Variational Monte Carlo for the Hubbard model

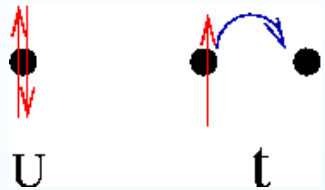
Luca F. Tocchio

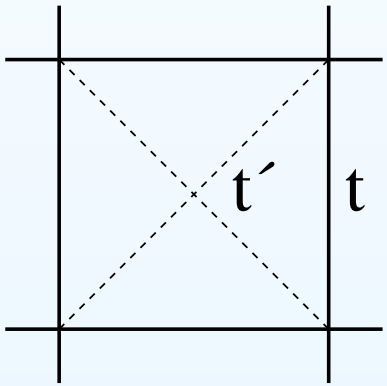
Politecnico di Torino

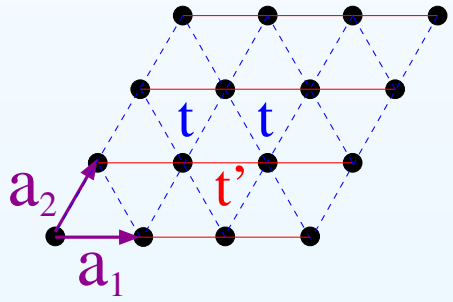
Jülich, 24th September 2026

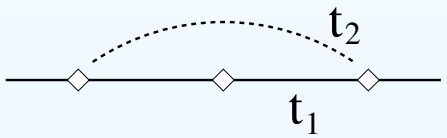
The single-orbital Hubbard model

$$\mathcal{H}_{\text{Hub}} = \sum_{i,j,\sigma} t_{ij} c^{\dagger}_{i,\sigma} c_{j,\sigma} + \text{h.c.} + \sum_{i} U n_{i,\uparrow} n_{i,\downarrow}$$





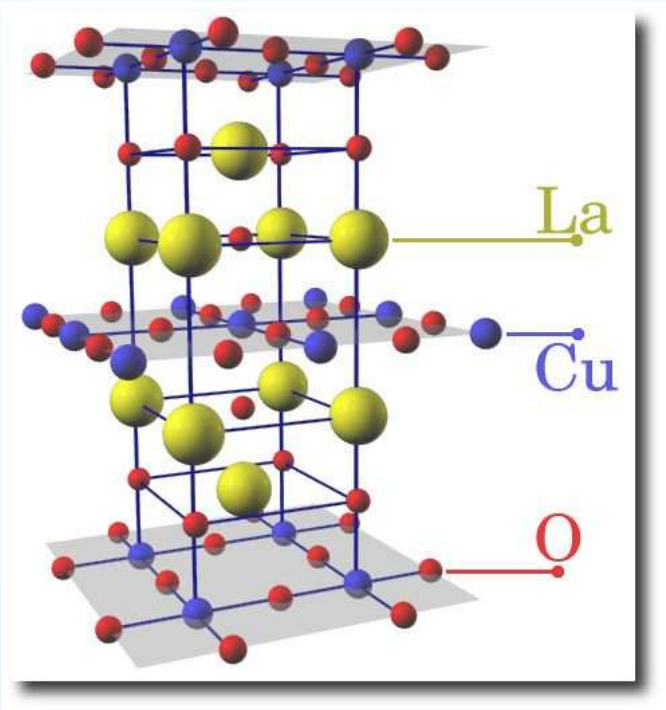


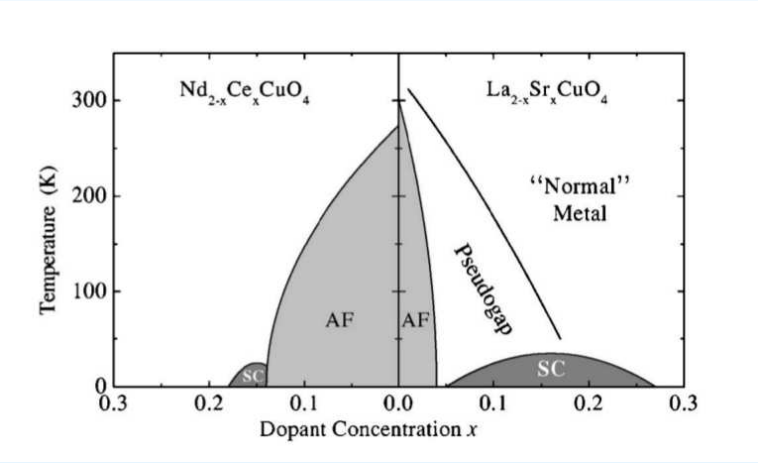


- Correlation: On-site Coulomb repulsion U
- Delocalization: Lattice-dependent competing hoppings
- Doping: Change the average electron density n
  (# electrons divided by the # sites)
- Minimal set of interactions describes complicated phenomena: Mott insulators, magnetism, superconductivity

High-temperature superconductors: cuprates

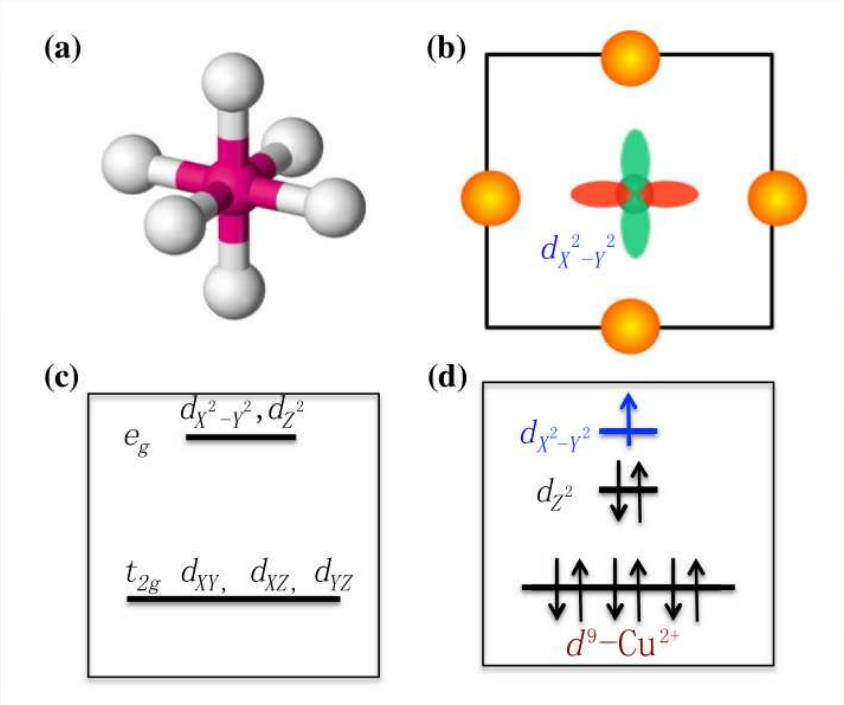
- Cuprates are antiferromagnetic Mott insulators e.g., La_2CuO_4
- Holes are added in the CuO planes, e.g., by replacing trivalent La with divalent Sr ($La_{2-x}Sr_xCuO_4$)
- Superconductivity develops in the CuO planes (that have a square lattice geometry)



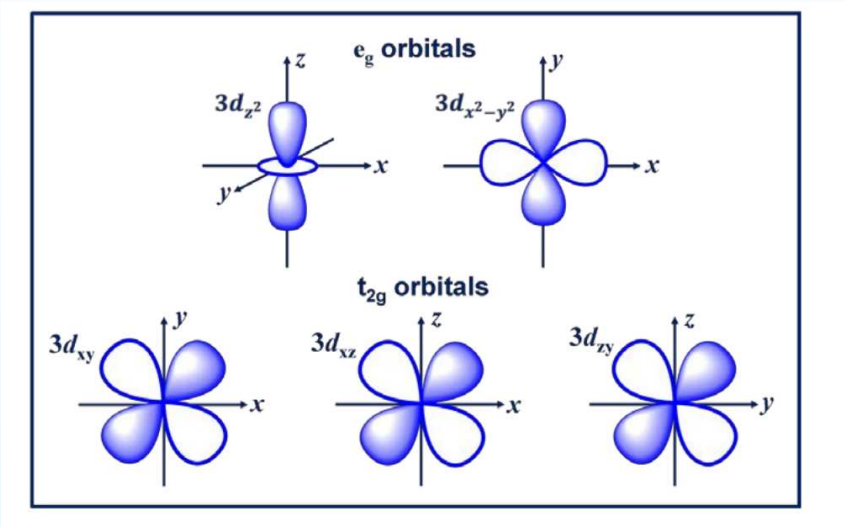


Damascelli *et al.*, RMP 75, 473 (2003)

A single-orbital physics



J. Hu, Sci. Bull. 61, 561 (2016)

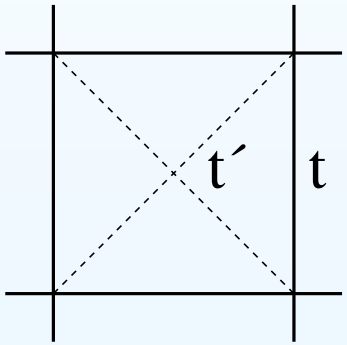


A. Biswas *et al.*, arXiv:1508.04929

- Cu atom surrounded by six oxygens (octahedron)
- Octahedral symmetry favors the t_{2g} orbitals: Lower Coulomb repulsion among electrons in the d orbitals and in the ligands (oxygens)
- Further lattice distortion along z favors the d_{Z^2} over the $d_{X^2-Y^2}$
- 9 electrons per Cu atom leave the $d_{X^2-Y^2}$ orbitals half-filled (one e^- per site)

Extended (single-orbital) Hubbard model

$$\mathcal{H}_{\text{Hub}} = -t \sum_{\langle R,R'\rangle,\sigma} c^{\dagger}_{R,\sigma} c_{R',\sigma} - t' \sum_{\langle\langle R,R'\rangle\rangle,\sigma} c^{\dagger}_{R,\sigma} c_{R',\sigma} + \text{H.c.} + \sum_{R} U n_{R,\uparrow} n_{R,\downarrow}$$



- The Hubbard model on the square lattice is appropriate for the cuprates
- In cuprates there is a realistic next-nearest neighbor hopping t' with $t'/t < 0$, estimated by Density Functional Theory

  Pavarini, Dasgupta, Saha-Dasgupta, Jepsen, and Andersen, PRL 87, 047003 (2001)
- It reproduces an antiferromagnetic Mott insulator at half filling
- Superconductivity emerges by doping the Hubbard model away from half filling (even if competition with charge and spin modulations)

"Solving" the Hubbard model: The variational approach

- The general model is not exactly solvable except in 1D
- Each site has four states: $|0\rangle$, $|\uparrow\rangle$, $|\downarrow\rangle$, $|\uparrow,\downarrow\rangle$
- The Hilbert space grows up exponentially as 4^L: Exact diagonalization only for very small sizes
- Alternative approach: Construct wave functions $|\Psi_{\text{var}}\rangle$ that capture the correct low-energy properties of correlated systems
- Select the best approximation with the variational principle

$$E_{\text{var}} = \frac{\langle\Psi_{\text{var}}|\mathcal{H}|\Psi_{\text{var}}\rangle}{\langle\Psi_{\text{var}}|\Psi_{\text{var}}\rangle} \geq E_0 \text{ (exact energy)}$$

- The best approximation has the lowest energy
- How do we compute these expectation values? Monte Carlo!

Variational Monte Carlo for the ground-state properties of correlated Hamiltonians I

Luca F. Tocchio

Politecnico di Torino

Main reference: Becca and Sorella: Quantum Monte Carlo Approaches for Corrletaed Systems (Cambridge University Press, 2017)

Variational Monte Carlo (VMC)

Pros:

- Variational principle: the estimated energy is always an upper bound of the exact one
- Non-perturbative in the electonic correlation
- Performances not significantly affected by dimensionality or size of the local Hilbert space
- Spatial correlations and superconductivity can be properly described
- No sign problem

Cons:

- Ground state method: no dynamical correlations. VMC can be extended to work at finite temperatures and time dependent, not discussed here
- Difficult to quantify the systematic errors that come from the choice of the trial state: Results may be biased from the trial state
- Usually not the most accurate estimate for the variational energy. It may improved (e.g. with Green's Function Monte Carlo)

Variational Monte Carlo

- Consider an orthonormalized basis set $\{|x\rangle\}$
- It is convenient to choose a local basis, defined on the lattice sites
- On each lattice site the basis set has four elements (for the Hubbard model): $\{|\xi\rangle_i\} = \{|0\rangle_i, |\uparrow\rangle_i, |\downarrow\rangle_i, |\uparrow\downarrow\rangle_i\}$
- A generic element of the lattice basis set $|x\rangle$ is given by:

$$|x\rangle = |\xi\rangle_1 \cdots |\xi\rangle_L$$

- Convenient choice: There is a small number of non-vanishing matrix elements of the Hamiltonian on this basis since interactions are local in space (Hubbard-U) and hoppings are short range

Variational Monte Carlo

In quantum mechanics, we need to compute expectaction values of operators $\mathcal{O}$, as the energy

$$\langle \mathcal{O} \rangle = \frac{\langle \Psi | \mathcal{O} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_x \langle \Psi | x \rangle \langle x | \mathcal{O} | \Psi \rangle}{\sum_{x'} \langle \Psi | x' \rangle \langle x' | \Psi \rangle}$$

The number of elements in the sum grows exponentially with the system size (4^L in the Hubbard model): we cannot compute $\langle \mathcal{O} \rangle$ exactly! We can use a Monte Carlo sampling:

- Recast $\langle \mathcal{O} \rangle$ in a more convenient way

$$\langle \mathcal{O} \rangle = \frac{\sum_x |\langle \Psi | x \rangle|^2 \frac{\langle x | \mathcal{O} | \Psi \rangle}{\langle x | \Psi \rangle}}{\sum_{x'} |\langle \Psi | x' \rangle|^2} = \sum_x \mathcal{P}(x) \mathcal{O}_L(x)$$

Variational Monte Carlo

- $\langle \mathcal{O} \rangle$ is the average of the local estimator $\mathcal{O}_L(x)$ of the operator $\mathcal{O}$ over the probability distribution in quantum mechanics (non-negative and normalized, so no sign problem)

$$\langle \mathcal{O} \rangle = \sum_x \mathcal{P}(x) \mathcal{O}_L(x) \quad \mathcal{O}_L(x) = \frac{\langle x|\mathcal{O}|\Psi\rangle}{\langle x|\Psi\rangle} \quad \mathcal{P}(x) = \frac{|\langle \Psi|x\rangle|^2}{\sum_{x'} |\langle \Psi|x\rangle|^2}$$

- We construct a Markov chain (with the Metropolis algorithm) that generates a set of configurations $\{x_n\rangle\}$ distributed according to $\mathcal{P}(x)$
- We compute the quantum expectation value $\langle \mathcal{O} \rangle$ as the average of the random variable $\mathcal{O}_L(x)$:

$$\langle \mathcal{O} \rangle \sim \frac{1}{N} \sum_{n=1}^{N} \mathcal{O}_L(x_n)$$

Zero-variance property

- The most important observable is the energy (variational principle). The local estimator e_L is:

$$e_L(x) = \frac{\langle x|\mathcal{H}|\Psi\rangle}{\langle x|\Psi\rangle}$$

- Assume $|\Psi\rangle$ to be an eigenstate of $\mathcal{H}$ ($\mathcal{H}|\Psi\rangle = E|\Psi\rangle$):

$$e_L(x) = \frac{\langle x|\mathcal{H}|\Psi\rangle}{\langle x|\Psi\rangle} = E\frac{\langle x|\Psi\rangle}{\langle x|\Psi\rangle} = E$$

  E is exact with zero variance

- This is an extreme case, but, in general, the variance decreases when $|\Psi\rangle$ approaches an exact eigenstate.

Metropolis algorithm

A practical way to construct configurations x_n that are distributed according to $\mathcal{P}(x)$, after an initial equilibration time

- Propose a move generating a configuration x' according to a transition probability $T(x'|x)$ (the move can be chosen to be very simple,i.e., flip a spin in the Ising model, move an electron in the Hubbard one)
- Extract a random number $\eta \in (0, 1]$
- Accept the move ($x_{n+1} = x'$) if $\eta < A(x'|x) = \min\left\{1, \frac{\mathcal{P}(x')}{\mathcal{P}(x)}\right\}$. Otherwise $x_{n+1} = x$
- Good news: We only need to know probability ratios $\mathcal{P}(x')/\mathcal{P}(x)$: No need to normalize the probability distribution!

Calculation of observables

- Compute local estimators (e.g., $e_L(x)$ for the energy) every $O(N_e)$ steps (N_e is the number of electrons and one step is the proposal of a single electron move)
- In correlated electron systems, the acceptance probability is usually small $\Rightarrow$ Perform a binning technique to uncorrelate measures
- First write the observables on the hard disk and then perform statistical analysis, since correlation time is unknown

Correlated measures: a bit of statistics

Local estimators are computed every $O(N_e)$ steps. Still, configurations are not independent because the acceptance ratio in the Hubbard model is small!

- Consider the variance estimator s^2 for a single Monte Carlo run

$$s^2 = \frac{1}{N}\sum_i (x_i - \bar{x})^2 = \frac{1}{N}\sum_i x_i^2 - \left(\frac{1}{N}\sum_i x_i\right)^2$$

- Compute its average $\langle s^2 \rangle$:

$$\langle s^2 \rangle = \frac{1}{N}\sum_i \langle x_i^2 \rangle - \frac{1}{N^2}\sum_{i,j} \langle x_i x_j \rangle$$

Correlated measures

- When the $\langle x_i \rangle$ are independent and identically distributed $\langle s^2 \rangle$ is a good estimator of the variance:

  Using $\langle x_i x_j \rangle = \langle x_i \rangle \langle x_j \rangle = \langle x \rangle^2$ for $i \neq j$; $\langle x_i x_j \rangle = \langle x^2 \rangle$ for $i = j$

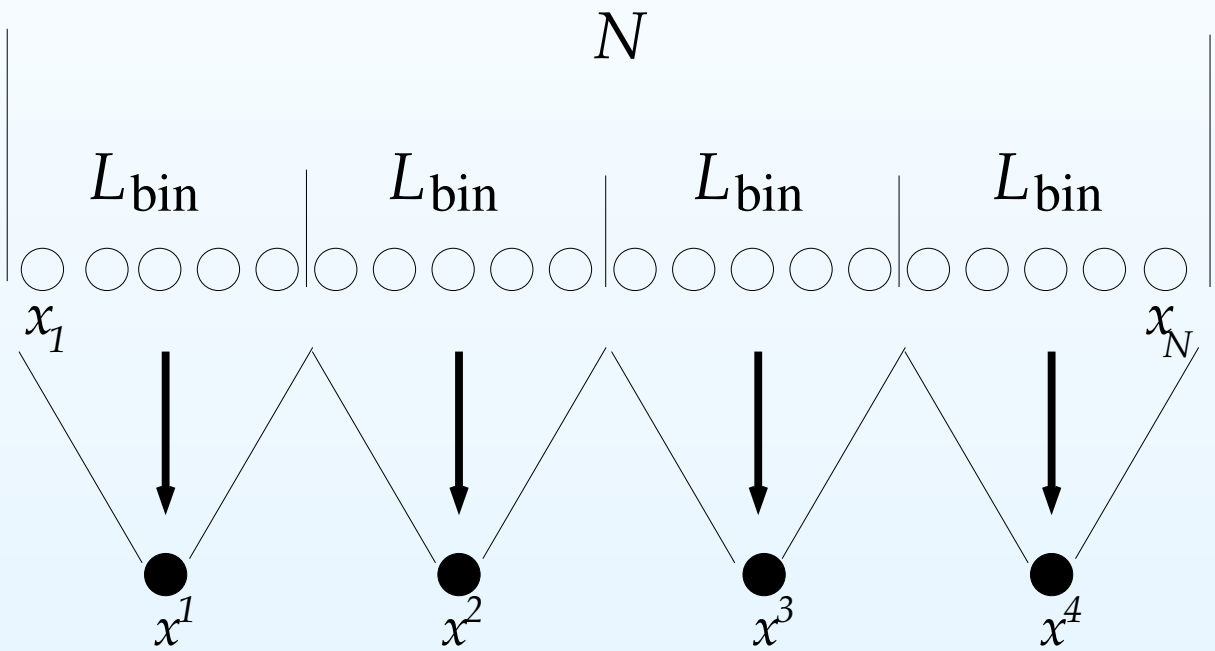
$$\langle s^2 \rangle = \frac{1}{N} \sum_i \langle x_i^2 \rangle - \frac{1}{N^2} \sum_{i,j} \langle x_i x_j \rangle = \left(1 - \frac{1}{N}\right) \left(\langle x^2 \rangle - \langle x \rangle^2\right) =$$

$$= \left(\frac{N-1}{N}\right) \sigma^2$$

- When $\langle x_i x_j \rangle \neq \langle x_i \rangle \langle x_j \rangle$, $\langle s^2 \rangle$ is no longer a proper estimator of the variance
- Solve the problem with the block analysis!

Binning technique (block analysis)

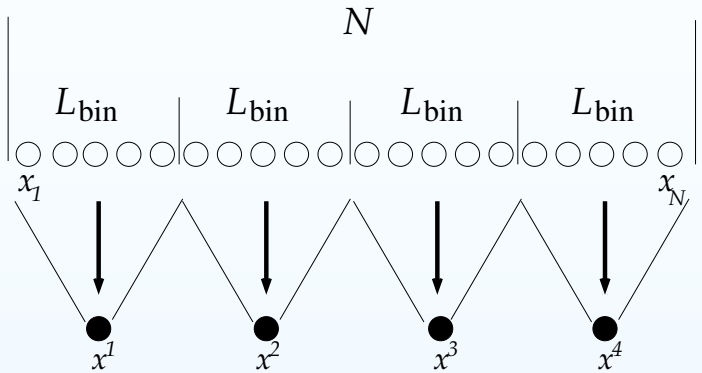
Divide the data set $\{x_i\}$ $i = 1, \ldots, N$ from a Markov chain in N_{bin} bins of length $L_{\text{bin}} = N/N_{\text{bin}}$ to uncorrelate data



- Consider the partial average on each bin j ($j = 1, \ldots, N_{\text{bin}}$)

$$x^j = \frac{1}{L_{\text{bin}}} \sum_{i=(j-1)L_{\text{bin}}+1}^{jL_{\text{bin}}} x_i$$

Binning technique (block analysis)



- The average over the bins is equal to the original average (linearity of the average)

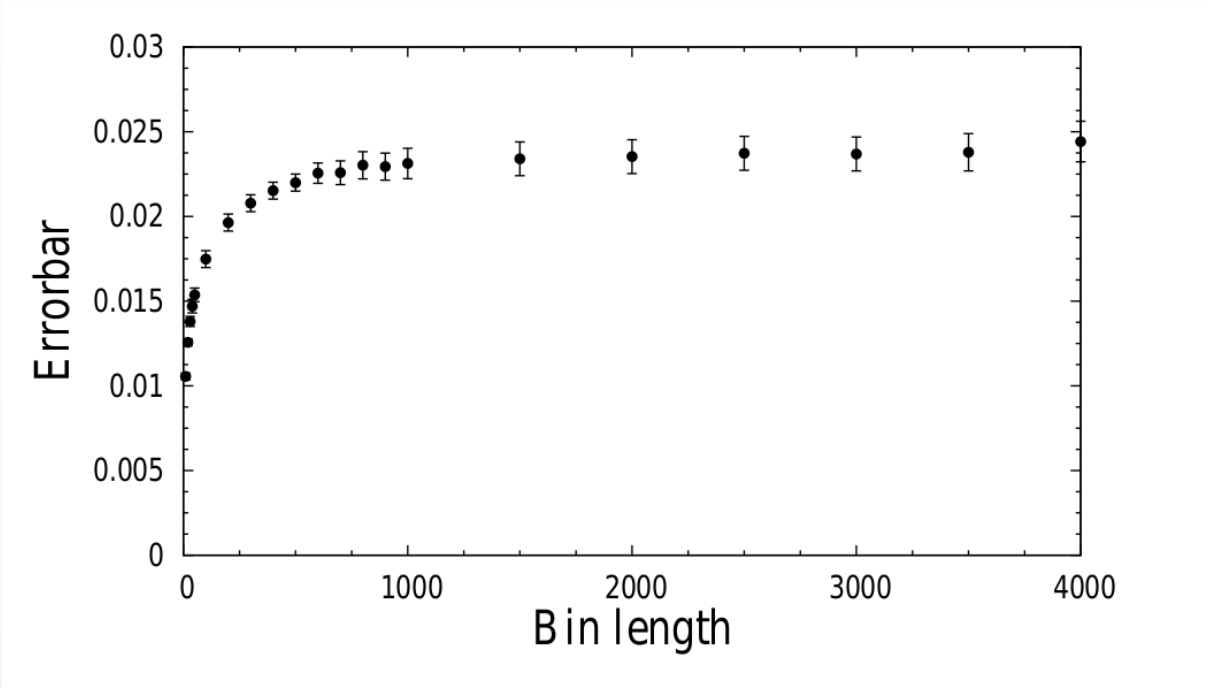
$$\bar{x^j} = \frac{1}{N_{\text{bin}}} \sum_{j=1}^{N_{\text{bin}}} x^j = \bar{x}$$

- If L_{bin} is large enough, the bin averages x^j are independent random variables

$$s^2_{\text{bin}} = \frac{1}{N_{\text{bin}}} \sum_{j=1}^{N_{\text{bin}}} (x^j - \bar{x})^2$$

- The error bar on the average value is given by $s_{\bar{x}} = s_{\text{bin}}/\sqrt{N_{\text{bin}}}$

Binning technique (block analysis)



Error bar on the average $s_{\bar{x}}$ as a function of L_{bin} along a Markov chain

(keeping the length of the Markov chain N fixed)

Variational wave functions for the single-orbital Hubbard model

Luca F. Tocchio

Politecnico di Torino

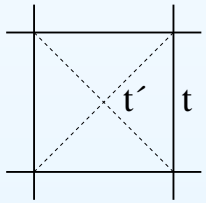
Variational MC: superconductivity and magnetism

Jastrow-Slater wave function: $|\Psi\rangle = \mathcal{J}|\Phi_0\rangle$

$|\Phi_0\rangle$ is the ground state of a non-interacting Hamiltonian $\mathcal{H}_{\text{aux}}$
It contains the relevant physical ingredients!

$\mathcal{J}$ is a Jastrow factor to include correlation effects

$$\mathcal{H}_{\text{aux}} = \mathcal{H}_0 + \mathcal{H}_{\text{BCS}} + \mathcal{H}_{\text{Néel}}$$



$$\mathcal{H}_0 = \sum_{\boldsymbol{k},\sigma} [-2t(\cos k_x + \cos k_y) - 4t' \cos k_x \cos k_y - \mu] c^\dagger_{\boldsymbol{k},\sigma} c_{\boldsymbol{k},\sigma}$$

$$\text{H}_{\text{BCS}} = \sum_{\boldsymbol{k}} 2(\Delta_x \cos k_x - \Delta_y \cos k_y) c^\dagger_{\boldsymbol{k},\uparrow} c^\dagger_{-\boldsymbol{k},\downarrow} + \text{H.c.} \quad \text{(superconductivity with } d\text{-wave pairing)}$$

$$\mathcal{H}_{\text{Néel}} = \Delta_{\text{AF}} \sum_R (-1)^{R_x+R_y} \left(c^\dagger_{R,\uparrow} c_{R,\uparrow} - c^\dagger_{R,\downarrow} c_{R,\downarrow} \right) \quad \text{(Néel antiferromagnetism)}$$

Gros, PRB 38, 931 (1988); Yokoyama and Shiba, JPSJ 57, 2482 (1988)

The variational wave functions: Jastrow factor

$$|\Psi\rangle = \mathcal{J}|\Phi_0\rangle$$

$$\mathcal{J} = \exp[-\frac{1}{2}\sum_{i,j} v_{ij} n_i n_j] \; ; v_{ij} > 0$$

- $\mathcal{J}$ Long-range density-density correlator

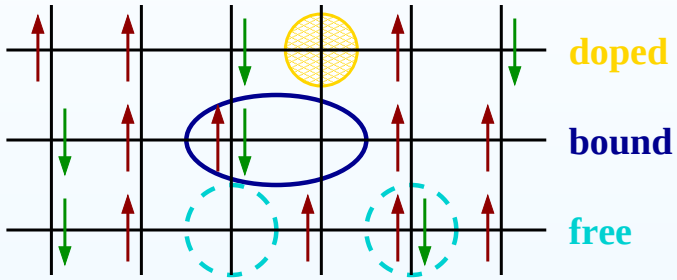
  v_{ij} variational parameters independently optimized for each distance $|i-j|$
- The local term v_{ii} accounts for the local Coulomb repulsion
- The optimal v_{ij} are maximal at short distance and decrease to zero when $|i-j|$ increases
- Long range density-density correlations are crucial to describe strongly correlated physics (see next)

The Jastrow factor $\mathcal{J}$

At half filling, pure Mott insulator, with no antiferromagnetism, above a finite value of U/t

M. Capello *et al.*, PRL 94, 026406 (2005)

Mott-Hubbard insulator: bounded empty and doubly occupied sites



Effective holon-doublon attraction and holon-holon (doublon-doublon) repulsion

$$\mathcal{J} = \exp[-\frac{1}{2}\sum_{i,j} v_{ij} n_i n_j] \; ; v_{ij} > 0 \text{ and maximal at short distance}$$

$$n_i n_j = D_i D_j + H_i H_j - H_i D_j - D_i H_j + n_i + n_j - 1$$

$$\begin{cases} n_i = n_{i,\uparrow} + n_{i,\downarrow} \\ D_i = n_{i,\uparrow} n_{i,\downarrow} \\ H_i = (1 - n_{i,\uparrow})(1 - n_{i,\downarrow}) \end{cases}$$

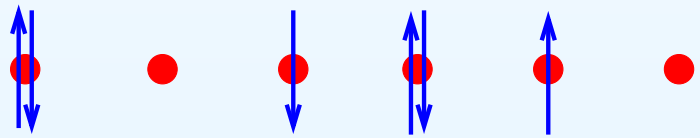
A simpler variational state: Gutzwiller factor

Suppress the weight of configurations with double occupations

Gutzwiller wf $$|\Psi_G\rangle = \mathcal{P}_G|\Phi_0\rangle = \exp[-g\sum_i n_{i,\uparrow}n_{i,\downarrow}]|\Phi_0\rangle; g > 0$$

M.C. Gutzwiller PRL 10, 159 (1963)

Define the basis $\{|x\rangle\}$ of configurations of electrons in the lattice. i.e., $|\bar{x}\rangle$:

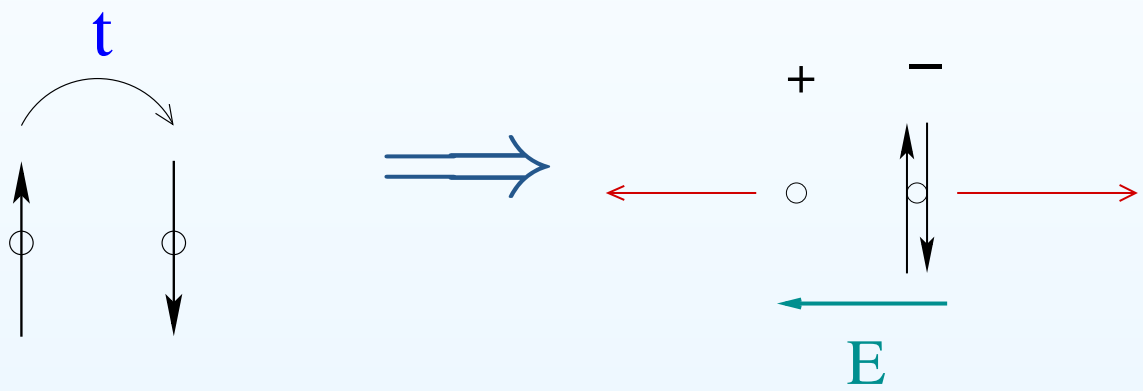


$$\langle x|\Psi_G\rangle = \mathcal{P}_G(x)\langle x|\Phi_0\rangle$$

$\mathcal{P}_G(x) \leq 1$ is a number, smaller when the number of doubly occupied sites in $|x\rangle$ is larger: the amplitude of the non-interacting state $\langle x|\Phi_0\rangle$ is renormalized by $\mathcal{P}_G(x)$

The Gutzwiller factor

At half filling, empty sites (Holons, H) and doubly occupied sites (Doublons, D) play a crucial role for the conduction



H and D must be correlated otherwise E induces a current

The Gutzwiller wf cannot describe a Mott insulator:
The formation of H and D pairs is penalized, but possible once g is finite, and there is no correlation among H and D

We have only a trivial insulator with $g = \infty$

Backflow correlations

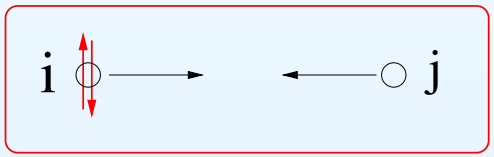
R.P. Feynman and M. Cohen, PR 102, 1189 (1956)

Y. Kwon, D.M. Ceperley, and R.M. Martin, PRB 48, 12037 (1993); M. Holzmann *et al.*, PRE 68, 046707 (2003)

Tocchio, Becca, Parola, Sorella PRB 78, 041101(R) (2008); Tocchio, Becca, Gros PRB 83, 195138 (2011)

$\Longrightarrow$ Many-body effects inside the mean-field variational wave functions $|\Phi_0\rangle$

$$\phi_k^b(r_{i,\sigma}) = \tilde{\epsilon}\phi_k(r_{i,\sigma}) + \eta_1 \sum_{j \text{n.n.} i} D_i H_j \phi_k(r_{j,\sigma}) + \dots$$



ϕ_k single-particle orbital

$D_i = n_{i,\uparrow} n_{i,\downarrow}$ $H_i = h_{i,\uparrow} h_{i,\downarrow}$

$h_{i,\sigma} = (1 - n_{i,\sigma})$

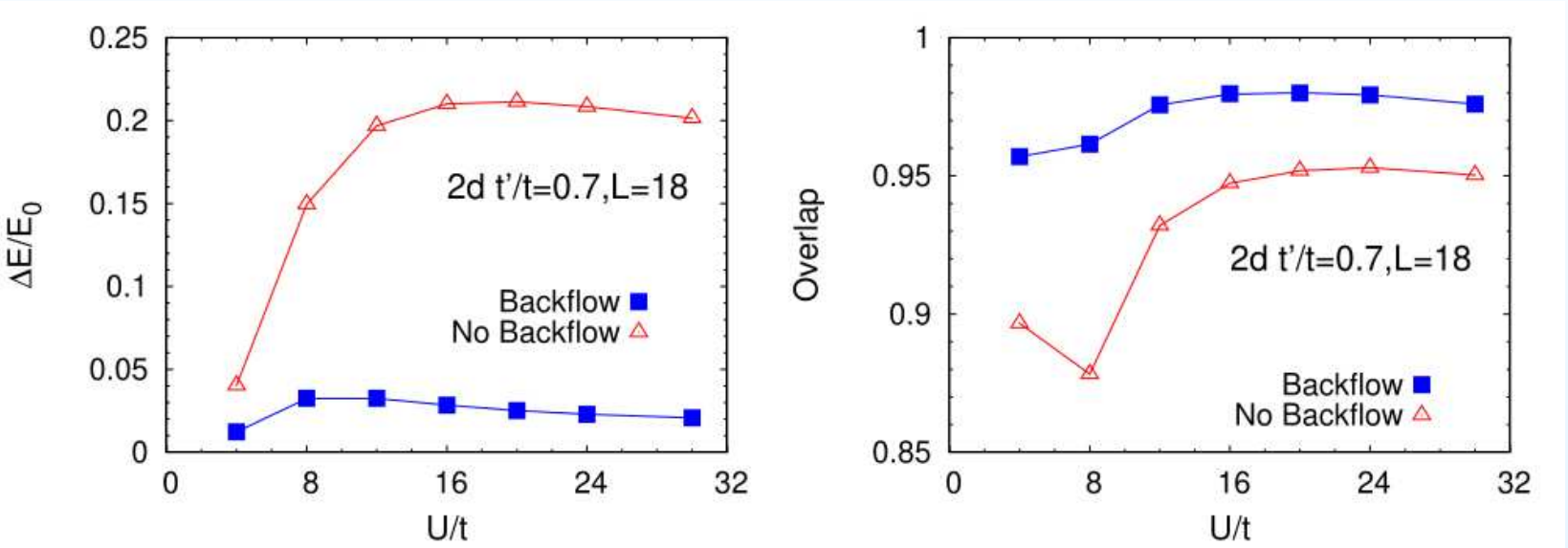
Summing up Slater determinants, according to the many-body configuration in the lattice:

The determinant part of the wave function contains correlation

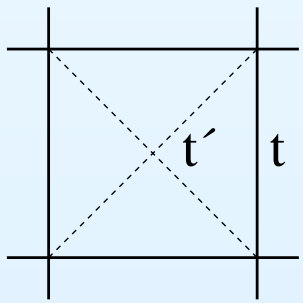
Backflow correlations: accuracy of the variational state

Comparison with exact diagonalization (Lanczos) at half filling

Overlap $\langle \Psi_{\text{var}} | \Psi_0 \rangle$ $\Delta E = E_{var} - E_0$; E_0 exact energy



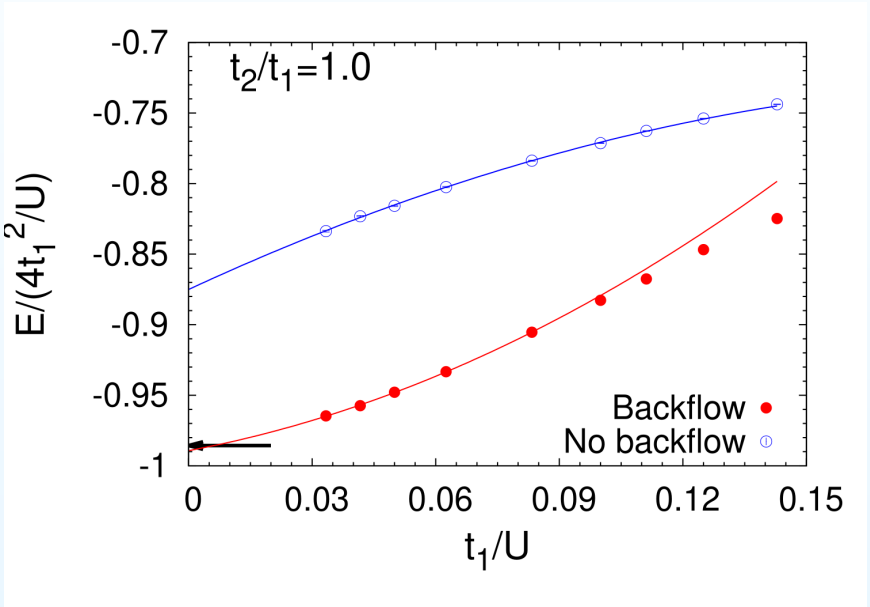
- Energy improvement in the insulating state ($U/t \gtrsim 8$)
- Overlap improvement at the metal-insulator transition



Tocchio, Parola, Becca, Sorella PRB 78, 041101(R) (2008)

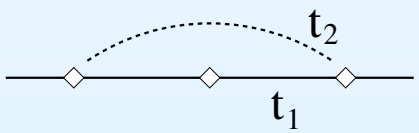
Backflow correlations: accuracy of the variational state

Large U limit in one dimension for the variational energy



- Proper extrapolation to the Heisenberg limit
- Black arrow is Density-matrix renormalization group (DMRG) - Exact in 1D
- Backflow correlations are the starting point for backflow-based neural network states

Luo and Clark, PRL 122,226401 (2019)



Tocchio, Becca, Gros PRB 83, 195138 (2011)

Variational Monte Carlo II

Luca F. Tocchio

Politecnico di Torino

Variational Monte Carlo

We need to compute:

- Ratio of probabilities for the Metroplis algorithm

$$\frac{\mathcal{P}(x')}{\mathcal{P}(x)} = \left| \frac{\langle x'|\Psi\rangle}{\langle x|\Psi\rangle} \right|^2$$

- The local estimator, i.e., for the energy:

$$e_L(x) = \frac{\langle x|\mathcal{H}|\Psi\rangle}{\langle x|\Psi\rangle} = \sum_{x'} \langle x|\mathcal{H}|x'\rangle \frac{\langle x'|\Psi\rangle}{\langle x|\Psi\rangle}$$

a) Is $\langle x|\mathcal{H}|x'\rangle$ hard to compute? No, thanks to the locality of the Hamiltonian: A given $|x\rangle$ is connected only to states that differ for the hopping of one electron from one site to one of its neighbors. Given $|x\rangle$, the non-zero matrix elements of $\langle x|\mathcal{H}|x'\rangle$ are $O(L)$ (# sites L × # bonds ×2 for the spin)

Variational Monte Carlo

b) What about the ratio $\langle x'|\Psi\rangle/\langle x|\Psi\rangle$?

- Let us consider the usual wave function $|\Psi\rangle = \mathcal{J}|\Phi_0\rangle$ ($\langle x|\Psi\rangle = \mathcal{J}(x)\langle x|\Phi_0\rangle$), where $|\Phi_0\rangle$ is the ground state of a quadratic Hamiltonian and $\mathcal{J}$ is the Jastrow factor:

$$\mathcal{J} = \exp(-\frac{1}{2}\sum_{i,j} v_{i,j} n_i n_j)$$

- $\langle x|\Phi_0\rangle$ is a determinant. Using standard linear algebra formulas it scales $O(L!)$. Using the LU decomposition into triangular matrices, it scales $O(L^3)$.
  Important message: Never do linear algebra by yourself, use always the libraries!
- It can be shown that $\langle x'|\Phi_0\rangle/\langle x|\Phi_0\rangle$ scales $O(L^2)$ if $|x'\rangle$ is close to $|x\rangle$

Jastrow ratio computation

- $\mathcal{J}(x) = \exp(-\frac{1}{2}\sum_{i,j} v_{i,j} n_i(x) n_j(x))$ scales $O(L^2)$
- When $|x'\rangle$ is close to $|x\rangle$, it is possible to define a fast computation of $\mathcal{J}(x')/\mathcal{J}(x)$ that is $O(L)$:

$$|x\rangle \equiv |n_1, \ldots, n_k, \ldots, n_l, \ldots, n_L\rangle \qquad |x'\rangle \equiv |n_1, \ldots, n_{k+1}, \ldots, n_{l-1}, \ldots, n_L\rangle$$

$$\frac{\mathcal{J}(x')}{\mathcal{J}(x)} = \frac{\mathcal{J}(n_1, \ldots, n_{k+1}, \ldots, n_{l-1}, \ldots, n_L)}{\mathcal{J}(n_1, \ldots, n_k, \ldots, n_l, \ldots, n_L)} = \frac{\exp(-\sum_i v_{k,i} n_i)}{\exp(-\sum_i v_{l,i} n_i)} e^{v_{k,l} - v_{k,k}}$$

Using $v_{i,j} = v_{j,i}$ and $v_{k,k} = v_{l,l}$ (translational invariance)

- The ratio $\frac{\mathcal{J}(x')}{\mathcal{J}(x)}$ costs $O(1)$ defining $T_{\text{Jastrow}}(j) = \sum_i v_{j,i} n_i$
- Updating $T_{\text{Jastrow}}(j)$ to the new configuration $|x'\rangle$ costs $O(L)$:

$$n'_i = n_i + \delta_{k,i} - \delta_{l,i} \Rightarrow T'_{\text{Jastrow}}(j) = T_{\text{Jastrow}}(j) + v_{k,j} - v_{l,j}$$

- From time to time, recompute $T_{\text{Jastrow}}(j)$ from scratch to avoid rounding errors

VMC summary

- We can compute expectation values of operators with a Monte Carlo sampling

$$\langle \mathcal{O} \rangle = \sum_x \mathcal{P}(x) \mathcal{O}_L(x) \quad \mathcal{O}_L(x) = \frac{\langle x|\mathcal{O}|\Psi\rangle}{\langle x|\Psi\rangle} \quad \mathcal{P}(x) = \frac{|\langle \Psi|x\rangle|^2}{\sum_{x'} |\langle \Psi|x\rangle|^2}$$

- Electronic states $|x_n\rangle$ generated with the Metropolis algorithm $\langle \mathcal{O} \rangle \sim \frac{1}{N} \sum_{n=1}^{N} \mathcal{O}_L(x_n)$
- Binning technique to uncorrelate measures
- Fast computation of ratios of determinants and Jastrow factors

$$\frac{\langle x'|\Psi\rangle}{\langle x|\Psi\rangle} = \frac{\mathcal{J}(x')\langle x'|\Phi_0\rangle}{\mathcal{J}(x)\langle x|\Phi_0\rangle}$$

Optimization of variational parameters

- We have learned how to compute expectation values over a given wave function
- The Jastrow-Slater wave function depends on many variational parameters that can be optimized to minimize the energy

$$|\Psi\rangle = \mathcal{J}|\Phi_0\rangle$$

$$\mathcal{J} = \exp[-\frac{1}{2}\sum_{i,j} v_{ij} n_i n_j] \; ; v_{ij} > 0$$

- We cannot perform loops over each parameter! We need a method to optimize a large number of parameters at the same time: Stochastic reconfiguration
- More efficient than steepest descent
- Still may have problems with local minima

Stochastic reconfiguration

- Starting state: $|\Psi^0\rangle = |\Psi(\{p_k^0\})\rangle$ with $\{p_k^0\}$ initial set of variational parameters
- Updated state: $|\Psi'\rangle = |\Psi(\{p_k'\})\rangle$ with $\{p_k'\}$ new set of variational parameters

For each variational parameter p_k^0, the update to p_k' is defined as:

$$\delta p_k = \sum_{j=1}^{N_p} S_{jk}^{-1} f_j \qquad p_k' = p_k^0 + \tau \delta p_k.$$

$$S_{jk} = \langle \Psi^0 | \mathcal{O}_j \mathcal{O}_k | \Psi^0 \rangle - \langle \Psi^0 | \mathcal{O}_j | \Psi^0 \rangle \langle \Psi^0 | \mathcal{O}_k | \Psi^0 \rangle$$

$$f_j = \langle \Psi^0 | \mathcal{O}_j | \Psi^0 \rangle \langle \Psi^0 | \mathcal{H} | \Psi^0 \rangle - \langle \Psi^0 | \mathcal{O}_j \mathcal{H} | \Psi^0 \rangle$$

$\mathcal{O}_k$ is the logarithmic derivative of $|\Psi^0\rangle$

$$\mathcal{O}_k(x) = \frac{1}{\Psi^0(x)} \frac{\partial \Psi^0(x)}{\partial p_k} = \frac{\partial \ln \langle x | \Psi^0 \rangle}{\partial p_k}$$

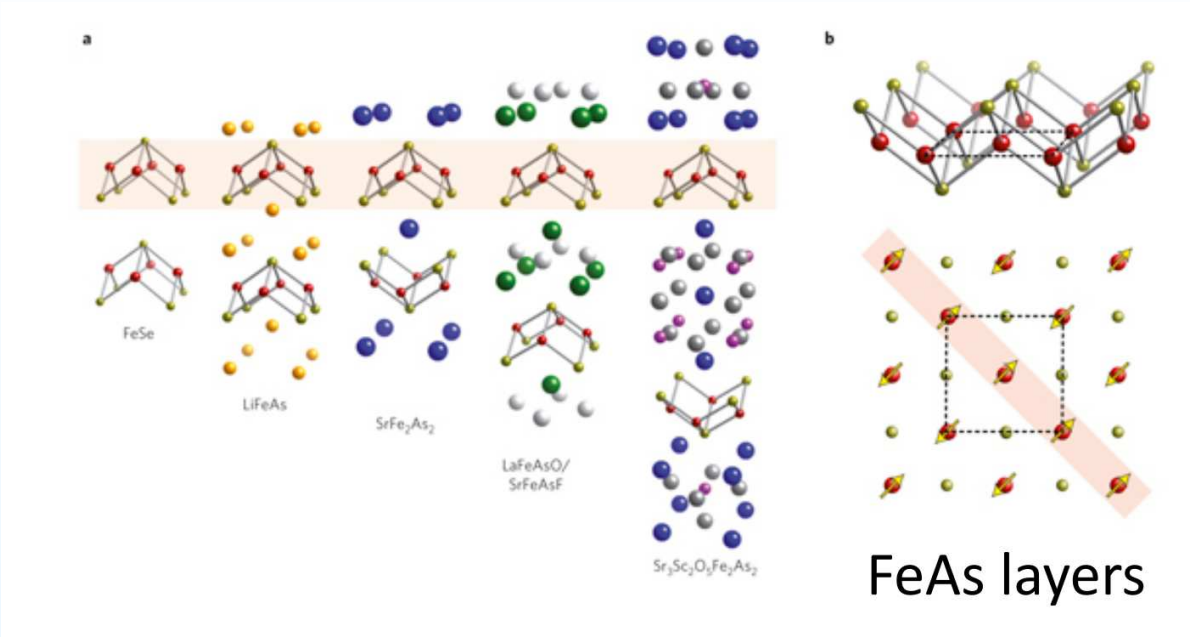
Main message: minimization through computing expectation values

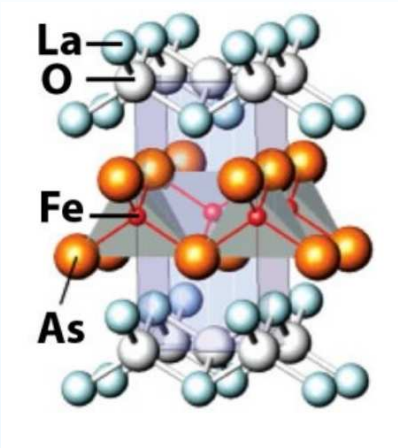
Multiorbital Hubbard model

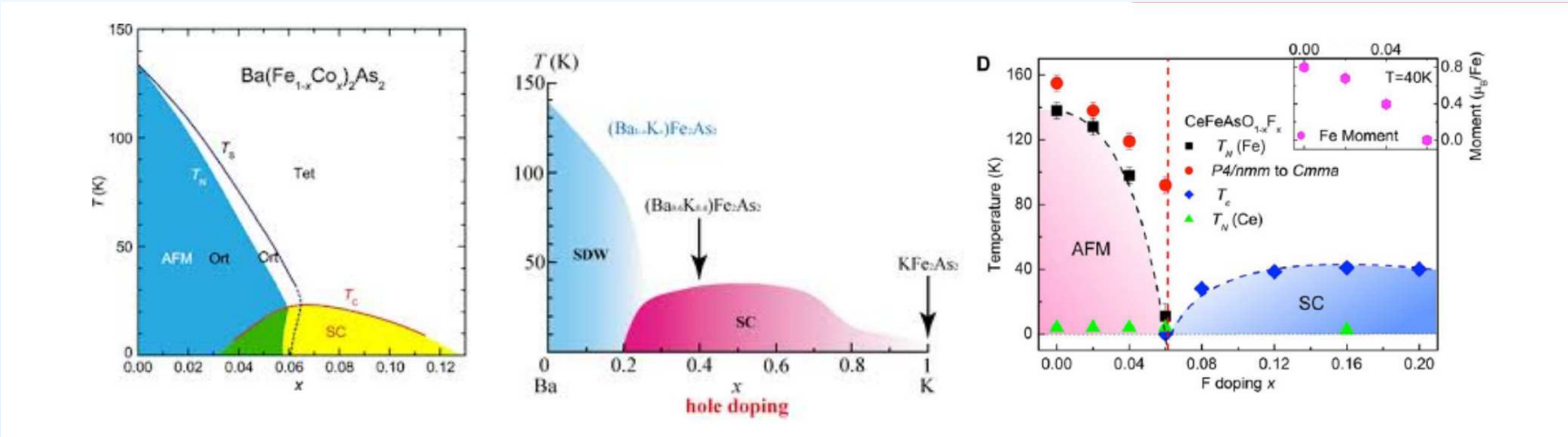
Luca F. Tocchio

Politecnico di Torino

The iron pnictide superconductors

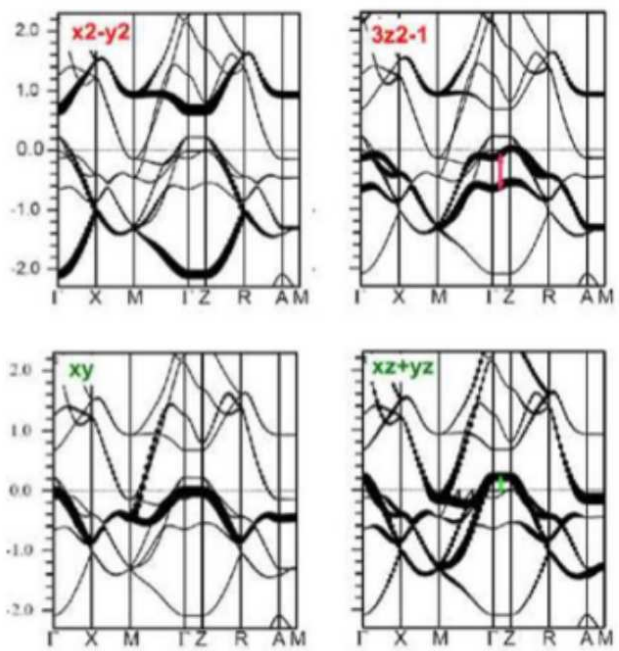






Paglione and Green, Nat. Phys. 6, 645 (2010); Stewart, Rev. Mod. Phys. 83, 1589 (2011)

The multiorbital physics



All the 5 d orbitals are close to the Fermi level. Multiorbital description is necessary

- Effective model with only the d_{xy}, d_{xz} and d_{yz} orbitals
  (light enough for numerics, but keeps main features of Iron-based superc.)
- Role of Hund's physics (correlation) and band structure

The 3-orbital Hubbard model $\mathcal{H}_k + \mathcal{H}_{\text{int}}$

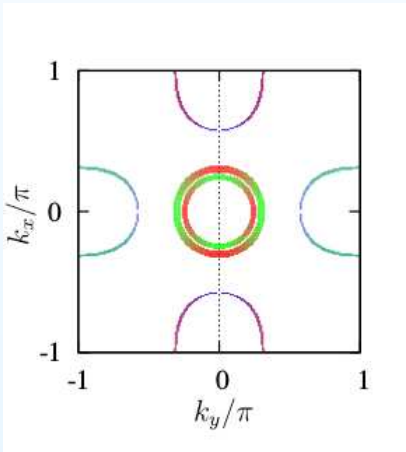
$$\mathcal{H}_k = \sum_k \sum_\sigma \sum_{\alpha,\beta} c^\dagger_{k\alpha\sigma} T_{\alpha\beta}(k) c_{k\beta\sigma}$$

α, β are orbital indices for the $1 = xz, 2 = yz, 3 = xy$ orbitals of Fe

$$T_{11}(k) = 2t_2 \cos k_x + 2t_1 \cos k_y + 4t_3 \cos k_x \cos k_y$$
$$T_{22}(k) = 2t_1 \cos k_x + 2t_2 \cos k_y + 4t_3 \cos k_x \cos k_y$$
$$T_{33}(k) = 2t_5(\cos k_x + \cos k_y) + 4t_6 \cos k_x \cos k_y + \Delta_{xy}$$
$$T_{12} = (T_{21})^* = 4t_4 \sin k_x \sin k_y$$
$$T_{13} = (T_{31})^* = 2it_7 \sin k_x + 4it_8 \sin k_x \cos k_y$$
$$T_{23} = (T_{32})^* = 2it_7 \sin k_y + 4it_8 \sin k_y \cos k_x$$

Orbital d_{xy} has a crystal field

Orbtials d_{xz} and d_{yz} exchange x and y directions



Fermi surface at $n = 4$ (parent compound in the three-orbital model)

Minimal model to reproduce qualitative features of Fermi surface

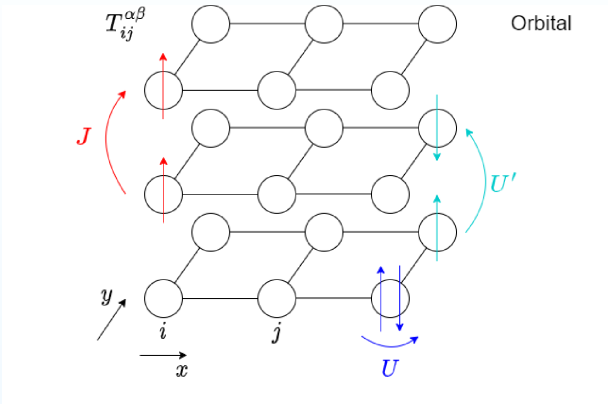
Daghofer *et al.*, PRB 81, 014511 (2010); Fanfarillo *et al.*, PRL 125, 177001 (2020).

The 3-orbital Hubbard model $\mathcal{H}_k + \mathcal{H}_{\text{int}}$

$$\mathcal{H}_{\text{int}} = U \sum_R \sum_\alpha n_{R\alpha\uparrow} n_{R\alpha\downarrow} + U' \sum_R \sum_{\alpha \neq \beta} n_{R\alpha\uparrow} n_{R\beta\downarrow} +$$

$$+ (U' - J) \sum_R \sum_{\alpha > \beta} \sum_\sigma n_{R\alpha\sigma} n_{R\beta\sigma} +$$

$$- J \sum_R \sum_{\alpha \neq \beta} c^\dagger_{R\alpha\uparrow} c_{R\alpha\downarrow} c^\dagger_{R\beta\downarrow} c_{R\beta\uparrow} + J \sum_R \sum_{\alpha \neq \beta} c^\dagger_{R\alpha\uparrow} c^\dagger_{R\alpha\downarrow} c_{R\beta\downarrow} c_{R\beta\uparrow}$$



- Intra-orbital Coulomb repulsion U
- Inter-orbital Coulomb repulsion U'
- Hund's coupling J (including spin-flip and pair-hopping terms)

Rotational invariance $U' = U - 2J$

Georges, de' Medici, Mravlje, Annu. Rev. Condens. Matter Phys. 4, 137 (2013)

The variational wave function (Jastrow factors)

$$|\Psi\rangle = \mathcal{J}_U \mathcal{J}_{U'} \mathcal{J}_J |\phi_0\rangle.$$

$|\phi_0\rangle$ is the ground state of a non-interacting Hamiltonian $\mathcal{H}_{\text{BCS}}$
$\mathcal{J}$ is a Jastrow factor to include correlation effects:

Capello, Becca, Fabrizio, Sorella, and Tosatti, PRL 94, 026406 (2005)

Generalization to a multiorbital model

$$\mathcal{J}_U = \exp\Big(-\frac{1}{2}\sum_{\alpha}\sum_{R,R'} v^{\alpha,\alpha}_{R,R'} n_{R\alpha} n_{R'\alpha}\Big)$$ Long-range intra-orbital density-density correlator

$$\mathcal{J}_{U'} = \exp\Big(-\frac{1}{2}\sum_{\alpha\neq\beta}\sum_{R,R'} v^{\alpha,\beta}_{R,R'} n_{R\alpha} n_{R'\beta}\Big)$$ Long-range inter-orbital density-density correlator

$$\mathcal{J}_J = \exp\Big(-\frac{1}{2}\sum_{\alpha\neq\beta} u^{\alpha,\beta}\sum_{R} S^z_{R\alpha} S^z_{R\beta}\Big)$$ On-site inter-orbital spin-spin correlator: Hund's coupling

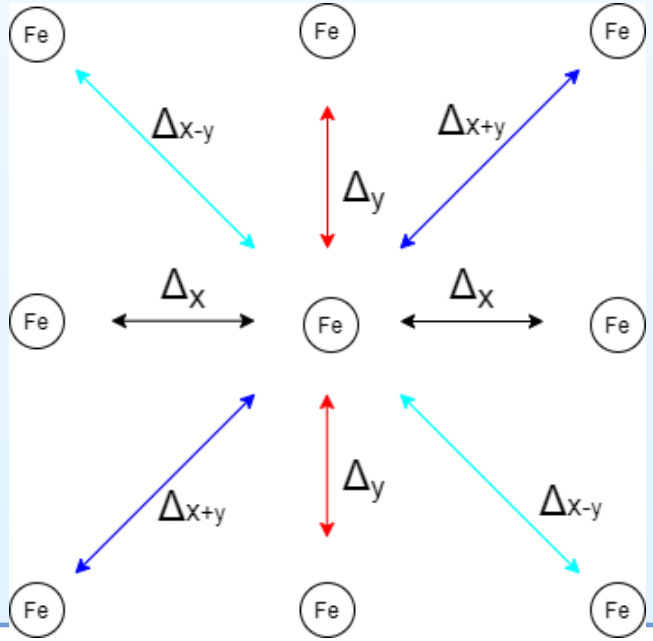
Marino, Scazzola, Becca, Capone, Tocchio, PRL 134, 196502 (2025)

The variational wf ($|\phi_0\rangle$ as ground state of $\mathcal{H}_{\text{BCS}}$)

$$\mathcal{H}_{\text{BCS}} = \mathcal{H}_k - \sum_{R,\alpha,\sigma} \mu_\alpha c^\dagger_{R\alpha\sigma} c_{R\alpha\sigma} + \sum_{R,\alpha,\delta} \left[\Delta_{\alpha,\delta} \left(c_{R,\alpha,\downarrow} c_{R+\delta,\alpha,\uparrow} - c_{R,\alpha,\uparrow} c_{R+\delta,\alpha,\downarrow} \right) + \text{h.c.} \right]$$

$$(\delta = \pm x, \pm y, \pm(x+y), \pm(x-y))$$

- The hopping term $\mathcal{H}_k$ is the same as in the Hubbard Hamiltonian
- Intraorbital pairing amplitudes with no a priori symmetry (optimized along the simulation): (no interorbital pairing Misawa and Imada, Nat Comm 5, 5738 (2014))



Correlation functions

Luca F. Tocchio

Politecnico di Torino

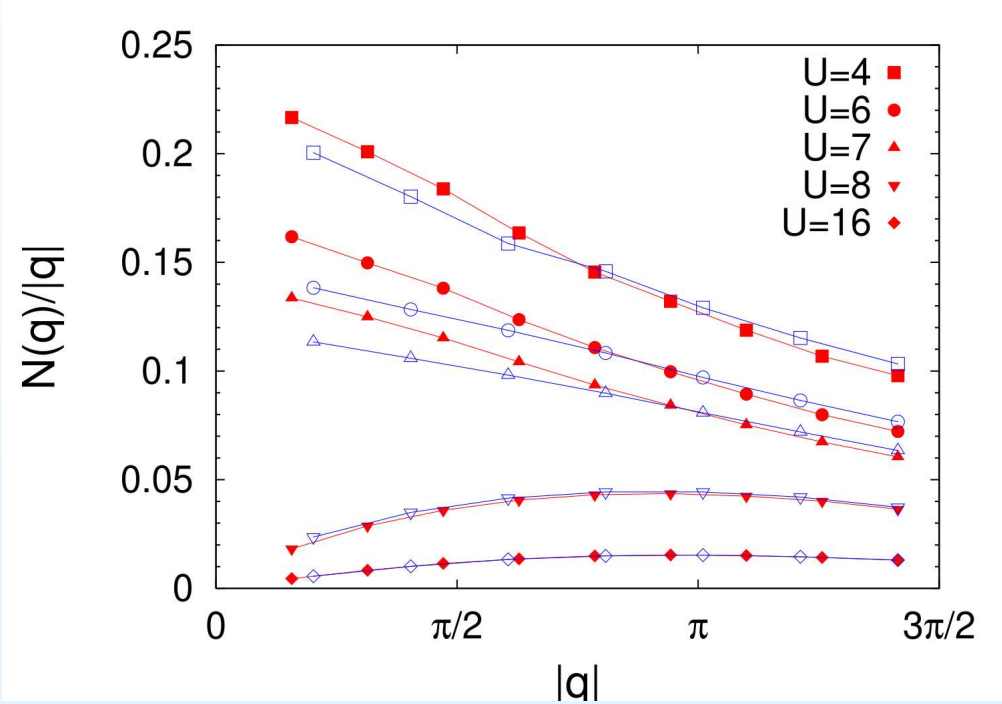
Metal or insulator?

Density-density correlations $N(\boldsymbol{q}) = \frac{1}{L}\sum_{R,R'}\langle n_R n_{R'}\rangle e^{i\boldsymbol{q}\cdot(\boldsymbol{R}-\boldsymbol{R}')}$

Excited state $|\Psi_{\boldsymbol{q}}\rangle = n_{\boldsymbol{q}}|\Psi\rangle \Rightarrow$ Charge gap $\propto \lim_{|\boldsymbol{q}|\to 0} \frac{\boldsymbol{q}^2}{N(\boldsymbol{q})}$

Metal: $N(\boldsymbol{q}) \propto |\boldsymbol{q}|$ for $|\boldsymbol{q}| \to 0$ Insulator: $N(\boldsymbol{q}) \propto |\boldsymbol{q}|^2$ for $|\boldsymbol{q}| \to 0$

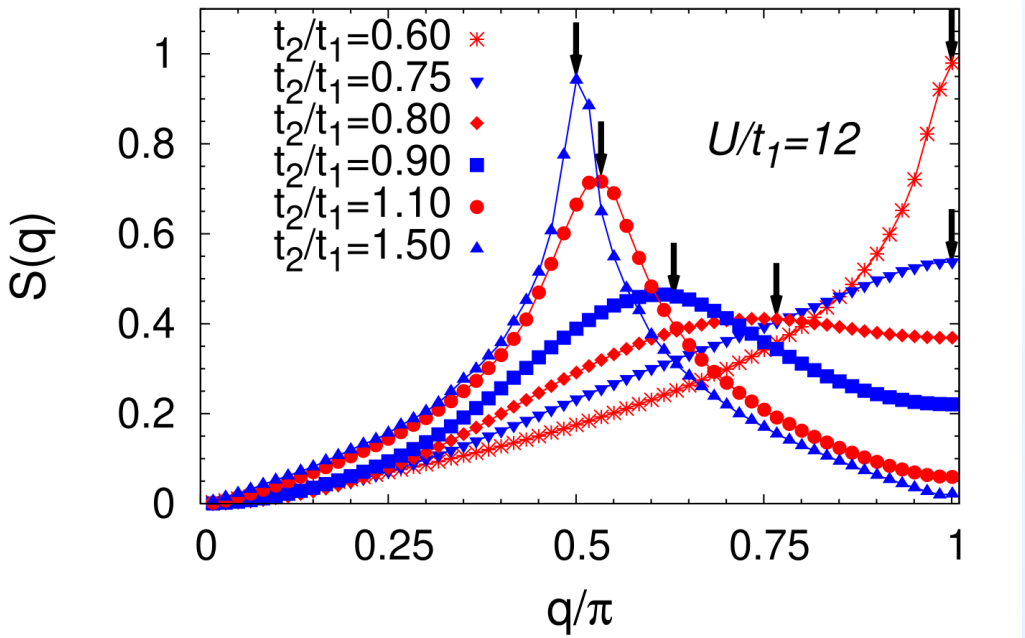
Feynman PR 94, 262 (1954); Tocchio, Becca, Gros, PRB 83, 195138 (2011)

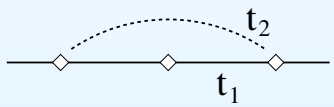


- Single-orbital Hubbard model on the square lattice with $t'/t = 0.75$
- Insulator for $U/t \geq 8$; Metal for $U/t \leq 7$ Tocchio, Parola, Becca, Sorella PRB 78, 041101(R) (2008)

Charge or magnetic order

- Spin-spin correlations $S(\boldsymbol{q}) = \frac{1}{L} \sum_{R,R'} \langle S^z_R S^z_{R'} \rangle e^{i\boldsymbol{q}\cdot(\boldsymbol{R}-\boldsymbol{R}')}$
- Density-density correlations $N(\boldsymbol{q}) = \frac{1}{L} \sum_{R,R'} \langle n_R n_{R'} \rangle e^{i\boldsymbol{q}\cdot(\boldsymbol{R}-\boldsymbol{R}')}$
- Long-range Order: Peak at $\boldsymbol{Q}$, diverging with the system size
- Short-range correlations: Peak at $\boldsymbol{Q}$, *non* diverging with the system size





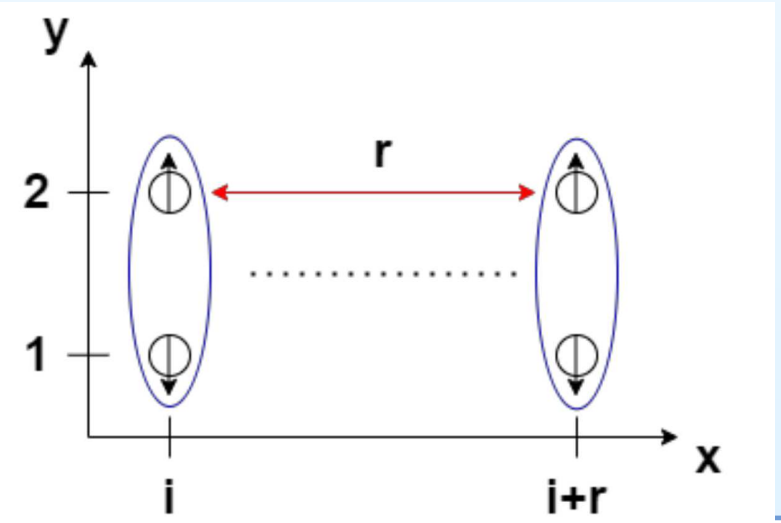
- Example for the 1D Hubbard model: No long-range order in 1D, only short-range spin-spin correlations Tocchio, Becca, Gros PRB 81, 205109 (2011)

Superconductive singlet-singlet correlations

- Finite pairing amplitudes $\Delta_{\alpha,\delta}$ in the variational wave function are necessary condition for superconductivity
- Superconductivity is present when singlet-singlet correlations are finite at large distances

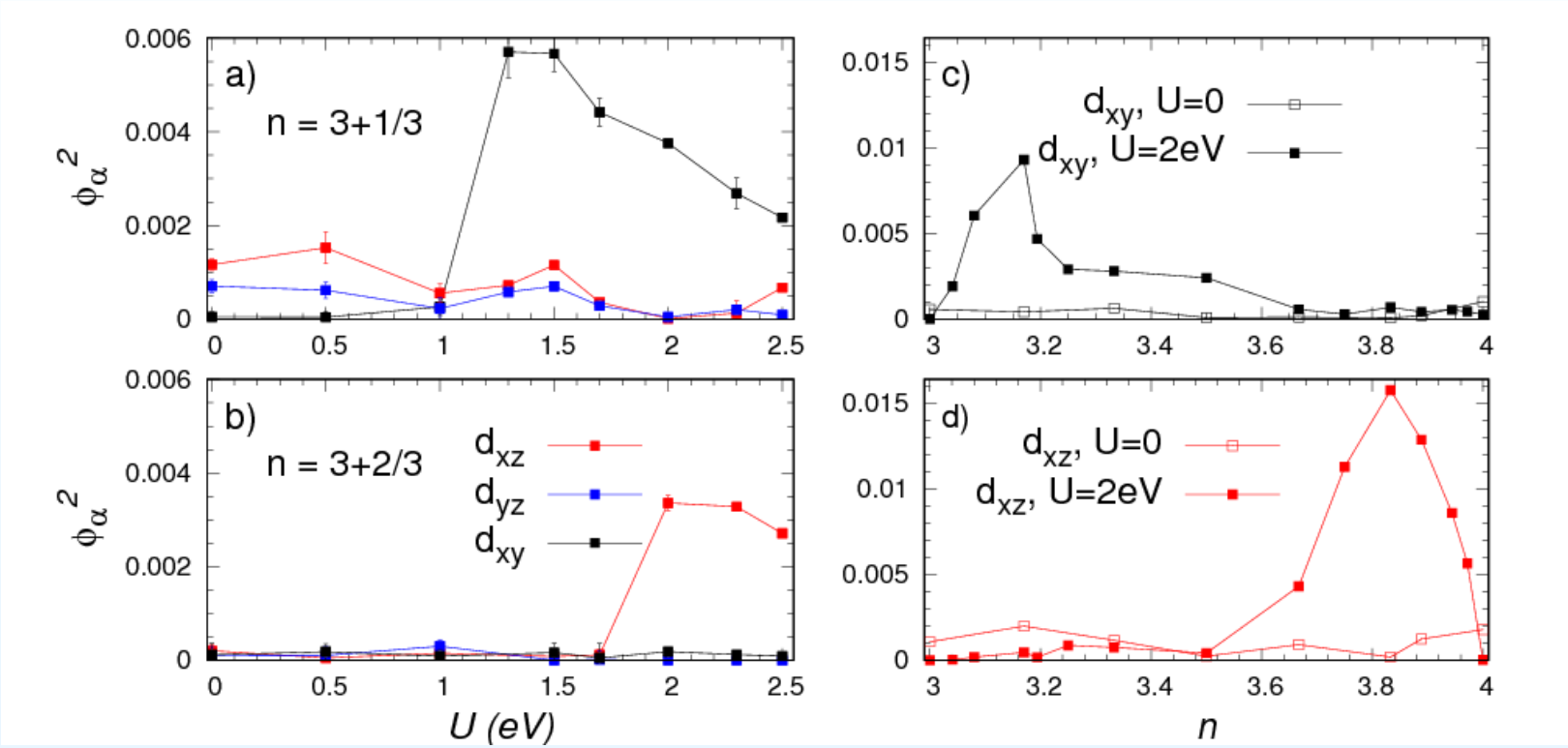
$$D_\alpha(r) = \frac{1}{L}\sum_R \langle P_{R,\alpha} P^\dagger_{R+rx,\alpha}\rangle \quad P_{R,\alpha} = c_{R+y,\alpha,\downarrow} c_{R,\alpha,\uparrow} - c_{R+y,\alpha,\uparrow} c_{R,\alpha,\downarrow}$$

Finite order parameter ϕ_α: $\lim_{r\to\infty} D_\alpha(r) = \phi_\alpha^2$ for superconductivity



Superconductivity in the multiorbital model

Marino, Scazzola, Becca, Capone, Tocchio, PRL 134, 196502 (2025)



- Superconductivity in one of the three orbitals, for U larger than a critical value d_{xy} for $3 < n \lesssim 3.5$; d_{xz} for $3.5 \lesssim n < 4$
- Compare with the finite-size effects at $U = 0$ (Role of U and J (not shown) in driving superconductivity)

Quo vadis VMC?

- Computation of physical observables over a variational approximation of the true ground state
- One path: Apply simple Jastrow-Slater states to multi-orbital Hubbard models where nonperturbative approaches are lacking

  Di Sante *et al.*, Kagome metals, RMP 98, 015002 (2026)
- Another path: Improve the variational states with neural-network quantum states based on the transformer architecture (at the moment limited to single-orbital Hubbard models on simple lattice geometries)

  Viteritti *et al.*, arXiv:2604.21978