_{R'} \rangle \right\}$$

However, by definition, spin liquids have a zero magnetization:

$$\langle \mathbf{S}_R \rangle = 0$$

How can we construct a mean-field approach for such disordered states?

We need to construct a theory in which all classical order parameters are vanishing

From spins to electrons...

- Consider the spin-1/2 Heisenberg model on a generic lattice

$$\mathcal{H} = \sum_{R,R'} J_{R,R'} \mathbf{S}_R \cdot \mathbf{S}_{R'}$$

- A faithful representation of spin-1/2 is given by

$$S^a_R = \frac{1}{2} c^\dagger_{R,\alpha} \sigma^a_{\alpha,\beta} c_{R,\beta}$$

SU(2) gauge redundancy
e.g., $c_{R,\beta} \to e^{i\theta_R} c_{R,\beta}$

- The spin model is transformed into a purely interacting electronic system

$$\mathcal{H} = \sum_{R,R'} J_{R,R'} \sum_{\sigma,\sigma'} \left(\sigma\sigma' c^\dagger_{R,\sigma} c_{R,\sigma} c^\dagger_{R',\sigma'} c_{R',\sigma'} + \frac{1}{2} \delta_{\sigma',\bar{\sigma}} c^\dagger_{R,\sigma} c_{R,\sigma'} c^\dagger_{R',\sigma'} c_{R',\sigma} \right)$$

- One spin per site $\to$ we must impose the constraint

$$c^\dagger_{i,\uparrow} c_{i,\uparrow} + c^\dagger_{i,\downarrow} c_{i,\downarrow} = 1$$

- The SU(2) symmetric mean-field approximation gives a BCS-like form

$$\mathcal{H}_0 = \sum_{R,R',\sigma} t_{R,R'} c^\dagger_{R,\sigma} c_{R',\sigma} + \sum_{R,R'} \Delta_{R,R'} c^\dagger_{R,\uparrow} c^\dagger_{R',\downarrow} + h.c.$$

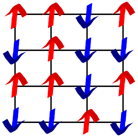
$\{t_{R,R'}\}$ and $\{\Delta_{R,R'}\}$ define the mean-field Ansatz $\longrightarrow$ BCS spectrum $\{\epsilon_\alpha\}$

The constraint is no longer satisfied locally (only on average)

- The constraint can be inserted by the Gutzwiller projector $\to$ **RVB**

$$\boxed{|\Psi_0\rangle = \mathcal{P}_G |\Phi_0\rangle}$$

$$\mathcal{P}_G = \prod_R (n_{R,\uparrow} - n_{R,\downarrow})^2$$



- The exact projection can be treated within the variational Monte Carlo approach

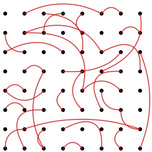
F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems*

The projected wave function

- The mean-field wave function has a BCS-like form

$$|\Phi_0\rangle = \exp\left\{\sum_{i,j} f_{i,j} c^\dagger_{i,\uparrow} c^\dagger_{j,\downarrow}\right\}|0\rangle = \left[1 + \sum_{i,j} f_{i,j} c^\dagger_{i,\uparrow} c^\dagger_{j,\downarrow} + \frac{1}{2}\left(\sum_{i,j} f_{i,j} c^\dagger_{i,\uparrow} c^\dagger_{j,\downarrow}\right)^2 + \dots\right]|0\rangle$$

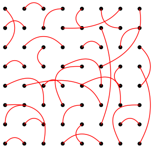
It is a linear superposition of all singlet configurations (that may overlap)

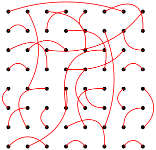


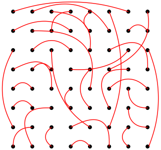
$+ \dots$

- After projection, only non-overlapping singlets survive: the resonating valence-bond (RVB) wave function

Anderson, Science **235**, 1196 (1987)







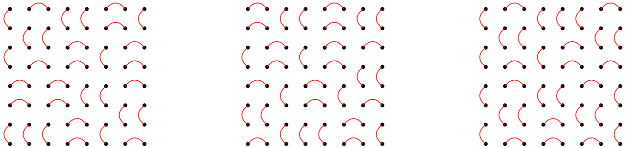
$+ \dots$

The projected wave function

- The mean-field wave function has a BCS-like form

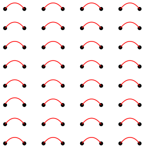
$$|\Phi_0\rangle = \exp\left\{\sum_{i,j} f_{i,j} c^\dagger_{i,\uparrow} c^\dagger_{j,\downarrow}\right\}|0\rangle = \left[1 + \sum_{i,j} f_{i,j} c^\dagger_{i,\uparrow} c^\dagger_{j,\downarrow} + \frac{1}{2}\left(\sum_{i,j} f_{i,j} c^\dagger_{i,\uparrow} c^\dagger_{j,\downarrow}\right)^2 + \dots\right]|0\rangle$$

- Depending on the pairing function $f_{i,j}$, different RVB states may be obtained...



$+ \dots$

- ...even with valence-bond order (valence-bond crystals)



- For a non-magnetic (spin liquid or valence-bond solid) state

$$\boxed{|\Psi_0\rangle = \mathcal{P}_G |\Phi_0\rangle}$$

$$\mathcal{H}_0 = \sum_{R,R',\sigma} t_{R,R'} c^\dagger_{R,\sigma} c_{R',\sigma} + \sum_{R,R'} \Delta_{R,R'} c^\dagger_{R,\uparrow} c^\dagger_{R',\downarrow} + h.c.$$

- For an antiferromagnetic state

$$\boxed{|\Psi_0\rangle = \mathcal{P}_{S_z} \mathcal{J} \mathcal{P}_G |\Phi_0\rangle}$$

$$\mathcal{H}_0 = \sum_{R,R',\sigma} t_{R,R'} c^\dagger_{R,\sigma} c_{R',\sigma} + \Delta_{\rm AF} \sum_{R} e^{iQR} \left(c^\dagger_{R,\uparrow} c_{R,\downarrow} + c^\dagger_{R,\downarrow} c_{R,\uparrow} \right)$$

In analogy with the Jastrow wave function, the magnetic moment in the $x - y$ plane

$\mathcal{J} = \exp \left(\frac{1}{2} \sum_{R,R'} v_{R,R'} S^z_R S^z_{R'} \right)$ is the spin-spin Jastrow factor

Towards the exact ground state

How can we improve the variational state?
By the application of a few Lanczos steps!

$$|\Psi_{p-LS}\rangle = \left(1 + \sum_{m=1,\ldots,p} \alpha_m \mathcal{H}^m\right) |\Psi_{VMC}\rangle$$

- For $p \to \infty$, $|\Psi_{p-LS}\rangle$ converges to the exact ground state, provided $\langle \Psi_0 | \Psi_{VMC} \rangle \neq 0$
- On large systems, only FEW Lanczos steps are affordable: We can do up to $p = 2$

The variance extrapolation

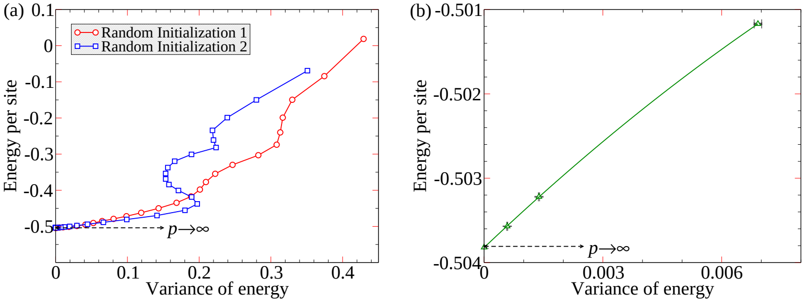
- A zero-variance extrapolation can be done

Whenever $|\Psi_{VMC}\rangle$ is sufficiently close to the ground state:

$$E \simeq E_0 + \text{const} \times \sigma^2$$

$$E = \langle \mathcal{H} \rangle / N$$
$$\sigma^2 = (\langle \mathcal{H}^2 \rangle - E^2)/N$$

How does it work?



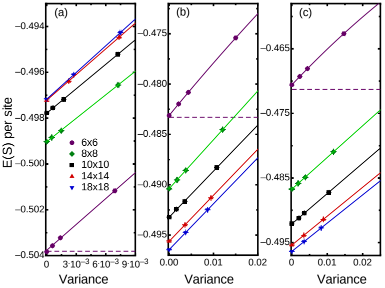
A few energies on $L \times L$ clusters with PBC

$J_2 = 0.40$	DMRG (8192)	VMC ($p = 0$)	VMC ($p = 2$)	VMC ($p = \infty$)
$L = 6$	−0.529744	−0.52715(1)	−0.52957(1)	−0.52972(1)
$L = 8$	−0.525196	−0.52302(1)	−0.52539(1)	−0.52556(1)
$L = 10$	−0.522391	−0.52188(1)	−0.5240(1)	−0.52429(2)
$J_2 = 0.45$	DMRG (8192)	VMC ($p = 0$)	VMC ($p = 2$)	VMC ($p = \infty$)
$L = 6$	−0.515655	−0.51364(1)	−0.51558(1)	−0.51566(1)
$L = 8$	−0.510740	−0.50930(1)	−0.51125(1)	−0.51140(1)
$L = 10$	−0.507976	−0.50811(1)	−0.51001(1)	−0.51017(2)
$J_2 = 0.50$	DMRG (8192)	VMC ($p = 0$)	VMC ($p = 2$)	VMC ($p = \infty$)
$L = 6$	−0.503805	−0.50117(1)	−0.50357(1)	−0.50382(1)
$L = 8$	−0.498175	−0.49656(1)	−0.49886(1)	−0.49906(1)
$L = 10$	−0.495530	−0.49521(1)	−0.49755(1)	−0.49781(2)
$J_2 = 0.55$	DMRG (8192)	VMC ($p = 0$)	VMC ($p = 2$)	VMC ($p = \infty$)
$L = 6$	−0.495167	−0.48992(1)	−0.49399(1)	−0.49521(7)
$L = 8$	−0.488160	−0.48487(1)	−0.48841(2)	−0.48894(3)
$L = 10$	−0.485434	−0.48335(1)	−0.48693(3)	−0.48766(6)

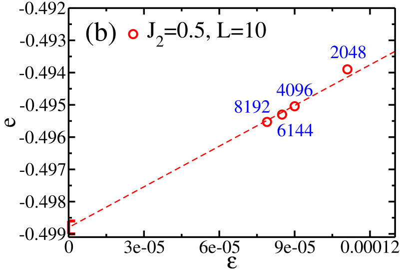
Hu, Becca, Parola, and Sorella, Phys. Rev. B **88**, 060402 (2013)

Gong, Zhu, Sheng, Motrunich, and Fisher, Phys. Rev. Lett. **113**, 027201 (2014)

Extrapolations to the ground state energy



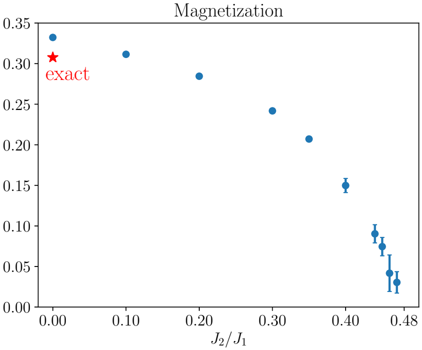
W.-J. Hu *et al.*, Phys. Rev. B **88**, 060402 (2013)

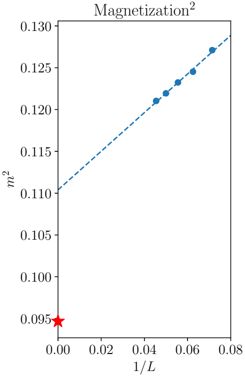


S.-S. Gong *et al.*, Phys. Rev. Lett. **113**, 027201 (2014)

$$m^2 = \lim_{r \to \infty} \langle \mathbf{S}_r \cdot \mathbf{S}_0 \rangle$$

- Magnetization computed for finite clusters from 10×10 to 22×22





- A finite staggered magnetization is related to a finite $\Delta_{\rm AF}$ in the wave function

The present understanding of the magnetically disordered phase

- Valence-bond solid

Read and Sachdev, Phys. Rev. Lett. **62**, 1694 (1989)

Sachdev and Bhatt, Phys. Rev. B **41**, 9323 (1990)

Singh, Weihong, Hamer, and Oitmaa, Phys. Rev. B **60**, 7278 (1999)

Capriotti and Sorella, Phys. Rev. Lett. **84**, 3173 (2000)

Mambrini, Lauchli, Poilblanc, and Mila, Phys. Rev. B **74**, 144422 (2006)

Gong *et al.*, Phys. Rev. Lett. **113**, 027201 (2014)

- Gapped or gapless spin liquid

Capriotti, Becca, Parola, and Sorella, Phys. Rev. Lett. **87**, 097201 (2001)

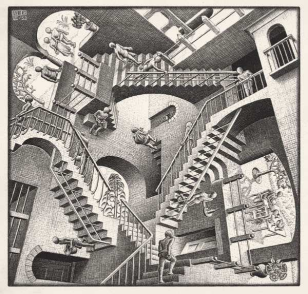
Jiang, Yao, and Balents, Phys. Rev. B **86**, 024424 (2012)

Wang, Poilblanc, Gu, Wen, and Verstraete, Phys. Rev. Lett. **111**, 037202 (2013)

Poilblanc and Mambrini, Phys. Rev. B **96**, 014414 (2017)

Haghshenas and Sheng, Phys. Rev. B **97**, 174408 (2018)

Wang and Sandvik, Phys. Rev. Lett. **121**, 107202 (2018)

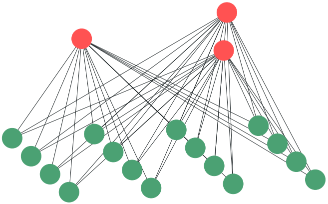


Restricted Boltzmann Machines (RBMs) entered into the game...

MANY-BODY PHYSICS

Solving the quantum many-body problem with artificial neural networks

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



$$|\Psi_{\rm RBM}\rangle = \sum_{h_a=\pm 1} \exp \left[\sum_{R,a} W_{R,a} S^z_R h_a + \sum_a b_a h_a \right] |\Phi_{\rm cl}\rangle$$

$$|\Psi_{\rm RBM}\rangle \propto \prod_a \exp \left\{ \log\cosh \left[b_a + \sum_R W_{R,a} S^z_R \right] \right\} |\Phi_{\rm cl}\rangle$$

- Hidden spin variables $(h_1, \dots, h_\alpha)$
- Network parameters (b, W)
- Generalization of the Jastrow factor that includes many-body interactions

The "sign problem"

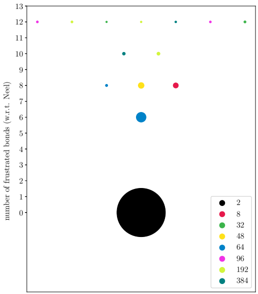
- With a real parametrization (b and W), the sign structure is fixed by the reference state
- A complex parametrization is often needed to "learn" the correct signs

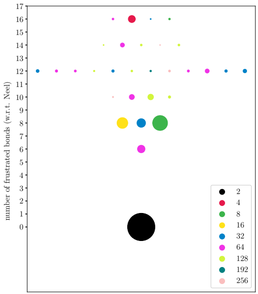
J_2/J_1	$\langle s \rangle$
0.00	1
0.05	1
0.10	1
0.15	1
0.20	1
0.25	1
0.30	1
0.35	0.9999937
0.40	0.9995104
0.45	0.9927903
0.50	0.9608835
0.55	0.8704279
0.60	0.6144326

The average Marshall sign on the 6×6 cluster

$$\langle s \rangle = \sum_x |\langle x|\Psi_{\rm ex}\rangle|^2 \mathrm{sign}\{M(x)\langle x|\Psi_{\rm ex}\rangle\}$$

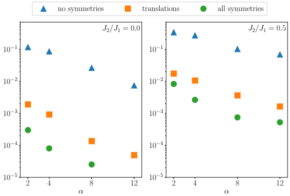
Weights of the exact ground state on the 4×4 cluster





Learning signs and amplitudes on the 4 × 4 cluster

- Fixing the sign to the exact one and optimizing amplitudes



- Optimizing only the sign

$$F(x) = \prod_a \exp\left\{ i \log\cosh\left[b_a + \sum_R W_{R,a} S^z_R(x) \right] \right\}$$

$$C = 1 - \left| \sum_x |\Psi_{\rm ex}(x)|^2 \text{sign}\{F(x)\Psi_{\rm ex}(x)\} \right|$$

α	C for $J_2/J_1 = 0.0$
1	0.30381655
4	0.00000004

α	C for $J_2/J_1 = 0.5$
1	0.02770868
4	0.00312562

- We combine **Gutzwiller-projected fermionic states and RBMs**

$$\langle x|\Psi_{\rm RBM}\rangle = \prod_T \prod_a \exp\left\{ \log\cosh\left[b_a + \sum_R W_{R,a} S^z_{T(R)} \right] \right\} \langle x|\Phi_0\rangle$$

where $|\Phi_0\rangle$ is the ground state of a quadratic Hamiltonian

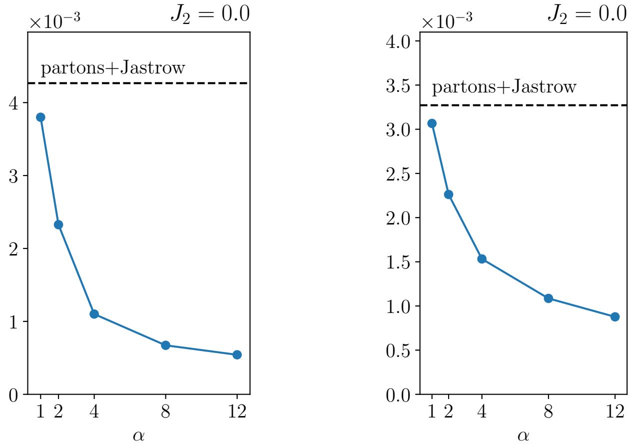
Different from Choo, Carleo, and Neupert, talk at the conference

- We impose **translational symmetry ($Q = 0$) on the RBM**

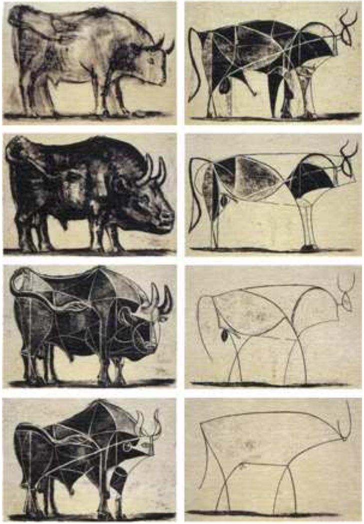
- We consider **real parameters for $J_2 = 0$** to impose the Marshall-sign rule

- We consider **complex parameters for $J_2 > 0$** to change the fermionic signs

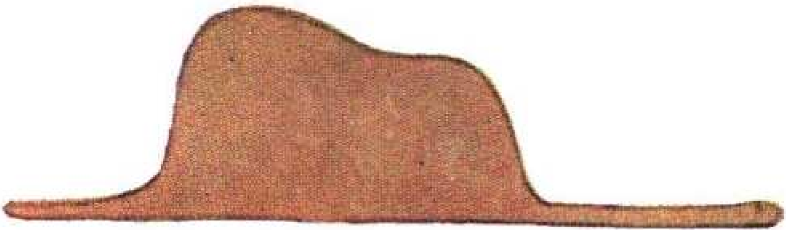
The unfrustrated Heisenberg model



The unfrustrated Heisenberg model



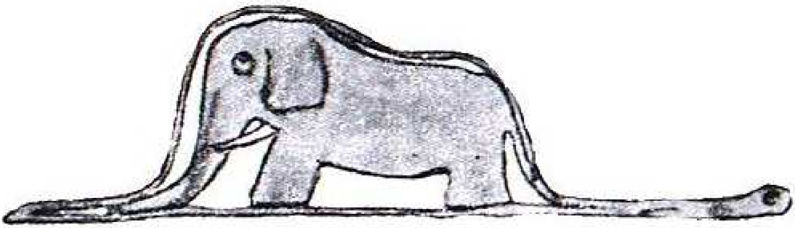
With a poor accuracy we see a hat...



Antoine de Saint-Exupéry, *Le Petit Prince* (1943)

Problems where a good accuracy is needed: the highly-frustrated region

By increasing the accuracy we identify an elephant!



Antoine de Saint-Exupéry, *Le Petit Prince* (1943)

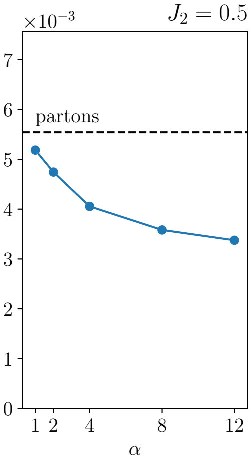
Problems where a good accuracy is needed: the highly-frustrated region

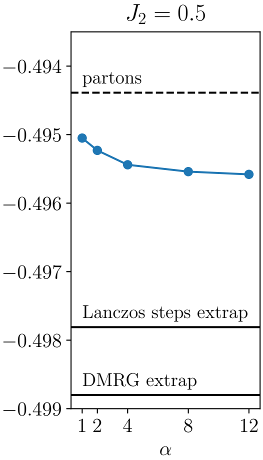
Maybe by further improving the accuracy we will discover the truth...

Antoine de Saint-Exupéry, *Le Petit Prince* (1943)

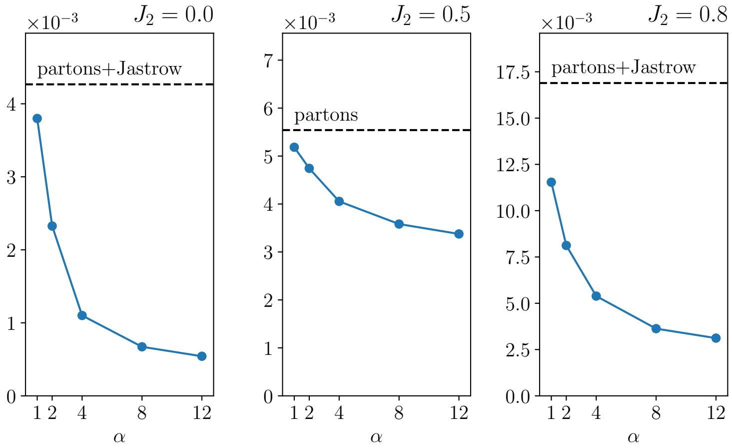
The highly-frustrated case $J_2/J_1 = 0.5$

$$\langle x|\Psi_{\rm RBM}\rangle = \prod_T \prod_\alpha \exp\left\{\log\cosh\left[b_\alpha + \sum_R W_{R,\alpha} S^z_{T(R)}\right]\right\}\langle x|\Phi_0\rangle$$

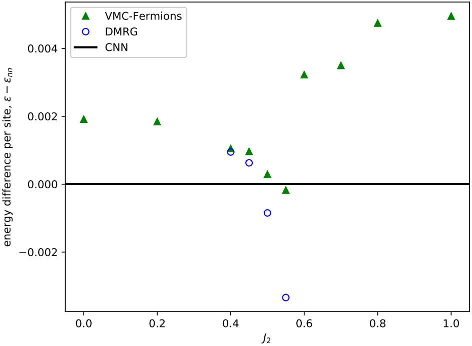




A summary on the 6×6 cluster



Comparison with Choo, Carleo, and Neupert



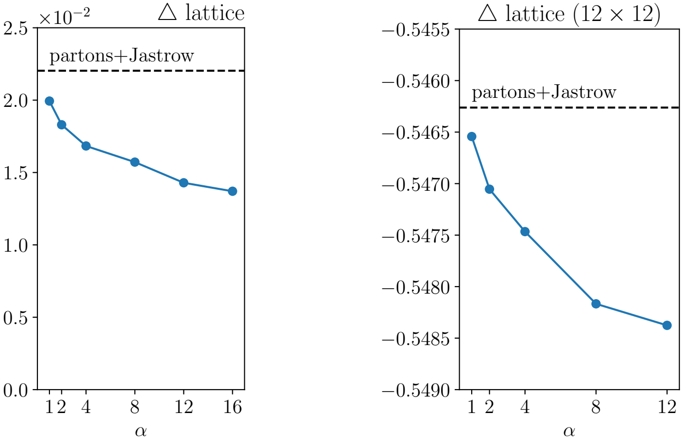
- CNN with about 4000 variational parameters
- Fermionic state with about 40 variational parameters

What about non-collinear order?

- Heisenberg model on the triangular lattice

  The exact sign structure is not known

  The ground state has coplanar magnetic order



What is wrong with **these** RBMs?

These RBMs assume that

- **Spin degrees of freedom S_R^z are the relevant objects**
- **A particular form of the spin-spin correlation is present** $\log\cosh(z)$

The first assumption is correct for (collinear) magnetically ordered phase

The second assumption limits the flexibility of the wave function

Many variational parameters

- **Difficult optimizations**
- **No transparent description to understand the physical properties**

Often there are many local minima, with completely different parameters

Calculations are limited to $O(100)$ sites

A more educated guess would be desirable

- **Parametrization in terms of spinons and not spins**