Variational Wave Functions
for Strongly-Correlated Fermionic Systems

Federico Becca

Autumn School on Correlated Electrons:
Many-Body Methods for Real Materials
Jülich, September 19th
1 Introduction

2 Hubbard and Heisenberg Models
- Variational wave functions for the Heisenberg model
- Variational Monte Carlo method
- Wave functions for the Hubbard model: the (density) Jastrow factor
- Variational wave functions for the Hubbard model: the backflow terms

3 Further Developments and Generalizations
- Generalization to multi-orbital Hubbard models
- Application of a few Lanczos steps
- Restricted Boltzmann Machines for the Heisenberg model
Correlation effects in real materials

Band theory + Odd number of el. per unit cell ⇒ Metal

Many materials with an odd number of electrons per unit cell are insulators for example transition-metal compounds

These are called Mott insulators

We will consider lattice models
Motivations and strategy

GOAL: Describe an insulating state of purely Mott type

In Mott insulators localization is induced by strong correlation

Failure of the single-particle picture

The variational approach gives insight into the ground state properties

Until very recently a consistent Mott insulating state was not available

- Long-range (density) Jastrow factor
- “Backflow” correlations

Main result:

Metal-insulator transition and strong-coupling phase

But also:

Accurate metallic and/or superconducting phase when doping Mott insulators
The Hubbard model is the prototype for correlated electrons on the lattice [like the Ising model for classical magnetism] NO exact solution in D>1

Antiferromagnetic super-exchange
\[ J = \frac{4t^2}{U} \]

NO charge fluctuations, only spin

\[ \mathcal{H} = -t \sum_{\langle i,j \rangle, \sigma} c_{i,\sigma}^\dagger c_{j,\sigma} + h.c. + U \sum_i n_{i,\uparrow} n_{i,\downarrow} \]
Antiferromagnetic order?

At large $U/t$ there are antiferromagnetic correlations

At $T = 0$ (ground state), magnetic order may be present

- Long-range magnetic order on square and honeycomb lattices (QMC)
  - Triangular lattice: convincing evidence from different numerical methods

Magnetic frustration: a way to destabilize magnetic AF order

Non-magnetic ground states may exist

- Evidence for the absence of magnetic order on the kagome lattice
  - Frustrated square and triangular lattices: evidence from different numerical methods
**The infinite-\(U\) limit (Heisenberg model)**

- **M magnetically ordered state: the (spin) Jastrow factor**

\[ \mathcal{H}_{\text{cl}} = \Delta_{AF} \sum_j S_j \cdot n_j \]

\[ n_j = \{ \cos(Q \cdot R_j), \sin(Q \cdot R_j) \} \]

\[ \mathcal{J}_s = \exp[-\frac{1}{2} \sum_{i,j} u_{i,j} S_i^z S_j^z] \]

\[ |\psi_{AF}\rangle = \mathcal{J}_s |\Phi_{\text{cl}}\rangle \]

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

- **Non-magnetic state (spin liquid): the Gutzwiller projector**

\[ \mathcal{H}_{\text{BCS}} = \sum_{i,j,\sigma} t_{i,j} c_{i,\sigma}^\dagger c_{j,\sigma} + \sum_{i,j} \Delta_{i,j} [c_{i,\uparrow}^\dagger c_{j,\downarrow}^\dagger + c_{j,\uparrow}^\dagger c_{i,\downarrow}^\dagger] + h.c. \]

\[ |\psi_{\text{SL}}\rangle = P_G |\Phi_{\text{BCS}}\rangle \]

\[ P_G = \prod_i (1 - n_{i,\uparrow} n_{i,\downarrow}) \]

Jastrow wave functions for magnetically ordered phases

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

\[ |\Phi_{cl}\rangle = \prod_j \left( |\uparrow\rangle_j + e^{iQj} |\downarrow\rangle_j \right) \]

No correlation

Q determines the periodicity

• Include a two-body (spin) Jastrow factor to modify the weights

\[ |\Psi_{AF}\rangle = \exp \left[ -\frac{1}{2} \sum_{i,j} u_{i,j} S_i^z S_j^z \right] |\Phi_{cl}\rangle \]

The Jastrow factor creates correlations

\(u_{i,j}\) is a pseudo-potential to be optimized

This wave function corresponds to the one of the spin-wave approximation
Gutzwiller-projected states for non-magnetic phases

• The mean-field wave function has a BCS form

$$|\Phi_{\text{BCS}}\rangle = \exp \left\{ \frac{1}{2} \sum_{i,j} f_{i,j} c_{i,\uparrow}^\dagger c_{j,\downarrow}^\dagger \right\} |0\rangle$$

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

• With $$P_G = \prod_i (1 - n_{i,\uparrow} n_{i,\downarrow})$$, only non-overlapping singlets survive

The wave function corresponds to the resonating valence-bond (RVB) state
These wave functions cannot be treated by using analytical approaches

- They can be treated within quantum Monte Carlo

\[
E(\Psi) = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \sum_x P(x) e_L(x) \approx \frac{1}{M} \sum_{i=1}^{M} e_L(x_i)
\]

\[
P(x) = \frac{|\langle x | \Psi \rangle|^2}{\langle \Psi | \Psi \rangle} \text{ ("classical" Monte Carlo)}
\]

\[
e_L(x) = \frac{\langle x | H | \Psi \rangle}{\langle x | \Psi \rangle} = \sum_y \langle x | H | y \rangle \langle y | \Psi \rangle \langle x | \Psi \rangle
\]

Requirements

- \(\langle x | \Psi \rangle\) must be efficiently computed
- The Hamiltonian must be local

Becca and Sorella, Quantum Monte Carlo Approaches for Correlated Systems (Cambridge University Press, 2017)
• **Magnetically ordered state** $\rightarrow |\Phi_{cl}\rangle$ is a product state

\[
\mathcal{J}_s = \exp[-\frac{1}{2} \sum_{i,j} u_{i,j} S_i^z S_j^z]
\]

$|\Psi_{AF}\rangle = \mathcal{J}_s |\Phi_{cl}\rangle$

$|\chi\rangle$ is the (Ising) basis with fixed $S^z$ on each site

• $\mathcal{J}_s$ is diagonal and $\langle \chi | \Phi_{cl} \rangle$ is a number $\rightarrow \langle \chi | \Psi_{AF} \rangle$ computed in $O(N^2)$

• $e_L(\chi)$ is computed in $O(N^3)$, for a short-range Hamiltonian

**Hasting-Metropolis algorithm: Markov chain $|\chi\rangle \rightarrow |\chi'\rangle$**

• \(\frac{P(\chi')}{P(\chi)}\) is computed in $O(1)$ for local moves!

• $e_L(\chi)$ is computed in $O(N)$

• Updating is done in $O(N)$
Accuracy of Jastrow wave function

- Size consistent wave function
  
  $O(N)$ variational parameters (with translational invariance): $u_{i,j} \rightarrow u_r$
  
  $O(N^2)$ 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: $\Delta E/E_{\text{ex}} \approx 1\%$
  
  Accuracy on observables follows ($\epsilon$ on $E \rightarrow \sqrt{\epsilon}$ on $O$): $\Delta M/M_{\text{ex}} \approx 10\%$

- 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)

- Goldstone modes from the Feynman construction
  
  For small momenta: $\langle \Psi_{\text{AF}} | S^z_{-q} S^z_q | \Psi_{\text{AF}} \rangle / \langle \Psi_{\text{AF}} | \Psi_{\text{AF}} \rangle \propto q$
  
  $|\psi_q\rangle = S^z_q |\Psi_{\text{AF}}\rangle$ gives $E_q - E \propto \frac{q^2}{S_q}$
Variational Monte Carlo for the RVB wave function

- Non-magnetic state (spin liquid) $\Rightarrow |\Phi_{BCS}\rangle$ is an entangled state

$$P_G = \prod_i (1 - n_i, \uparrow n_i, \downarrow)$$

$|\psi_{SL}\rangle = P_G |\Phi_{BCS}\rangle$

$|x\rangle$ is the basis with one electron per site, fixed $S^z$ on each site

- $P_G$ is the identity and $\langle x|\Phi_{BCS}\rangle$ is a determinant $\Rightarrow \langle x|\psi_{SL}\rangle$ computed in $O(N^3)$

- $e_L(x)$ is computed in $O(N^4)$, for a short-range Hamiltonian

Hasting-Metropolis algorithm: Markov chain $|x\rangle \rightarrow |x'\rangle$

- $\frac{P(x')}{P(x)}$ is computed in $O(1)$ for local moves!

- $e_L(x)$ is computed in $O(N)$

- Updating is done in $O(N^2)$
Accuracy of RVB wave function

• Size consistent wave function

\[ O(1) \] variational parameters (few distances): \( t_{i,j} \rightarrow t_r \) and \( \Delta_{i,j} \rightarrow \Delta_r \)

\[ O(N^3) \] scaling for sampling: easy calculations up to \( N \approx 100 \div 400 \) (on a desktop)

• The accuracy depends upon the lattice

Rather good variational energy for frustrated lattices: \( \Delta E/E_{\text{ex}} \approx 1\% \)

Accuracy on observables follows (\( \epsilon \) on \( E \rightarrow \sqrt{\epsilon} \) on \( O \))

• It does not break spin SU(2) symmetry

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

Good for the thermodynamic limit (if the ground state does not break the symmetry)

• Fractional \( S = 1/2 \) spinon excitations and “gauge” excitations

Free (“deconfined”) \( S = 1/2 \) objects are expected

In addition, neutral \( S = 0 \) excitations should exist

**FINITE U repulsion: the antiferromagnetic case**

Magnetic wave function (stable for $t' = 0$ and $U > 0$)

$$\mathcal{H}_{MF} = -t \sum_{\langle i,j \rangle, \sigma} c_{i,\sigma}^{\dagger} c_{j,\sigma} + h.c. + \Delta_{AF} \sum_{j} e^{iQ \cdot R_{j}} S_{j}^{x} \quad \Rightarrow \left| \Phi_{MF} \right\rangle$$

$$\mathcal{J}_{s} = \exp \left[ -\frac{1}{2} \sum_{i,j} u_{i,j} S_{i}^{z} S_{j}^{z} \right] \left| \Psi_{AF} \right\rangle = \mathcal{J}_{s} \left| \Phi_{MF} \right\rangle$$

![Graph showing the energy as a function of $1/U$ for different $S^z$ and $S^x$](image-url)
Finite $U$ repulsion: the non-magnetic case

Gutzwiller wave function

$$\mathcal{H}_0 = -t \sum_{\langle i,j \rangle, \sigma} c_{i,\sigma}^\dagger c_{j,\sigma} + h.c. \quad \Rightarrow |D\rangle$$

$$|\Psi_g\rangle = e^{-g \sum_i n_{i,\uparrow} n_{i,\downarrow}} |D\rangle$$

- $g = 0$, the non-interacting wave function is recovered
- $g = \infty$, the full Gutzwiller projector is obtained

No exact calculations, except 1D


Monte Carlo sampling is possible

The Gutzwiller wave function is metallic for ANY $g \neq \infty$

It does not correlate charge fluctuations (empty and doubly occupied sites)

In any realistic insulator there are charge fluctuations

Empty sites (Holons, H) and doubly occupied sites (Doublons, D) play a crucial role for the conduction

H and D must be correlated otherwise an electric field would induce a current
**Finite $U$ repulsion: the non-magnetic case**

Short-range holon-doublon wave function

\[ |\Psi_{hd}\rangle = e^{f \sum_{\langle l,m \rangle} h_l d_m} |\Psi_g\rangle = e^{f \sum_{\langle l,m \rangle} h_l d_m} e^{-g \sum_i n_i,\uparrow n_i,\downarrow} |D\rangle \]

- Put nearest-neighbor correlation among holons and doublons

**Exact calculations on small clusters**


Monte Carlo sampling is possible


H and D farer than nearest neighbors are uncorrelated: metallic for ANY $f$
The (density) Jastrow wave function

\[ |\psi\rangle = J_c |D\rangle \]

\[ J_c = \exp \left[ -\frac{1}{2} \sum_{i,j} v_{i,j} n_i n_j \right] = \exp \left[ -\frac{1}{2} \sum_q v_q n_{-q} n_q \right] \]

\[ |D\rangle \] is an uncorrelated determinant, possibly including BCS pairing

Find the optimal set of parameters \( v_{i,j} \) which minimizes the energy without any a-priori assumption


**Metal or insulator?**

**Ansatz for the low-energy excitations**

\[ |\psi_q\rangle = n_q |\psi_0\rangle \]

\[ N_q = \frac{\langle \psi_0 | n_{-q} n_q |\psi_0\rangle}{\langle \psi_0 | \psi_0 \rangle} \]

**f-sum rule**

\[ \Delta_q = \frac{\langle \psi_q | (H - E_0) |\psi_q\rangle}{\langle \psi_q | \psi_q \rangle} = \frac{\langle \psi_0 | [n_{-q}, [H, n_q]] |\psi_0\rangle}{2N_q} \sim \frac{q^2}{N_q} \]

\[ N_q \sim |q| \Rightarrow \Delta_q \to 0 \Rightarrow \text{metal} \]

\[ N_q \sim q^2 \Rightarrow \Delta_q \text{ is finite} \Rightarrow \text{insulator} \]

Gutzwiller (left) and Jastrow (right) wave functions for \( U = 4 \) and 10
Reatto-Chester relation for $N_q$ and Jastrow factor

$$N_q = \frac{\langle \Psi | n_{-q} n_q | \Psi \rangle}{\langle \Psi | \Psi \rangle}$$

RPA Reatto and Chester, Phys. Rev. 155, 88 (1967)

- For continuous systems
- In the weak-coupling regime

Two-dimensional (paramagnetic) Hubbard model

\[
N_q = \frac{N_q^0}{1 + 2v_q N_q^0} \approx \frac{1}{v_q}
\]
Strong-coupling limit of the energy per site

One dimension

<table>
<thead>
<tr>
<th>Energy per site E [4t^2/U]</th>
<th>1/U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gutzwiller L=18</td>
<td>♦</td>
</tr>
<tr>
<td>Gutzwiller L=82</td>
<td>□</td>
</tr>
<tr>
<td>Jastrow L=18</td>
<td>○</td>
</tr>
<tr>
<td>Jastrow L=82</td>
<td>⬤</td>
</tr>
</tbody>
</table>

Two dimensions

<table>
<thead>
<tr>
<th>Energy per site E [4t^2/U]</th>
<th>1/U</th>
</tr>
</thead>
<tbody>
<tr>
<td>t'=0</td>
<td>▲</td>
</tr>
<tr>
<td>t'/t=0.7</td>
<td>▼</td>
</tr>
</tbody>
</table>

Poor accuracy in 2D systems: especially in presence of frustration
Standard (very)-large-$U$ Approach

Suppose we have a good ansatz $|\Psi_H\rangle$ for $U = \infty$

$$E = \langle \Psi_H | H_{\text{Heis}} | \Psi_H \rangle$$

Then a good ansatz for the Hubbard model in the large-$U$ limit is

$$|\Psi\rangle = e^{iS} |\Psi_H\rangle \quad iS = \frac{1}{U} (T^+ - T^-)$$


**Difficult to treat**

- Expand $e^{iS} \approx 1 + iS$ (not size consistent)

- Perform the Hubbard-Stratonovich decoupling

$$\langle x_0 | \Psi \rangle = \langle x_0 | \Psi_H \rangle$$

$$\langle x_1 | \Psi \rangle = \frac{1}{U} \left\{ \langle x_0^{\uparrow\downarrow} | \Psi_H \rangle + \langle x_0^{\downarrow\uparrow} | \Psi_H \rangle \right\}$$
The backflow wave function in the continuum considers fictitious coordinates of the electrons

\[ r^b_{\alpha} = r_{\alpha} + \sum_{\beta} \eta_{\alpha,\beta} (r_{\beta} - r_{\alpha}) \]

- Proposed for roton excitations in liquid Helium
  Feynman and Cohen, Phys. Rev. 102, 1189 (1956)
- Implemented in Monte Carlo calculations to study bulk $^3$He
- Used to improve the electron gas

Apply backflow to a lattice model

\[ \phi_k(r^b_{\alpha}) \simeq \phi^b_k(r_{\alpha}) \equiv \phi_k(r_{\alpha}) + \sum_{\beta} c_{\alpha,\beta} \phi_k(r_{\beta}) \]

\[ \phi_k = \text{single particle orbitals} \]
The backflow wave function

To favor the recombination of neighboring charge fluctuations

\[ \phi_k^b(r_i, \sigma) \equiv \epsilon \phi_k(r_i, \sigma) + \sum_j \eta_{i,j} D_i H_j \phi_k(r_j, \sigma) \]

\[ D_i = n_{i,\uparrow} n_{i,\downarrow} \quad H_i = (1 - n_{i,\uparrow})(1 - n_{i,\downarrow}) \]

- The determinant part of the wave function includes correlations
- Backflow correlations can modify the nodes of the variational wave function
- Jastrow factor can change only amplitudes
Important backflow parameters up to the range of the Hamiltonian

Irrelevant backflow parameters for longer distances

Backflow parameters are particularly important in the insulating phase
Backflow correlations make it possible to reach the fully-projected state

In the frustrated regime, backflow terms are useful also in the AF wave function
Generalization to multi-orbital Hubbard models

\[ H_{\text{kin}} = -\sum_{\langle i,j \rangle, \sigma} \sum_{\alpha, \beta} t_{i,j}^{\alpha,\beta} c_{i,\alpha,\sigma}^\dagger c_{j,\beta,\sigma} + h.c. \]

\[ H_{\text{int}} = U \sum_i \sum_{\alpha} n_{i,\alpha,\uparrow} n_{i,\alpha,\downarrow} + U' \sum_i \sum_{\alpha < \beta} n_{i,\alpha} n_{i,\beta} \]

\[ H_{\text{Hund}} = -J \sum_{i,\sigma,\sigma', \alpha < \beta} c_{i,\alpha,\sigma}^\dagger c_{i,\alpha',\sigma'} c_{i,\beta,\sigma}^\dagger c_{i,\beta',\sigma'} - J' \sum_i \sum_{\alpha < \beta} c_{i,\alpha,\uparrow}^\dagger c_{i,\alpha',\downarrow} c_{i,\beta,\uparrow}^\dagger c_{i,\beta',\downarrow} + h.c. \]

\[ |\psi\rangle = J_c |D\rangle \]

- Orbital-dependednt Jastrow factor:

\[ J_c = \exp \left( -\frac{1}{2} \sum_{i,j} \sum_{\alpha, \beta} v_{i,j}^{\alpha,\beta} n_{i,\alpha} n_{j,\beta} \right) \]

(Similar for the spin Jastrow factor)
Towards the exact ground state

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

\[ |\psi_{p-LS} \rangle = \left( 1 + \sum_{m=1,\ldots,p} \alpha_m \mathcal{H}^m \right) |\psi \rangle \]

- For \( p \to \infty \), \( |\psi_{p-LS} \rangle \) converges to the exact ground state, provided \( \langle \psi_0 | \psi \rangle \neq 0 \)

- On large systems, only a FEW Lanczos steps are affordable:
  \[ \langle x | \mathcal{H}^m | \psi \rangle \text{ (with } m = 1, \ldots, p \text{) must be computed for a given } |x\rangle \]
  We can do up to \( p = 2 \)
A zero-variance extrapolation can be done

Whenever $|\Psi\rangle$ 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?
Restricted Boltzmann Machines (RBMs)

Solving the quantum many-body problem with artificial neural networks

Giuseppe Carlen¹ and Matthias Troyer²,³

\[ |\Psi_{\text{RBM}}\rangle = \sum_{h_a=\pm 1} \exp \left[ \sum_{i,a} W_{i,a} S^z_i h_a + \sum_a b_a h_a \right] |\Phi_{\text{cl}}\rangle \]

\[ |\Psi_{\text{RBM}}\rangle \propto \prod_a \exp \left\{ \log \cosh \left[ b_a + \sum_R W_{i,a} S^z_i \right] \right\} |\Phi_{\text{cl}}\rangle \]

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