UNIVERSITÀ DEGLI STUDI DI MILANO

Facoltà di Scienze e Tecnologie
Corso di Laurea Magistrale in Fisica

Variational Monte Carlo study of the two-band Hubbard model

Relatore:
Dott. Federico BECCA
Relatore interno:
Prof. Nicola MANINI

Tesi di laurea di:
Federico ARRIGONI
Matr. n. 827306
Codice PACS: 71.10.Fd

Anno Accademico 2013-2014

Contents

Introduction

One of the earliest applications of Quantum Mechanics was the theory of non-interacting electrons in a periodic potential. In 1928 F. Bloch derived the general form of the wave functions of the eigenstates of a periodic Hamiltonian [1] and thus sparked the beginning of the Quantum theory of electrons in solids. The independent-electron approximation has been applied to predict properties such as the conductivity and the specific heat of many metals; moreover, in this framework it is straightforward to determine whether a material should be a metal or an insulator [2]. However, electrons in a solid can hardly be seen as independent particles and the Coulomb interaction must be taken into account at least approximately. If this interaction can be considered a small perturbation, then one can think that the properties of the interacting system resemble those of the non-interacting one, for instance the concept of Fermi surface is still valid in the presence of the interaction; this statement is justified by the hypothesis that the non-interacting spectrum is adiabatically connected with the interacting one. Then, the interaction is treated perturbatively, starting from the non-interacting case. This is the rationale behind the Landau theory of Fermi liquids, which was phenomenologically developed by L.D. Landau [3] and subsequently formalized by A.A. Abrikosov and I.M. Khalatnikov in the language of Quantum Field Theory.

At the beginnings of Quantum theory of electrons in solids, the powerful capability of determining whether a solid should be a metal or an insulator soon turned out to be one of the greatest shortcomings of the independent-electron model: it was pointed out that some compounds, such as NiO, expected to be metals within that model, have insulating behavior [4]. In 1949 N.F. Mott [5] pointed out that for materials where conduction is determined by

electrons in $3d$ orbitals, the Heitler-London scheme should be the correct picture rather than Bloch theory: this is due to the fact that $3d$ atomic orbitals (as well as $4d$, $4f$,$5d$ and $5f$) are well-localized in space and describing the system in terms of a localized basis set is more consistent than using extended Bloch states. Under the effect of Coulomb repulsion, electrons prefer to localize into a so-called *Mott insulator*. Unlike a band insulator, such a state is driven by strong interactions and cannot be adiabatically connected to the ground-state of a free Hamiltonian: it is an example of a strongly-correlated electron system. A striking phenomenon due to strong correlations is that certain materials with metallic character undergo a Mott transition (*i.e.* they become Mott insulators) if subject to a decrease in pressure. The paradigmatic model that displays a Mott transition is the Hubbard model, introduced in 1963 M.C. Gutzwiller [6], J. Hubbard [7] and J. Kanamori [8]. Since then, a number of approximate solutions and numerical studies have been made, yet our understanding of the physics of the Hubbard model is not complete.

Fermi liquid theory can take into account even the phenomenon of superconductivity, which manifests itself most noticeably with a dramatic drop of electrical resistivity below a critical temperature. The discovery of superconducting materials dates back to 1911, but it took almost 50 years before J. Bardeen, L. Cooper and J.R. Schrieffer worked out the BCS theory of conventional superconductivity in 1957 [9]. In the presence of an attractive effective interaction, resulting from the exchange of virtual phonons, electron pairs with opposite momenta and spins condense into a coherent state, which is lower in energy than to the *normal* Fermi liquid state. BCS theory gives a microscopic explanation of the observed vanishing resistivity as well as of other phenomena typical of superconductors, such as the Meissner-Ochsenfeld effect, the isotopic effect and the exponentially vanishing specific heat as a function of temperature. Microscopically, the theory shows the existence of an energy gap for single-particle excitations of the ground-state. The predictions of BCS theory both on microscopic and macroscopic phenomena have found agreement with experiments for any superconducting material discovered up to 1978 [10]: this is the reason why superconductors whose behavior is explained by BCS theory are referred to as *conventional superconductors*. Elemental superconductors (Al, Hg, Sn,...), as well as many alloys (NbTi), fall in this family [11]. Noticeably, these materials all have low (*i.e.* below $77K$, the boiling point of N) critical temperatures, the highest being $39K$ for MgB_2 [12]: conventional superconductivity is essentially a low-temperature phenomenon.

In 1986, the discovery of high-temperature superconductivity in $Ba_xLa_{5-x}Cu_5O_{5(3-y)}$ by J.G. Bednorz and K.A. Müller [13] gave a twist to the understanding of superconductivity, although unconventional superconductors were already known before that date: heavy-fermion superconductors, discovered in 1979 [10], cannot be described by means of the BCS theory. One of the most relevant classes of such materials is that of layered-perovskite copper-oxide, the cuprates, which comprises the compound discovered by Bednorz and Müller. Among the other classes of unconventional superconductors we list the organic superconductors, the heavy-fermion compounds, the novel iron-pnictide superconductors and many others. Just

one year after Bednorz and Müller groundbreaking discovery, Anderson borrowed the concept of Resonating Valence Bond (RVB) from the field of quantum spin liquids and, from its basis, proposed a theory which sought to describe the mechanism for the superconductivity in the cuprates as a consequence of electronic interaction in the two-dimensional CuO_2 planes. Although promising, this RVB theory of a doped Mott insulator initially fell short of expectations. Later on, the theory was successfully employed to predict some of the most striking experimental evidence on cuprate superconductors, such as the existence of the *pseudogap* regime and the *d*-wave nature of the superconductive coupling [14, 15]. Although an appealing and influent concept, the RVB theory is not an accepted theory of superconductivity in cuprates, both because it has led to fallacious prediction in the past and due to the existence of competing theories. The ideas underlying the RVB picture were then reformulated in terms of a Renormalized Mean Field Theory, which follows as closely as possible the original BCS theory.

Superconductivity in the cuprates is often modeled in terms of a Hubbard model with three orbitals per unit cell out of half-filling, but it was shown by F.C. Zhang and T.M. Rice [16] that a single orbital can effectively describe the phenomenon. More recently, the study of more complex materials has revealed that the presence of more than one orbital is in fact required in order to describe certain physical phenomena, such as the existence of an Orbital Selective Mott Phase (OSMP) in some ruthenate compounds [17], the prominence of spin-orbit interaction in the iridate compounds [18] or the superconductivity in the iron pnictides [19]. From the theoretical point of view, these discoveries have sparked interest in a more general class of strongly-correlated systems, the *multi-band* (or *multi-orbital*) Hubbard models. Increasing efforts in studying these models have been made, but much remains to be explored and understood. Indeed, the great complexity and variety of these models, which are matched by the multitude of materials described in terms of one of them, require a great amount of work.

Nowadays, a large portion of this work can be accomplished by tackling the problem numerically. Thanks to the huge boost given to computer resources during the last decades, and thus the opportunity of giving predictive and accurate results, numerical simulations of physical systems have acquired a prominent importance in Condensed Matter Physics, complementing theory and experiment. One of the most widely used and resourceful methods for strongly-correlated systems is Dynamical Mean Field Theory (DMFT), which gives a mean-field treatment of spatial fluctuations, while local time fluctuations are taken into account completely. Density Functional Theory can also take into account strong correlations by means of the DFT+U scheme, where the Hubbard U term contributes to the energy functional. A class of numerical methods relies upon the variational principle in order to approximate the exact ground-state with guess states: this class includes the Density Matrix Renormalization Group (DMRG) and its latest developments and the Variational Monte Carlo (VMC) method. Both these methods can accurately describe spatial correlation functions, which are of great physical importance. In particular, the VMC method computes expectation values by sampling

the Hilbert space by means of a stochastic process.

The purpose of this thesis is to employ the VMC method to study the two-band Hubbard model and to gain insight on its properties, most importantly the existence of the OSMP. To this purpose, a projected BCS wave function, closely related to the RVB theory, is employed. Another issue addressed in the present work is whether superconductivity can be enhanced in the two-band model compared to the single-band case. This work also stands as an introductory step for an application of the VMC method to other, possibly more realistic, models: these include systems with non-trivial topological structure and multi-band models with more than two bands, which are relevant in the mentioned iron superconductors or iridium oxides. An immediate continuation of the present work involves studying whether spin-triplet superconductivity (observed for instance in Sr_2RuO_4) can be naturally described by means of VMC or not.

Outline

This thesis is organized as follows:

- Chapter 1 introduces strong correlations and the Hubbard model, briefly summarizing the RVB theory of cuprate superconductors. Subsequently the main object of this thesis, the two-band Hubbard model, is described from a general point of view, along with its properties and relevance.
- Chapter 2 is a general survey of the VMC method: the basis of the method and the rationale behind it are discussed; moreover, an optimization scheme for wave functions, the Stochastic Reconfiguration, is introduced.
- Chapter 3 describes how the VMC method applies to a model of strongly-correlated electrons on a lattice and specifically to the two-band model of interest.
- Chapter 4 resumes the results obtained during this work. The phase diagram at half-filling both in absence and in presence of Hund's coupling is shown. Then, superconducting properties of the model out of half-filling are addressed.

Acknowledgements

I am very grateful to Federico for accepting to act as supervisor of this thesis and at the same time introducing me to the field of strongly-correlated electron systems. Luca played a major role in the development of the present work, being always available to help and contributing to the results here displayed. Their contribution to this work is unvaluable.

I also want to thank Prof. Manini for kindly helping during the writing of this thesis, especially for exposing countless bugs in the draft versions.

The present thesis largely blossomed at the International School of Advanced Studies (ISAS/SISSA), thanks to a research training fellowship for undergraduate students.

1
Hubbard models

This chapter gives a general overview of the physics of the single-band and two-band Hubbard models. First, we discuss how Mott insulators are described in terms of the one-band Hubbard model, taking into account the symmetries and limiting cases of the model. This is useful in order to get acquainted with the typical issues of a strongly-correlated system. The strong coupling limit of the Hubbard model is also addressed, due to its relevance to the Resonating Valence Bond theory of high-temperature superconductivity. A detailed description of the two-band Hubbard model follows: its model Hamiltonian is derived from an *ab initio* Hamiltonian; subsequently the model is described from a general point of view. We then focus on the compound alloy series $Ca_{2-x}Sr_xRuO_4$, which is strictly related to the model Hamiltonian of the two-band model, and we conclude with the description of the Orbital Selective Mott Phase.

1.1 The Hubbard model and Mott insulators

Fermi liquids are effectively described by a picture of non-interacting quasi-particles. The Fermi liquid picture gives a good understanding of the conducting properties of numerous materials, where the electron-electron interaction leads only to a renormalization of the band structure. On the other hand, when correlations become dominant, such an effective picture may fail. This is the case of several materials, for instance the transition-metal oxides, where valence electrons are subject to Coulomb interactions which dominate over the kinetic energy, a feature typical of well-localized d or f orbitals. The Fermi liquid picture does not accu-

rately predict the properties of such systems: for instance, systems with an odd number of electrons per unit cell cannot be insulators according to band theory, in contrast to what is observed in many transition-metal oxides. This failure reflects the overall difficulty of the class of strongly-correlated electron systems. However, a wide range of interesting and exotic physical phenomena, for instance high-temperature superconductivity, arise as a consequence of correlations; sometimes such phenomena are concealed by more trivial ones, for instance antiferromagnetic ordering.

Mott transitions in transition-metal oxides
with the ABO_3 perovskite crystal structure

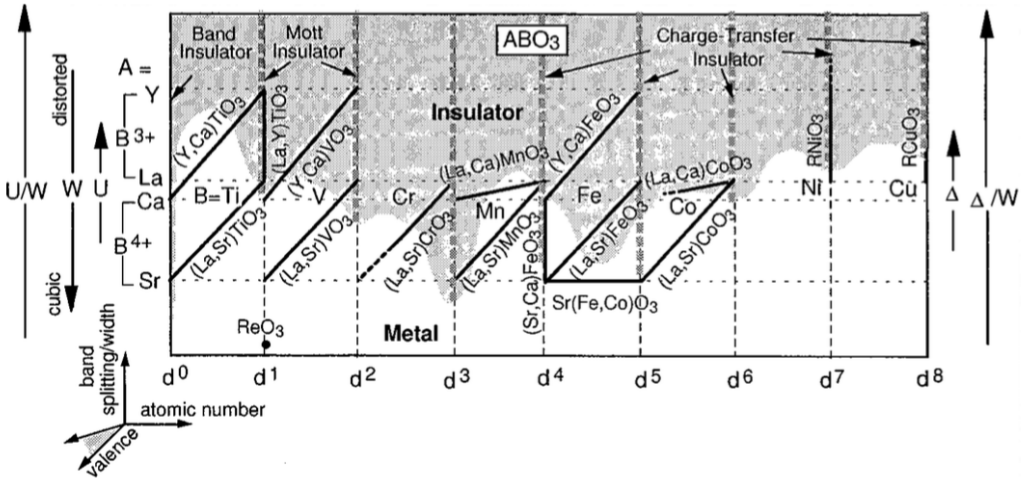


Figure 1.1: A graph summarizing metals and insulators with the perovskite crystal structure ABO_3, after reference [20]. On the horizontal line the group of the transition metal B and on the vertical axis either the bandwidth or the splitting.

The Mott transition is one of the most striking examples of the effects of strong correlations: it consists in a metal-to-insulator transition driven by the strong electron-electron interaction. Metal-insulator transitions can arise in band theory as well, but it turns out that band theory does not give the correct picture of the Mott insulating mechanism. In his pioneering work [5], Mott argued that in the case of some transition-metal compounds, where atomic orbitals are well-localized, the correct picture would be the Heitler-London one, while band theory would be deficient in the description. Consider a D-dimensional hyper-cubic lattice of atoms with one atomic s orbital, for instance Na atoms, and lattice spacing a; suppose each atom contributes with 1 valence electron, so that there are as many electrons as lattice sites L. If electrons were considered as non-interacting particles, the system would be conducting for any finite value of a: even if the atoms were separated by large distances, the tight-binding

s-band would still be half-filled, despite its bandwidth being an exponentially small function of a. There is clearly something wrong in this argument, because, when the interatomic distance is large enough, the system is expected to behave as an array of neutral, isolated Na atoms, which does not conduct. Evidently, the fallacy in this argument lies in neglecting the electron-electron repulsion: therefore, it is necessary to take it into account to some extent. The largest contribution of the electron-electron repulsion arises when one atomic orbital is doubly occupied. If this contribution is considered, charge fluctuations induced by the kinetic terms are suppressed because of their higher energy cost, at least when the band is narrow enough: increasing the lattice spacing above some critical value would then make the system insulating. Although the situation described above is oversimplified, it captures the essence of the physics of many transition-metal oxides, as depicted in figure 1.1.

The physical situation described above (sometimes referred to as Mott's *Gedankenexperiment*) is represented by the single-band Hubbard model Hamiltonian [6, 7, 8]:

$$H_{Hub} = -t \sum_{\langle i,j \rangle} \sum_{\sigma} \left(c^{\dagger}_{i,\sigma} c_{i,\sigma} + \text{h.c.} \right) + U \sum_{i} n_{i,\uparrow} n_{i,\downarrow} \,, \tag{1.1}$$

where $c_{i,\sigma}$ are fermionic annihilation operators at site i and with z-axis spin projection σ, satisfying canonical anti-commutation relations:

$$\left\{ c_{i,\sigma} \,, c^{\dagger}_{j,\sigma'} \right\} = \delta_{i,j} \, \delta_{\sigma,\sigma'} \,; \tag{1.2a}$$

$$\left\{ c_{i,\sigma} \,, c_{j,\sigma'} \right\} = 0 \,; \tag{1.2b}$$

where the curly braces indicate the anticommutator $\{A, B\} \equiv AB + BA$, while $n_{i,\sigma} \equiv c^{\dagger}_{i,\sigma} c_{i,\sigma}$ are the occupation number operators; $t > 0$ is the hopping matrix element and $U > 0$ is the on-site Coulomb repulsion; the notation $\langle i, j \rangle$ restricts the sum to nearest neighboring sites i and j: in general, one can define the Hubbard model by taking into account further neighbor hopping amplitudes (and off-site Coulomb interactions as well), but for simplicity we restrict only to nearest-neighbor ones. The local Fock space is 4-dimensional: in the occupation number basis, each site can either be empty $|0\rangle$, occupied by one spin-$\uparrow$ electron $|\uparrow\rangle = c^{\dagger}_{i,\uparrow} |0\rangle$, occupied by one spin-$\downarrow$ electron $|\downarrow\rangle = c^{\dagger}_{i,\downarrow} |0\rangle$ or doubly occupied $|d\rangle = c^{\dagger}_{i,\uparrow} c^{\dagger}_{i,\downarrow} |0\rangle$: doubly occupied sites are defined with the spin-$\uparrow$ operators on the left; the opposite convention differs by a change of sign in $|d\rangle$. The system is said to be at *half-filling* if the total number of electrons $N = \sum_{i,\sigma} n_{i,\sigma}$ equals the number of sites L.

The Hubbard model is probably the simplest strongly-correlated system of electrons[1]. It is a model Hamiltonian which implements the competing tendencies of the kinetic energy and Coulomb interaction in the most elementary way possible. The phase diagram of the Hubbard model depends on the total number of electrons and the ratio between the couplings U/t; therefore at half-filling, which is the case described above, the phase diagram depends only on the parameter U/t. Consider the limiting behaviors of the model at half-filling: for $U/t =$

[1] The even simpler Bose-Hubbard model has no spin degree of freedom, but does not apply to electrons in solids.

0 the non-interacting picture is restored and band theory correctly describes a conductor, whereas when $U/t = \infty$, electrons are localized at the lattice sites and the atomic picture is correct. It is clear that for some critical value of U/t a metal-insulator transition must occur. In the special case of one dimension ($D = 1$) the critical value of U/t was found by E.H. Lieb and F.Y Wu [21] to be exactly 0, whereas in higher dimensions the critical value is finite[2].

Symmetries play a significant role in Physics and the Hubbard model is no stranger to this fact. The simplest symmetry of the Hubbard model is a global $U(1)$-charge symmetry: multiplying each annihilation operator by the same phase factor gives a canonical transformation which leaves the Hamiltonian invariant; this is nothing but charge conservation, in other words the total number of electrons commutes with the Hamiltonian and is thus conserved. Applying a global $SU(2)$-spin transformation to the annihilation operators clearly leaves the kinetic term invariant. In order to see that the interaction term is invariant as well, it is useful to define the spin operator at site i:

$$\boldsymbol{s}_i \equiv \frac{1}{2} \sum_{\rho,\tau} c^{\dagger}_{i,\rho} \, \boldsymbol{\sigma}_{\rho\tau} \, c_{i,\tau} \,, \tag{1.3}$$

where $\boldsymbol{\sigma}$ is the vector of Pauli matrices (in the following m_i will denote the z-component of the spin operator). It then turns out that the interacting part of the Hamiltonian of equation (1.1) can be written as the sum of a term proportional to $-U\,(\boldsymbol{s}_i)^2$ plus a constant: consequently, the Hubbard model has a global spin-rotational invariance, not surprisingly, since no preferential axis is specified in the Hamiltonian. On bipartite lattices at half-filling the Hubbard model also has a particle-hole symmetry.

1.2 Superconductivity in the Hubbard model

It was mentioned that high-temperature superconductivity is a consequence of strong correlations between the electrons. This situation is completely different from the case of conventional superconductivity, which originates from the effective electron attraction mediated by phonons[3]. The existence of a different mechanism, which is supported by many experimentally observed phenomena (some of which are briefly presented here), calls for a theory of unconventional superconductors. In the case of cuprates, the Resonating Valence Bond theory developed by P.W. Anderson aims at explaining how superconductivity is a consequence of strong correlations.

[2] In the case of a two-dimensional square lattice at half-filling the system has an antiferromagnetic instability for all values of $U > 0$ due to the so-called *perfect nesting* of the Fermi surface, a feature which favors the insulating state.

[3] Phonons can play a role in enhancing high-temperature superconductivity, but they are not the main cause of its settling.

1.2.1 Superconductivity in the cuprates

Soon after the discovery of superconductivity in the cuprates, there was clear evidence of differences in relation to conventional BCS superconductors. As already mentioned, the cuprate superconductors have a layered-perovskite structure (see figure 1.6a) and most of them have an insulating antiferromagnetic parent compound which becomes superconducting upon doping with holes or electrons. It is now accepted that cuprate superconductors are doped Mott insulators, as the phase diagram in figure 1.2 displays. From this point on, the focus will be on the hole-doped case. As soon as the parent compound is doped with holes to few percent, superconductivity arises below a critical temperature, which displays a characteristic *dome* as a function of the doping percentage. Superconductivity takes place inside the CuO_2 planes, while the interlayer regions act as charge reservoirs. This behavior is clearly different from the picture of a conventional superconductor: indeed, the difference lies in the mechanism which induces the superconducting behavior. It is now well-accepted that strong correlations generated by the electron-electron repulsion are the cause of superconductivity in the cuprates.

Besides the relationship of the cuprate superconductors with their antiferromagnetic Mott insulating parent compounds, there is much more evidence for an unconventional mechanism: first of all, the symmetry of the superconducting order parameter does not have an s-wave symmetry, but rather a $d_{x^2-y^2}$ symmetry. This symmetry has been made manifest by various experimental techniques, such as phase-sensitive Josephson tunneling [15], which revealed the nodal points and the sign change upon $\pi/2$ rotations of the order parameter.

Another feature of the phase diagram of cuprate compounds is the *pseudogap* regime, a phase in the underdoped region which displays a gapped excitation spectrum only in certain regions of the Fermi surface, but no superconductivity. Angle resolved photoemission spectroscopy (ARPES) experiments have probed the existence of such a pseudogap, at the same time unveiling the d-wave symmetry of the order parameter. In the superconducting phase (below T_c), the gap has a strong directional dependence and a node along the $\Gamma - M$ direction; the node broadens into an arc while moving above T_c in the pseudogap regime and eventually becomes a fully developed Fermi surface above T^* [23]. The nature of this pseudogap is still a matter of discussion nowadays.

1.2.2 Strong coupling in the Hubbard model: the $t - J$ model

When the Coulomb interaction prevails over the kinetic energy, the Hubbard model at half-filling is a Mott insulator. It turns out that the strong coupling limit of the Hubbard model has its own relevance. The recipe to obtain the strong coupling Hamiltonian is applying a unitary transformation e^{iA} to the Hubbard Hamiltonian [24]:

$$e^{-iA}\, H_{Hub}\, e^{iA}; \tag{1.4}$$

this unitary transformation is then required to give an effective Hamiltonian for states of the low-energy subspace. The effective Hamiltonian is called the $t - J$ model Hamiltonian, which

General phase diagram of a hole-doped cuprate superconductor

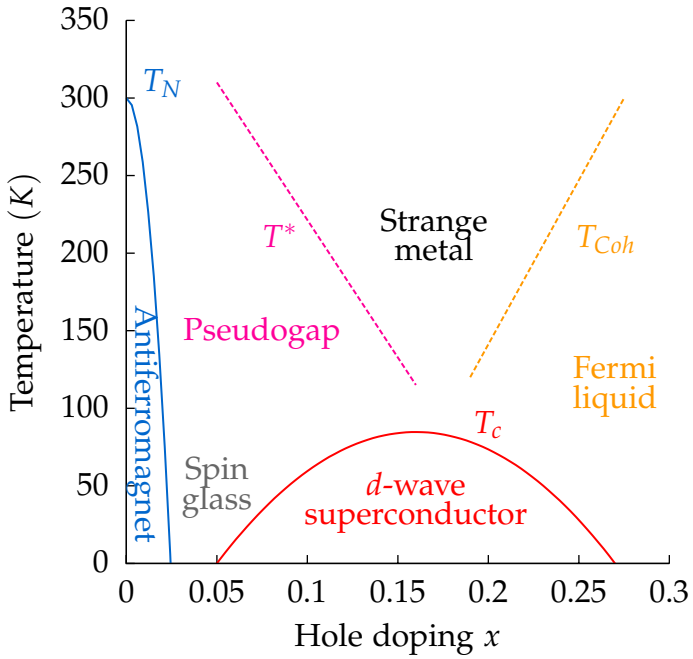


Figure 1.2: A semi-quantitative temperature-hole doping phase diagram for a typical cuprate superconductor, after reference [22]. The parent compound at $x = 0$ is an antiferromagnetic insulator, with critical Néel temperature T_N of about $300K$. When doped with holes, magnetic order disappears quite soon, leaving place to the spin-gapped *pseudogap* phase. The pseudogap phase is then replaced by the actual superconductor when a sufficiently high hole concentration is reached, provided that the temperature is lower than the critical temperature T_c. The temperature of onset of the pseudogap regime T^* can either enclose the superconducting dome (as predicted by RVB theory) or end at a quantum critical point at $T = 0$ inside the superconducting dome. In the over-doped regime, superconductivity is lost and the cuprates behave as a three-dimensional Landau-Fermi liquid. Above the pseudogap and superconducting phases there is a metallic phase dubbed *strange* metal, due to its non-Fermi liquid behavior. Charge degrees of freedom become coherent by decreasing the temperature below T_{Coh} in the over-doped regime, establishing the Fermi liquid state.

to leading order in $4t^2/U$ has the form:

$$\begin{aligned} H_{t-J} =& -t \sum_{\langle i,j \rangle} \sum_{\sigma} \left[\left(1 - n_{i,-\sigma}\right) c^{\dagger}_{i,\sigma} c_{j,\sigma} \left(1 - n_{j,-\sigma}\right) + \text{h.c.} \right] + \\ &+ \frac{4t^2}{U} \sum_{\langle i,j \rangle} \left(\boldsymbol{s}_i \cdot \boldsymbol{s}_j - \frac{n_i n_j}{4} \right) + \text{3-site terms} = \qquad (1.5) \\ =& -t \sum_{\langle i,j \rangle} \sum_{\sigma} \left[\mathcal{P}_G c^{\dagger}_{i,\sigma} c_{j,\sigma} \mathcal{P}_G + \text{h.c.} \right] + \\ &+ J_{ex} \sum_{\langle i,j \rangle} \left(\boldsymbol{s}_i \cdot \boldsymbol{s}_j - \frac{n_i n_j}{4} \right) + \text{3-site terms;} \qquad (1.6) \end{aligned}$$

where the antiferromagnetic exchange coupling $J_{ex} \equiv 4t^2/U > 0$ has been defined. The second form of the $t - J$ Hamiltonian is expressed in terms of the Gutzwiller projector:

$$\mathcal{P}_G \equiv \prod_i \left(1 - n_{i,\uparrow} n_{i,\downarrow}\right), \qquad (1.7)$$

whose role is to eliminate all states with doubly occupied sites. Usually, three-site terms are dropped in defining the model.

The exchange coupling of the $t - J$ model is positive and thus favors the antiferromagnetic alignment of neighboring spins. The half-filled limit of the $t - J$ model is nothing but the antiferromagnetic Heisenberg model, whose Hamiltonian is:

$$H_{Heis} = J_{ex} \sum_{\langle i,j \rangle} \boldsymbol{s}_i \cdot \boldsymbol{s}_j \,. \qquad (1.8)$$

Note that this model is the strong coupling limit of the half-filled Hubbard model. The ground-state in the atomic limit as a huge 2^L degeneracy, since each site is occupied by one electron, which can either be in a spin-$\uparrow$ or a spin-$\downarrow$ state. Turning on a small kinetic term lifts this degeneracy (to order t^2/U): indeed, an electron can gain kinetic energy by hopping from one site to a nearest-neighboring one and back to the original one, but only if nearest-neighboring electrons have opposite spin (otherwise this process would be forbidden by Pauli principle): this process is effectively represented by H_{Heis}. This result is satisfactory as most of the known Mott insulators exhibit antiferromagnetic order.

The $t - J$ model was shown by F.C. Zhang and T.M. Rice [16] to be a low-energy effective model for electrons in a three-band model describing oxygen and copper atoms in the CuO_2 planes of cuprate compounds (the superexchange coupling J_{ex} has a more involved form, in terms of the parameters of the 3-band Hubbard model). Therefore, the $t - J$ Hamiltonian is a candidate for modeling cuprate superconductors, but the mechanism of onset of superconductivity is in fact described by the Resonating Valence Bond theory.

1.2.3 Resonating Valence Bond theory

The concept of a Resonating Valence Bond (RVB) dates back to L.C. Pauling, who developed it in the framework of the Heitler-London picture of some molecules, such as benzene. The idea was revived by P.W. Anderson and P. Fazekas [25] in the field of quantum spin liquids and proposed as a variational ground-state for the $S = 1/2$ antiferromagnetic Heisenberg model. The RVB state is a linear superposition of states, each of them consisting of an arrangement of spin singlet dimers. To be more specific, consider two nearest-neighbor spins of the antiferromagnetic Heisenberg model: due to the antiferromagnetic exchange, they are likely, in terms of energy, to form a spin singlet (see figure 1.3b); then, one can think of arranging all the spins of the lattice into pairs, with each pair forming a singlet on its own: this is a *valence bond* state (see figure 1.3a). Of course, such a state has a huge degeneracy and each of the possible ways of forming the singlets will be equally probable. The RVB state is then constructed as a superposition of all of them (hence the term *resonating*).

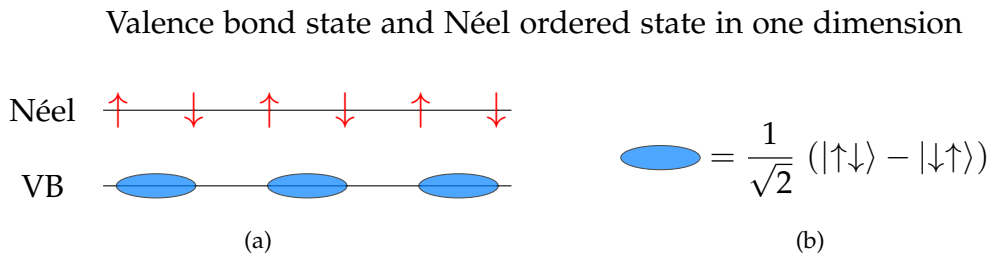


Figure 1.3: (a): Comparison of a Néel antiferromagnetic state and a singlet dimer state; the singlet dimer state has a lower energy than the Néel state. Indeed, it is one of the (highest weight) contributions in the true ground-state of the antiferromagnetic Heisenberg chain. (b): Explicit expression of one singlet dimer.

Indeed, this RVB state will be competitive in energy with the more usual ordered Néel antiferromagnet. Consider a D-dimensional spin-1/2 Heisenberg model on a hyper-cubic lattice: the Néel state with alternating spin-$\uparrow$ and spin-$\downarrow$ has an energy density of $-DJ/4$, while a single state of singlet dimers, has an average energy density of $-3J/8$: for $D = 1$ (see figure 1.3a) the valence bond state has a lower energy. As a matter of fact, the exact ground-state is a superposition of singlet bonds (not limited to nearest neighbor) with energy density $(-\ln 2 + 1/4)\, J$ [26]. In the two-dimensional Heisenberg model, the Néel state has a slightly better energy density of $-J/2$, while one singlet dimer state remains at $-3J/8$ and in higher dimensions the Néel state becomes more and more favorable. However, frustration can penalize the antiferromagnetic state, thus making the RVB state more favorable in energy.

The RVB theory of cuprate superconductivity, can be essentially seen as the theory of a hole-doped Mott insulator. Anderson proposed that the same RVB state that is relevant

to quantum spin liquids would lie close in energy to the Néel state for stoichiometric (*i.e.* undoped) cuprates [27]. Upon doping with a small percentage of holes, the RVB state becomes lower in energy than the Néel state, which is frustrated by hole hopping; eventually, holes will start to behave coherently and establish a superconducting order parameter (see figure 1.4). At the same time, an underlying idea of the theory is that the Mott transition in the Hubbard model exists regardless of the presence of antiferromagnetic order. Indeed, it is seen that a paramagnetic RVB state can describe the transition from a metal to a Mott insulator in the Hubbard model (see for instance subsection 4.1.1). These two facts about the RVB state suggest that it captures the features of the cuprate superconductors and that the Hubbard model, or rather its strong coupling limit the $t - J$ model, might be the correct effective model to describe them.

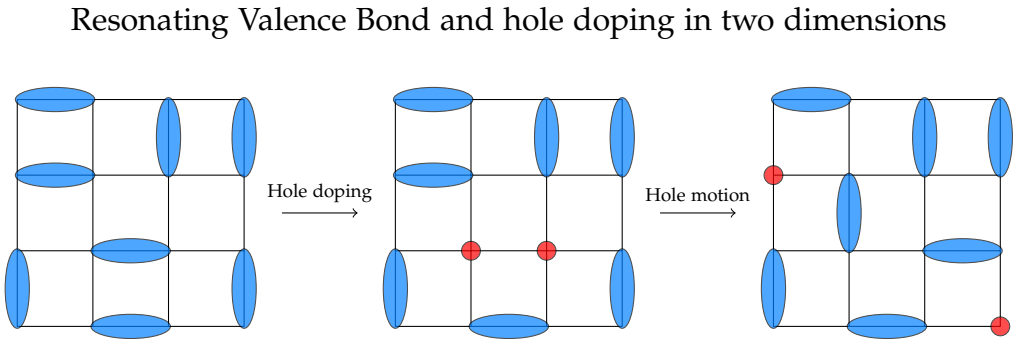


Figure 1.4: A snapshot of (*i.e.* one contribution to) the RVB state on a two-dimensional square lattice; only singlet dimers between nearest-neighboring sites are shown, according to the fact that they have the highest weight in the RVB superposition. Upon doping, some of the valence bonds are broken and the excitations, termed *holons*, are supposed to have a relatively easy way to move around in the lattice, because the singlet dimers can fluctuate and rearrange into different states of the RVB superposition. Eventually holes will behave coherently, taking advantage of the pairing and will contribute to superconductivity. Adapted from reference [28].

This is the underlying concept of RVB theory of cuprate superconductors. Unfortunately, as already realized by Anderson himself, RVB states are difficult to be handled analytically. Nonetheless, he suggested that they could be represented as:

$$|\Phi_{RVB}\rangle = \mathcal{P}_G \prod_k \left(u_k + v_k c^\dagger_{k,\uparrow} c^\dagger_{k,\downarrow}\right) |0\rangle \tag{1.9}$$

that is, projecting a BCS-like wave function by acting on it with the Gutzwiller projector. G. Kotliar and J. Liu represented this state by means of an auxiliary-boson representation and proved that a state with $d_{x^2-y^2}$-symmetric pairing provides the minimum Landau-Ginzburg energy functional [29], a result confirmed by numerical Variational Monte Carlo calculations

[30] and later experimentally confirmed; this is perhaps the most remarkable prediction of the theory. The pseudogap regime finds an easy explanation within the RVB picture: at a temperature scale T^* (see figure 1.2), singlets of RVB state are formed, but only when the temperature is lowered below T_c these preformed pairs experience coherent behavior, thus promoting superconductivity [31]. Besides these predictions, the RVB theory was shown to give a consistent account of a large number of phenomena observed in cuprate superconductors.

From the ideas of RVB, a Renormalized Mean Field Theory (RMFT) can be constructed. The RMFT follows a Hartree-Fock-BCS approximation of the wave function [32], but takes into account strong correlations by means of the Gutzwiller projector: from these grounds, a set of self-consistent equations for the BCS parameters ξ_k and Δ_k (which are functions of u_k and v_k) can be derived. At the same time, being founded on a guess state, the RVB lends itself to variational calculations. Unfortunately, both RMFT and numerical variational calculations allow only for an account of the $T = 0$ phase diagram and an RVB theory of T_c is not available yet [31].

1.3 The two-band Hubbard model

In the formulation of the Hubbard model, each lattice site was intended to host a single atomic s orbital. This picture is far from suitably describing transition-metal oxides and other compounds with strongly correlated electrons. The motivation is two-fold: first, in a compound, there usually is more than one significant orbital per unit cell; secondly, the symmetry of the orbitals may not be as trivial as in the s case, where there is full rotational symmetry. The description adopted here starts from an *ab-initio* Hamiltonian and thus can be applied to any model of interest, but the specific case of a two-band model is addressed. The basic properties of the model are discussed, as well as how such a Hamiltonian can describe transition-metal oxides.

1.3.1 Physical origin

Consider the second quantization Hamiltonian for a system of interacting electrons in a periodic potential[4]:

$$H = \sum_{\sigma} \int d^3\boldsymbol{x}\, \hat{\psi}^{\dagger}_{\sigma}(\boldsymbol{x}) \left(-\frac{\hbar^2}{2m_e} \boldsymbol{\nabla}^2 + \sum_i V(\boldsymbol{x} - \boldsymbol{R}_i) \right) \hat{\psi}_{\sigma}(\boldsymbol{x}) + \\ + \sum_{\sigma,\tau} \int d^3\boldsymbol{x}\, d^3\boldsymbol{y}\, \hat{\psi}^{\dagger}_{\sigma}(\boldsymbol{x})\, \hat{\psi}^{\dagger}_{\tau}(\boldsymbol{y})\, U(\boldsymbol{x} - \boldsymbol{y})\, \hat{\psi}_{\tau}(\boldsymbol{y})\, \hat{\psi}_{\sigma}(\boldsymbol{x}), \tag{1.10}$$

where $\hat{\psi}_{\alpha}(\boldsymbol{x})$ is the electron field operator, $U(\boldsymbol{x})$ is the screened Coulomb interaction, $V(\boldsymbol{x})$ is the effective potential generated by ions in the unit cell and the index i runs over all the unit

[4] To fix the ideas, consider a cluster of L unit cells per plane and periodic boundary conditions; this will result in a two-dimensional model with L sites.

cells, $\boldsymbol{R}_i$ being the corresponding Bravais lattice vector. Electrons in the inner atomic shells are taken into account effectively by $V(\boldsymbol{x})$, so that the Hamiltonian describes only a subset of valence electrons, which are thought to be of relevance *e.g.* for the conductive properties of the material. This Hamiltonian can be expressed in terms of a new basis of operators, associated to a new basis of states, which could be that of atomic orbitals or of Wannier orbitals. By introducing the unitary transformation for the operators to the Wannier basis $\{c_{i,\alpha,\sigma}\}_{i,\alpha,\sigma}$, the Hamiltonian is cast in the form:

$$\begin{aligned} H = &\sum_{i,j}\sum_{\alpha,\alpha'}\sum_{\sigma} t_{i,j}^{\alpha,\alpha'}\, c_{i,\alpha,\sigma}^{\dagger}\, c_{j,\alpha,\sigma} + \\ &+ \sum_{i_1,i_2,i_3,i_4}\ \sum_{\alpha_1,\alpha_2,\alpha_3,\alpha_4}\ \sum_{\sigma,\sigma'} U_{i_1,i_2,i_3,i_4}^{\alpha_1,\alpha_2,\alpha_3,\alpha_4}\, c_{i_1,\alpha_1,\sigma}^{\dagger}\, c_{i_2,\alpha_2,\sigma'}^{\dagger}\, c_{i_3,\alpha_3,\sigma'}\, c_{i_4,\alpha_4,\sigma}\,, \end{aligned} \tag{1.11}$$

where $t_{i,j}^{\alpha,\alpha'}$ and $U_{i_1,i_2,i_3,i_4}^{\alpha_1,\alpha_2,\alpha_3,\alpha_4}$ are matrix elements of the hopping term between orbital α at site i and orbital α' at site j and of the electron-electron interaction respectively. Some simplifications now come into play: first, the hopping term in the Hamiltonian can be cast in diagonal form in the orbital indices via a unitary transformation; secondly, one considers only the on-site interaction terms and the nearest-neighbor hopping matrix elements, which is a reasonable approximation if the orbitals are well-localized. When the system has a layered structure, which is the case of the materials of interest here (see section 1.4), each plane will be decoupled from the others to some reasonable approximation and it is therefore sufficient to consider the restriction of the Hamiltonian to a single plane[5]. For two degenerate atomic orbitals with the same orbital symmetry, one obtains the following Hamiltonian (a more detailed derivation of this Hamiltonian is relegated to appendix A):

$$\begin{aligned} H_{2B} = &\sum_{\alpha}\sum_{\sigma}\sum_{\langle i,j\rangle}\left(t^{\alpha} c_{i,\alpha,\sigma}^{\dagger} c_{j,\alpha,\sigma} + \text{h.c.}\right) + \\ &+ U\sum_{i}\sum_{\alpha} n_{i,\alpha,\uparrow} n_{i,\alpha,\downarrow} + U'\sum_{i}\sum_{\sigma,\sigma'} n_{i,1,\sigma} n_{i,2,\sigma'} + \\ &- J\sum_{i}\sum_{\sigma,\sigma'} c_{i,1,\sigma}^{\dagger} c_{i,2,\sigma'}^{\dagger} c_{i,1,\sigma'} c_{i,2,\sigma} + \\ &- J'\sum_{i}\left(c_{i,1,\uparrow}^{\dagger} c_{i,1,\downarrow}^{\dagger} c_{i,2,\downarrow} c_{i,2,\uparrow} + \text{h.c.}\right)\,, \end{aligned} \tag{1.12}$$

which is the *two-band* degenerate Hubbard model Hamiltonian (alternative forms have been used in literature [33, 34, 35, 36]) or just *two-band* Hamiltonian, as will be referred to in the following, in contrast to the simple one-band[6] Hubbard model of equation (1.1). The two-band model defined by H_{2B} is the object of interest of this thesis. In equation (1.12), $c_{i,\alpha,\sigma}$

[5] Formally, the Hamiltonian of the system would be written in a block diagonal form:

$$H = \bigotimes_{m} H_{2B}^{(m)}$$

with (m) indicating that the operator acts only on the subspace of states belonging to sites of the same layer. The fundamental object is then $H_{2B}^{(m)}$ and the superscript is understood.

[6] Similarly, the terms *N-band* model will be eventually used.

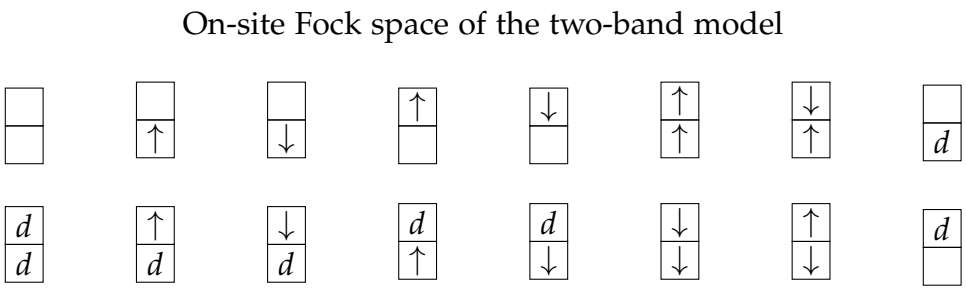


Figure 1.5: This figure shows the full 16-dimensional Fock space for a single site of the two-band model. For each rectangle, the upper and lower squares indicate the first and second orbitals respectively. Each of the squares can either be empty, occupied by one $\uparrow$ or $\downarrow$ electron or doubly occupied d. Just as in the one-band Hubbard model, it is necessary to define states by fixing the order of action of creation operators on the vacuum: for example, the state with both doubly occupied orbitals (leftmost of the second row) is $c^\dagger_{i,c,\uparrow} c^\dagger_{i,f,\uparrow} c^\dagger_{i,c,\downarrow} c^\dagger_{i,f,\downarrow} |0\rangle$.

and $n_{i,\alpha,\sigma}$ are fermionic annihilation operators and the associated occupation numbers, with the additional orbital index α. From this point on, the alphanumeric notations $\alpha = 1 = c$ and $\alpha = 2 = f$ will eventually be adopted[7]. On each site, an electron can be in one out of 4 possible states. The Fock Hilbert space on which H_{2B} acts is 16^L-dimensional, being the L-th tensor product of a 16-dimensional local Fock space $\mathscr{H}$, which is represented in figure 1.5. The first term of H_{2B} is the kinetic term, which allows nearest neighbor hopping term only between orbitals of the same type; the hopping parameters are chosen to be positive and in general different, so that the free bands are both centered in the origin of the Brillouin zone and have different bandwidths. The second line contains diagonal terms in the occupation-number operators, namely the on-site intra-orbital Hubbard U term and the on-site inter-orbital Hubbard U', which both have positive values according to their repulsive character. Finally, the last two lines in equation (1.12) represent the on-site ferromagnetic Hund exchange with $J > 0$ and the so-called *pair-hopping* term, which appears in the derivation of the Hamiltonian. Considering this description, the model can be usefully represented as two overlapping one-band layers with the addition of the inter-layer Coulomb repulsion and the Hund and pair-hopping terms. According to rotational symmetry of degenerate orbitals, the parameters satisfy the relations

[7]The alphabetic notation is reminiscent of the periodic Anderson model (PAM) [37], relevant in the physics of heavy-fermion systems and closely related to the Kondo lattice model. In the PAM, itinerant electrons in a c-band (conducting) interact with electrons localized in a lattice of f atomic orbitals via the Hamiltonian:

$$H_{PAM} = \sum_k \sum_\sigma \xi_k \, c^\dagger_{k,c,\sigma} c_{k,c,\sigma} + \sum_k \sum_l \sum_\sigma V_{k,l} \left(c^\dagger_{k,c,\sigma} c_{k,f,\sigma} + \text{h.c.} \right) + \frac{1}{2} U \sum_i \sum_{(l,\sigma) \neq (l',\sigma')} n_{i,l,\sigma} n_{i,l',\sigma'}$$

The relation with H_{2B} is clear.

[38]:

$$U' = U - 2J; \tag{1.13a}$$

$$J' = J. \tag{1.13b}$$

Relation (1.13b) will be relaxed in certain cases, in order to consider a broader range of physical phenomena.

Starting from the Hamiltonian of equation (1.10), the derivation of a multi-band model is, to a certain extent, arbitrary. This degree of arbitrariness is related to the precision of the description that one seeks to achieve for the model. One could in principle consider hoppings and interactions between all orbitals at each site. However, the full set of orbitals can be reduced to a subset of relevant ones and only the most important couplings are retained, not unlike for the Hubbard model: in this way, a low-energy model Hamiltonian is derived. Which orbitals are to be considered relevant? Consider an exemplary layered-perovskite material, like the one shown in figure 1.6a, which is the case of some cuprate superconductors and other transition-metal oxides: inside each transition-metal-oxygen plane, electrons can occupy either the d orbitals on the transiton-metal ions or the $2p$ orbitals on the oxygen ions. Within this picture, each lattice site should host 8 orbitals with suitably chosen couplings[8]; indeed, such orbitals would be modified by crystal fields and Jahn-Teller distortions and consequently some degeneracies would be lifted according to symmetry. An alternative to the $2p$ and d orbitals is a set of 5 less localized orbitals which resemble the original atomic $3d$ ones, but encode some information from the $2p$ as well. Of course, this would yield a different model from the previous one (it also has a different Hilbert space), which takes into account the hopping processes in a less accurate manner: this choice corresponds to a coarse graining of the degrees of freedom in unit cell. The procedure of dropping some degrees of freedom in order to obtain a simpler effective Hamiltonian emerges most significantly in the description of a copper-oxide plane with two $3d$ and one $2p$ orbital in terms of the (single-band) $t - J$ model [16]. In general, both the arrangement of atomic energy levels of the particular compound and the faithfulness sought for its description affect the choice of the model.

The simplest case where a two-band model could be of interest is when crystal fields with octahedral symmetry split the d orbitals into 3 low-energy t_{2g} orbitals and 2 high-energy e_g orbitals. The former three orbitals are assumed to be degenerate, which corresponds to neglecting the effect of Jahn-Teller distortions. The hybridization with the $2p$ orbitals from the oxygen ions gives the effective degrees of freedom of the model. This also explains the relations between the couplings that have been mentioned earlier. As described in the following, some ruthenate compounds do have 3 lower lying t_{2g} orbitals: as a consequence, one might think that a description in terms of a two-band model would be insufficient. However, despite the ultimate interest would be describing real materials, the present work is not focused on any of them, but rather aims at studying the two-band model from a general point of view.

[8] For example according to some *ab-initio* calculation, such as Density Functional Theory.

The rationale behind this viewpoint is twofold: first, the two-band model allows for a wider range of physical situations compared to the single-band model, at the same time being the easiest case of multi-band model. Secondly, modifying the couplings in the Hamiltonian allows to study the phase diagram of the model, thereby giving a comprehensive view of the competing tendencies in the model; eventually, real compounds are described by choosing a particular set of couplings.

In the limit of $t_2/t_1 = 1$ and $J = 0$, the two-band model becomes even more symmetric than the usual $SU(2) \times SU(2) \times U(1)$ case: it can be easily shown the system has an enhanced global $SU(4)$-(spin-orbital) invariance.

1.4 Experimental realizations

Studying a two-band Hubbard model is justified by the existence of materials which can be satisfactorily described in terms of it. The effective degrees of freedom described by a two-band model have already been pictured. Here, we describe the real situation of some ruthenium based oxides, which originally stimulated interest in the two-band model. Apart from these materials, there are lots of compounds which have more than one relevant band, from the alkali-doped fullerenes to the more recently discovered iron pnictides.

1.4.1 Ruthenates

One of the classes of compounds which fit a description in terms of two effectives bands is that of $4d$ Ru orbitals in ruthenates, in particular the isoelectronic alloy series $Ca_{2-x}Sr_xRuO_4$[9] [17]. The end member Sr_2RuO_4 ($x = 2$) is a 3-dimensional anisotropic Landau-Fermi liquid [39]. Its crystal structure is the undistorted layered-perovskite (K_2NiF_4) structure, depicted in figure 1.6a, and the Ru octahedra are slightly elongated along the z-axis. The t_{2g} shell of $4d$ Ru orbitals lies lower in energy than the e_g shell and is occupied by 4 electrons. Due to the z-axis elongation, there is a small splitting between the degenerate $\{xz, yz\}$-orbitals and the xy-orbital. The $\{xz, yz\}$-orbitals hybridize only with one pair of $2p$ oxygen orbitals in the xy-plane, while the xy-orbital hybridizes with all 4 of them, which results in a bandwitdh which is approximately twice as larger than those of the $\{xz, yz\}$-orbitals. Despite strong correlations lead to enhancements in the electron mass, a Landau-Fermi liquid behavior is observed. Remarkably, the compound Sr_2RuO_4 is interesting because of its spin-triplet superconductivity [40], which also arises in the heavy-fermion compounds.

Substitution of Sr^{2+} ions with the smaller Ca^{2+} ions gives the compound Ca_2RuO_4 ($x = 0$) which is an antiferromagnetic insulator (up to about $100K$). In this compound, the Ru octahedra experience a compression along the z-axis, combined with a tilt-rotation. The overall effect is a change of sign in the energy splitting between the $\{xz, yz\}$-orbitals and the xy-orbital,

[9] In this section x denotes the concentration of Sr ions.

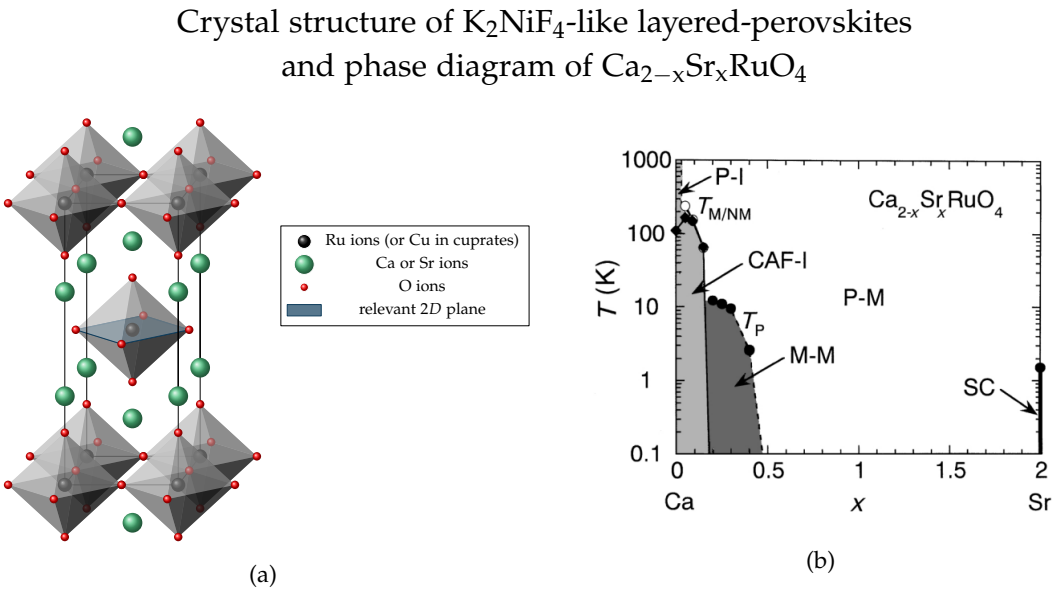


Figure 1.6: (a): Undistorted crystal structure (K_2NiF_4 layered-perovskite) of the alloy series $Ca_{2-x}Sr_xRuO_4$; this also applies to some of the cuprates, for instance $La_{2-x}Sr_xCuO_4$. (b): Phase diagram in the strontium doping percentage-temperature plane of the alloy series $Ca_{2-x}Sr_xRuO_4$, as reported by reference [39].

which now lies lower in energy. The antiferromagnetic and insulating behavior of Ca_2RuO_4 is well understood in terms of LDA+U calculations [17].

The appealing feature of this alloy series is the evolution from the Fermi liquid to the Mott insulator: the full phase diagram is shown in figure 1.6b. Superconductivity of Sr_2RuO_4 is readily suppressed upon doping with Ca and leads to a paramagnetic metal in the doping range $0.5 < x < 2$. At the critical value $x_c \simeq 0.5$, the compound shows anomalous properties, namely a metallic behavior associated with the onset of magnetism with moment $S = 1/2$, not $S = 1$ as could be expected by isolated Ru^{4+} ions and as observed in Ca_2RuO_4. These properties are interpreted in terms of an orbital-selective behavior: the $\{xz, yz\}$-orbitals host 3 electrons and undergo a Mott transition, displaying a local moment $S = 1/2$. The remaining electron occupies the xy-band and has itinerant behavior, resulting in a metallic character. This situation holds up to Sr concentration of $x = 0.2$; past this value, the compound becomes insulating and resembles the case of Ca_2RuO_4.

1.5 Orbital Selectivity in the two-band model

Despite the ruthenate compounds having three effective bands, a two-band model can adequately describe a state where two bands have dissimilar behavior. Such a state is said to be

orbital selective. Orbital selectivity essentially refers to the particular and clearest case of an Orbital Selective Mott Phase (OSMP)[10], which consists in a phase where electrons in one band are itinerant, while those belonging to another band form a Mott insulating state. This OSMP corresponds to the coexistence of itinerant and localized electrons in different bands suggested to be realized in ruthenate compounds; evidence for its existence has been recently found in iron pnictides [41] as well.

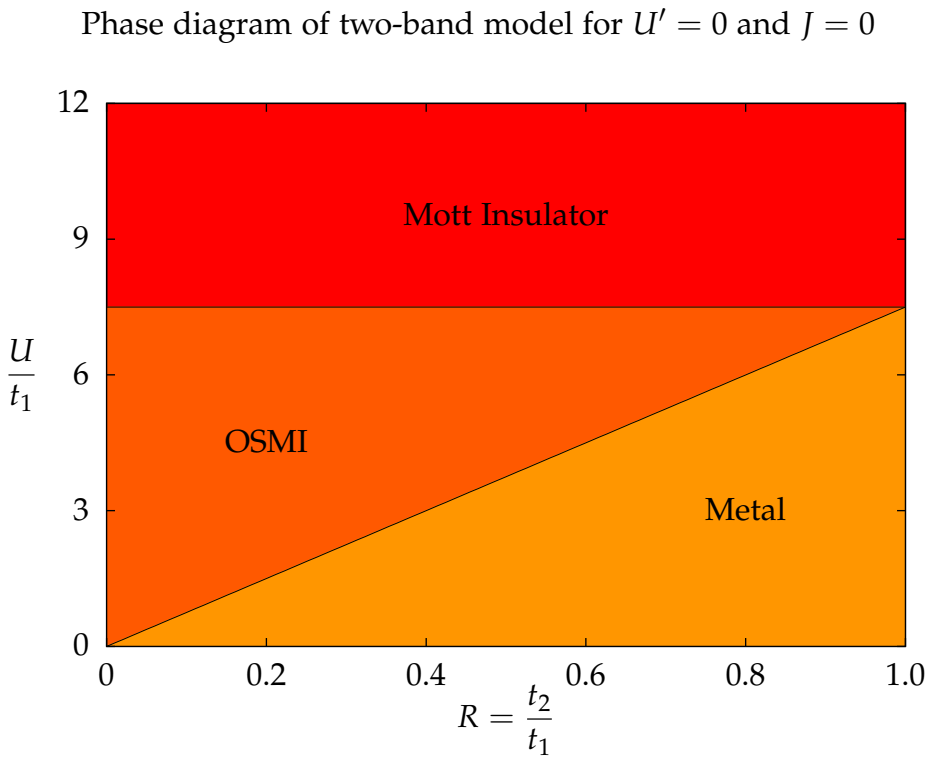
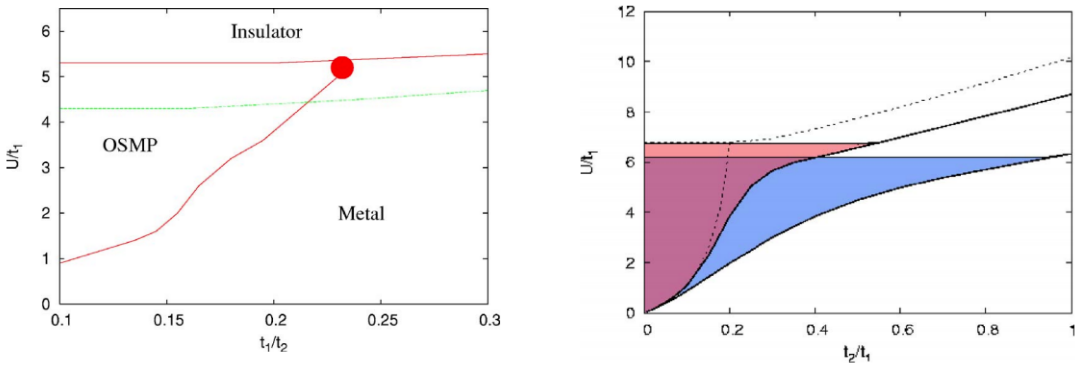


Figure 1.7: At $U' = 0$ and $J = 0$ the phase diagram at half-filling is easily drawn in terms of the phase diagram of the one-band model. Here, coherently with results on the Hubbard model of the present work (see section 4.1), the Mott transition is assumed to be at $(U/t_1)_c = 7.5 \pm 0.5$. The transition where the c-band becomes insulating is always at $(U/t_1)_c = 7.5$. The two bands do not interact with each other: consequently the Mott transition of the f-band occurs for $U/t_1 = (U/t_1)_c\, R$, which corresponds to the oblique line in figure. All transition lines represent continuous second-order phase transitions.

How does the two-band model describe an OSMP? In order to see this, consider the case of vanishing Hund's coupling $J = 0$. It is useful to introduce two couplings which span

[10] Or Orbital Selective Mott Insulator (OSMI), while the corresponding transition from a two-band Mott insulator or a two-band metal is termed Orbital Selective Mott Transition (OSMT).

Phase diagram of two-band model obtained within DMFT and mean-field slave-spin representation



(a) (b)

Figure 1.8: (a): Close up of the phase diagram of the two-band model according to DMFT in a region where all three phases are present, according to reference [33]; the red lines are transition lines, while the green line is the coexistence line of the Mott insulator. (b): Phase diagram with $J = 0.01U$ (pink) and $J = 0.1U$ (light blue) obtained with mean-field slave-spin representation.

the zero-temperature phase diagram, namely the ratio between the hopping matrix elements $R \equiv t_2/t_1$ and the ratio U/t_1: this choice corresponds to fixing the bandwidth of the first band as the energy scale and exploring the phase diagram by modifying t_2 and U. It is possible to restrict R to the range $[0,1]$, since each point with $R > 1$ corresponds to one in $[0,1]$ upon swapping the bands: in this range the f band will be the most correlated one. If the constrain $U' = U$ is relaxed and the inter-band Coulomb repulsion is neglected ($U' = 0$), the two-band model becomes equivalent to two non-interacting single-band models (see figure 1.7). In this case, it is clear that an OSMI is present in a large region of the phase diagram and the same should hold as long as $U' \ll U$. The case when $U' = U$ is not as straightforward. Although the presence of an OSMI was challenged in the first place [42, 43], it has been predicted by a variety of methods, such as DMFT [44, 36], the Gutzwiller approach [45] and mean-field slave-spin representation [33], as can be seen in figure 1.8. It is now generally accepted that such a phase exists for $U' = U$ at least for a very asymmetric bandwidths ($R \ll 1$).

Another common belief is that turning on the Hund's coupling J stabilizes the OMSP on a wider range of R, while favoring the Mott insulator up to lower values of U/t_1; this fact is supported by multiple calculations [33, 45]. In general, the effect of the Hund's coupling has a striking dependence on the filling factor of a degenerate Hubbard model: DMFT calculations show that the critical value of the insulating phase is decreased only for global half-filling [46, 47] while acting as a *band-decoupler* and favoring the OSMP over the metallic state.

2

Variational Monte Carlo

Strongly-correlated electron systems rarely allow exact solutions. Even the simplest single-band Hubbard model can be solved analytically only in one dimension [21]. Thus one must usually rely upon non-analytical approaches, either approximate solutions or numerical methods. A large class of methods (including for instance the Density Matrix Renormalization Group) are variational in nature, since they are all based on a fundamental theorem, namely the so-called *variational principle*. Consider a Hilbert space $\mathscr{H}$ and let H be a Hamiltonian with eigenstates $|\psi_i\rangle$ and eigenvalues E_i, $|\psi_0\rangle$ being the ground-state[1]; then for any trial state $|\Psi_V\rangle = \sum_i a_i |\psi_i\rangle$ with normalization condition $\sum_i |a_i|^2 = 1$:

$$E_V \equiv \frac{\langle\Psi_V| H |\Psi_V\rangle}{\langle\Psi_V|\Psi_V\rangle} = \sum_i |a_i|^2 E_i \geq \sum_i |a_i|^2 E_0 = E_0. \tag{2.1}$$

The variational energy E_V, *i.e.* the expectation value $\langle\Psi_V| H |\Psi_V\rangle$, is always an upper bound to the ground-state energy of the system. One then has to empirically build a variational wave function with properties that correctly describe the physics of the system. The better the trial wave function, the closer the variational energy will be to the true ground-state energy. Such a wave function is often taken of a parametric form: in this case, it must be optimized by finding the set of parameters which minimizes the expectation value of the Hamiltonian. Due to the fact that the wave function is correlated, a numerical technique is required to compute expectation values. The Variational Monte Carlo (VMC) method is a numerical technique which computes the expectation values of observables on the trial wave function by sampling

[1] Throughout this chapter this general case will be taken into account.

the states of the Hilbert space via the Metropolis algorithm. Consequently, the outcomes of a VMC computation are intrinsically endowed with an error bar. Since VMC is based on the trial wave function, it is clear that the outcomes will always reflect the variational *ansatz*: this is the major shortcoming of the method, since one will never find new physics behind what was already expected. On the other hand, the method can be used to compare different variational *ansätze*: thanks to the parametric dependence, the variational wave function can be made fairly general, allowing one to take into account a broad range of phenomena. Moreover, since the VMC method is based on the variational principle (2.1), the expectation value of the energy is an upper bound to the exact value, a feature that allows improving the accuracy of the calculation in a controlled manner.

This chapter aims at presenting the VMC method from a general point of view: first, the evaluation of expectation values is addressed and it is shown how it can be accomplished by means of a Markov chain; subsequently, we explain the standard sampling technique, namely the Metropolis algorithm; the last section reviews the Stochastic Reconfiguration algorithm, which allows on-the-fly optimization of the wave function.

2.1 Evaluation of observables

What one ultimately needs to know about a physical system are expectation values of observables and their correlation functions. These quantities are respectively of the forms:

$$\langle O \rangle_V \equiv \frac{\langle \Psi_V | \, O \, | \Psi_V \rangle}{\langle \Psi_V | \Psi_V \rangle} \tag{2.2}$$

and:

$$\mathcal{C}_V^{i_1 \dots i_N} \equiv \frac{\langle \Psi_V | \, O_{i_1} \dots O_{i_N} \, | \Psi_V \rangle}{\langle \Psi_V | \Psi_V \rangle} \,, \tag{2.3}$$

the variational energy E_V being a special case of expectation value with $O = H$. Consider the case of expectation values. Let $\{|x\rangle\}_{x \in X}$ be a suitable (complete) basis of the Hilbert space, with X being a certain set; then one can write the expectation value of O as:

$$\langle O \rangle_V = \frac{\sum_x \langle \Psi_V | \, O \, |x\rangle \, \langle x | \Psi_V \rangle}{\sum_x \langle \Psi_V | x \rangle \, \langle x | \Psi_V \rangle} = \frac{\sum_x |\langle \Psi_V | x \rangle|^2 \, O_V(x)}{\sum_x |\langle \Psi_V | x \rangle|^2} = \sum_x O_V(x) \, p_V(x); \tag{2.4}$$

where the following quantities have been defined:

$$O_V(x) = \frac{\langle \Psi_V | \, O \, |x\rangle}{\langle \Psi_V | x \rangle} \tag{2.5}$$

and:

$$p_V(x) = \frac{|\langle \Psi_V | x \rangle|^2}{\sum_x |\langle \Psi_V | x \rangle|^2} \,, \tag{2.6}$$

which are respectively the variational *local* density of O and the probability density $p_V(x)$. The quantum expectation value can thus be written as the expectation value of the local density

(or local value) of O on a probability density p_V defined on X. Therefore one should compute overlaps and matrix elements on every state of the Hilbert space, a task that becomes impossible as soon as one is taking into account large enough sizes of the system. In place of this, one can hope to compute those values on a significative multisubset $\mathcal{S} \subset X$ of N elements of the Hilbert space, which of course must be distributed according to the probability $p_V(x)$, that is:

$$\langle O \rangle_V \simeq \frac{1}{N} \sum_{x \in \mathcal{S}} O_V(x). \tag{2.7}$$

This approximation introduces the stochastic evaluation of observables. The approximate equation (2.7) is a consequence of the law of large numbers: the expectation value is approximated by the mean value of $O_V(x)$ on the configurations in $\mathcal{S}$; these are generated according to the probability distribution $p_V(x)$, which is a great advance with respect to picking them evenly distributed in the whole Hilbert space and weigh them. At this point, the only question is: how to generate a set of configurations distributed according to a given $p_V(x)$? This question will be answered in the next section.

2.2 Sampling through Markov chains

To start, a brief reminder of stochastic processes is needed. A *discrete* stochastic process is a collection of random variables $\{X_n\}_n$, each of them valued in a *sample* space Ω and distributed according to a probability distribution $p_n(\omega)$ (n plays the role of a discrete time index). In this framework, it is sufficient to consider Ω as a finite set. A stochastic process is said to be in state ω at time n if the random variable X_n takes the value ω. We denote by $p(X_n = \omega_n, \ldots, X_1 = \omega_1)$ the probability that the the stochastic process has been through states $\omega_1, \ldots, \omega_n$ at times $1, \ldots, n$ respectively; this quantity is termed *joint* probability. Similarly, let $p(X_n = \omega_n | X_{n-1} = \omega_{n-1}, \ldots, X_1 = \omega_1)$ denote the *conditional* probability that the stochastic process be in state ω_n at time n, provided that it was in states $\omega_{n-1}, \ldots, \omega_1$ at times $n-1, \ldots, 1$ respectively. There is a fundamental relation between the joint and the conditional probabilities:

$$p\left(X_n = \omega_n | X_{n-1} = \omega_{n-1}, \ldots, X_1 = \omega_1\right) = \frac{p\left(X_n = \omega_n, \ldots, X_1 = \omega_1\right)}{p\left(X_{n-1} = \omega_{n-1}, \ldots, X_1 = \omega_1\right)}; \tag{2.8}$$

the probability at the denominator is the so-called *marginal* probability with respect to time n.

A stochastic process such that:

$$\begin{aligned} p(X_n = \omega_n | X_{n-1} = \omega_{n-1}, \ldots, X_1 = \omega_1) = p(X_n = \omega_n | X_{n-1} = \omega_{n-1}) = \\ \equiv K(\omega_n | \omega_{n-1}), \end{aligned} \tag{2.9}$$

is said to be a *Markov process* or *Markov chain*. Markov processes are said to have no memory in the sense that the probability of the n-th outcome being ω_n depends only on the $(n-1)$-th outcome and not on all the previous history of the process. $K(\omega|\xi)$ is the *transition probability*

and defines the Markov process by quantifying its probability to go from state ξ to state ω in one step[2]. The transition probability satisfies the following properties:

$$K(\omega|\xi) \geq 0 \quad \forall \omega, \xi \in \Omega\,; \tag{2.10a}$$

$$\sum_{\omega \in \Omega} K(\omega|\xi) = 1 \quad \forall \xi \in \Omega\,; \tag{2.10b}$$

i.e. it is a probability distribution itself as a function of its first argument. By considering the fundamental relation between the joint and conditional probabilities, equation (2.8), the master equation for a Markov process follows:

$$p_{n+1}(\omega_{n+1}) = \sum_{\omega_n} K(\omega_{n+1}|\omega_n)\, p_n(\omega_n), \tag{2.11}$$

with the notation $p_n(\omega_n) \equiv p(X_n = \omega_n)$. A straightforward consequence of the master equation is the Chapman-Kolmogorov property for Markov chains. From both these relations, it is clear that even if the n-th outcome of the stochastic process is unknown, its probability distribution is deterministically given by the transition probability and the initial probability distribution. Given a probability density $p(\omega)$ and a transition probability $K(\eta|\xi)$, they are said to satisfy the condition of detailed balance if:

$$K(\omega|\xi)\, p(\xi) = K(\xi|\omega)\, p(\omega), \tag{2.12}$$

which means that the probability of being in ω and moving to state ξ is invariant if ω and ξ are interchanged. Detailed balance turns out to be extremely useful in the following.

2.2.1 Convergence of a Markov chain

Consider now the limiting behavior of a Markov chain: it is possible, under some given assumptions, to have a Markov chain sample a desired probability density in the large n limit. This subsection aims at proving this and finding the required assumptions under which this happens. First of all, does a limiting stationary probability density exist? To see this, write down the master equation, assume the detailed balance condition to hold and exploit the normalization condition for the transition probability:

$$p_{n+1}(\omega) = \sum_{\xi} K(\omega|\xi)\, p_n(\xi) = \sum_{\xi} K(\xi|\omega)\, p_n(\omega) = p_n(\omega). \tag{2.13}$$

Therefore it is sufficient for a probability distribution to satisfy detailed balance in order to be stationary. A more subtle question is whether, given any initial distribution p_0, the stochastic process will eventually converge to and sample a stationary distribution p_{eq}. In order to address this problem, it is necessary to state the Perron-Frobenius theorem. A square matrix $\boldsymbol{A}$ is said to be irreducible if, for any i and j, there exists an integer m such that $(\boldsymbol{A}^m)_{ij} \neq 0$.

[2]Note: the fact that $K(\omega|\xi)$ does not depend explicitly on the time index n means that the Markov process is in fact stationary. However, this does not mean that the probability distribution is stationary: this issue is discussed in the next subsection.

Perron-Frobenius theorem. *Given any irreducible, real symmetric matrix $\boldsymbol{A}$ with non-negative entries $A_{ij} \geq 0$, its highest eigenvalue λ_{Max} is non-degenerate and the components of the corresponding eigenvector all have the same sign.*

The following serves both as a proof of this theorem and as its application to the convergence of Markov chains. Consider the quantity:

$$A(\xi|\omega) \equiv K(\xi|\omega)\sqrt{\frac{p_{eq}(\omega)}{p_{eq}(\xi)}}\,, \tag{2.14}$$

which is symmetric under the exchange of $\xi \leftrightarrow \omega$, due to the detailed balance property; note that in order to define this quantity it is required that $p_{eq}(\omega) \neq 0$ for any ω. At this point it is at least insightful to switch to a matrix form notation, namely:

$$A(\xi|\omega) \to A_{\xi\omega}\,; \tag{2.15a}$$

$$K(\xi|\omega) \to K_{\xi\omega}\,; \tag{2.15b}$$

$$\sqrt{p_{eq}(\xi)} \to g_\xi\,; \tag{2.15c}$$

then $\boldsymbol{A}$ is a symmetric matrix with non-negative entries. It can be seen that the matrices $\boldsymbol{A}$ and $\boldsymbol{K}$ have the same spectrum, while their eigenvectors are different. It follows from equation (2.14) and the normalization condition of the transition probability that:

$$\sum_\xi A_{\xi\omega}\, g_\xi = \sum_\xi K(\xi|\omega)\,\sqrt{p_{eq}(\omega)} = \sqrt{p_{eq}(\omega)} = g_\omega\,, \tag{2.16}$$

that is, the vector $\boldsymbol{g}$ is eigenvector[3] of the matrix $\boldsymbol{A}$ with eigenvalue equal to 1. It is shown in the following that $\boldsymbol{g}$ is the eigenvector with highest eigenvalue and that, under the further assumption of ergodicity, 1 is a non-degenerate[4] eigenvalue. First of all, there are no other eigenvalues greater than one. In order to prove this statement, assume *ad absurdum* that there exists a normalized eigenvector $\boldsymbol{a}$ of $\boldsymbol{A}$ with maximum eigenvalue $\lambda_{Max} > 1$. Then:

$$\lambda_{Max} = \left|\sum_{\xi\omega} a_\xi\, A_{\xi\omega}\, a_\omega\right| \leq \sum_{\xi\omega} |a_\xi|\, A_{\xi\omega}\, |a_\omega| = \sum_{\xi\omega} b_\xi\, A_{\xi\omega}\, b_\omega, \tag{2.17}$$

where the vector $\boldsymbol{b}$ is defined by $b_\xi = |a_\xi|$. The previous inequality states that the expectation value of $\boldsymbol{A}$ on the vector $\boldsymbol{b}$ is greater than or equal to λ_{Max}. At the same time, since $\boldsymbol{A}$ is symmetric, its expectation value on any vector must be lower than its maximum eigenvalue (this statement corresponds to the variational principle). As a consequence:

$$\sum_{\xi\omega} b_\xi\, A_{\xi\omega}\, b_\omega = \lambda_{Max}\,, \tag{2.18}$$

[3] Since $\boldsymbol{A}$ is symmetric, left and right eigenvectors coincide.

[4] $\boldsymbol{g}$ resembles the ground-state of the opposite of the symmetric matrix $\boldsymbol{H} = -\boldsymbol{A}$, which in turns recalls a Hamiltonian from Quantum Mechanics.

or rather, $\boldsymbol{b}$ is an eigenvector of $\boldsymbol{A}$ itself, with eigenvalue λ_{Max} and all positive elements. However, $\boldsymbol{g}$ is an eigenvector with all positive elements as well: therefore, the inner product between $\boldsymbol{g}$ and $\boldsymbol{b}$ must be strictly positive and thus they are not orthogonal. This is absurd, because eigenvectors relative to different eigenvalues must be orthogonal. In conclusion, the maximum eigenvalue of $\boldsymbol{A}$ is 1.

Under the further assumption of ergodicity, it can be shown that 1 is a non-degenerate eigenvalue. Ergodicity states that any configuration can be eventually reached from any other: in other words, the probability of transitioning between any two configurations ξ and ω is non-zero upon applying m times the matrix $\boldsymbol{A}$, which is nothing but stating that $\boldsymbol{A}$ is an irreducible matrix. Suppose *ad absurdum* that $\boldsymbol{a}$ is an eigenvector with eigenvalue 1, which differs from g by more than a multiplicative constant. Then by linearity $\boldsymbol{a} + c\boldsymbol{g}$ is an eigenvector with maximum eigenvalue for any value of the real constant c and, by the previous argument, the vector $\boldsymbol{f}$ such that $f_\omega = |a_\omega + c g_\omega|$ must be eigenvector with eigenvalue 1 as well. The constant c can be chosen in such a way that $f_\eta = 0$ for a given η. Then:

$$\sum_\omega A_{\eta\omega} f_\omega = f_\eta = 0\,, \tag{2.19}$$

but this implies that $\boldsymbol{f}$ must be 0 for all ω connected to η by $\boldsymbol{A}$, since both $\boldsymbol{A}$ and $\boldsymbol{f}$ have non-negative components; under the hypothesis of ergodicity, repetitively applying the last equation implies that $\boldsymbol{f} = \boldsymbol{0}$ for all configurations, which is absurd. Therefore, $\boldsymbol{f}$ differs from $\boldsymbol{g}$ only by a multiplicative constant, in other words $\boldsymbol{g}$ is unique.

In order to conclude the argument, consider any initial probability density $p_0(\omega)$ (the original notation is restored for the probabilities). Then, after n steps the probability density would be:

$$p_n(\omega') = \sum_\omega (\boldsymbol{K}^n)_{\omega'\omega}\; p_0(\omega) = \sum_\omega \sqrt{\frac{p_{eq}(\omega')}{p_{eq}(\omega)}}\; (\boldsymbol{A}^n)_{\omega'\omega}\; p_0(\omega), \tag{2.20}$$

where the transition matrix $\boldsymbol{K}$ has been applied n times to the initial probability density and then cast in terms of $\boldsymbol{A}$. Then, by introducing the spectral representation of $\boldsymbol{A}$:

$$A_{\omega'\omega} = \sum_i \lambda_i\, v_i(\omega')\, v_i(\omega), \tag{2.21}$$

it turns out that:

$$p_n(\omega') = \sum_\omega \sqrt{\frac{p_{eq}(\omega')}{p_{eq}(\omega)}} \Big[\sum_\alpha \lambda_i^n\, v_i(\omega')\, v_i(\omega) \Big]\, p_0(\omega) = \tag{2.22}$$

$$= \sqrt{p_{eq}(\omega')} \sum_i \lambda_i^n\, v_i(\omega') \left[\sum_\omega v_i(\omega)\, \frac{p_0(\omega)}{\sqrt{p_{eq}(\omega)}} \right]. \tag{2.23}$$

For large n, the term with largest eigenvalue (let it be the one with $i = 0$) dominates, while all the others decay exponentially with n; taking into account the fact that $v_0(\omega)$ is nothing but

$\sqrt{p_{eq}(\omega)}$:

$$p_n(\omega') \xrightarrow{n\to\infty} \sqrt{p_{eq}(\omega')}\, v_0(\omega') \left[\sum_\omega v_0(\omega) \frac{p_0(\omega)}{\sqrt{p_{eq}(\omega)}} \right] = \tag{2.24}$$

$$= \left(\sqrt{p_{eq}(\omega')} \right)^2 = p_{eq}(\omega'). \tag{2.25}$$

This is the required result. It has been proved that under the hypotheses of detailed balance and ergodicity the Markov chain will converge to a stationary probability distribution, provided that the initial condition has non-zero amplitude for any configuration.

2.2.2 The Metropolis algorithm

It has been proved in the previous section that the evaluation of observables according to equation (2.7) involves a stochastic process which satisfies detailed balance and ergodicity. In this section it is shown how to construct a Markov process which satisfies the required properties: this algorithm was proposed by N. Metropolis *et al.* [48]. In order to construct the suitable stochastic process, one divides the transition probability into a product of a *trial probability* and an *acceptance probability*:

$$K(x'|x) = T(x'|x) \cdot A(x'|x). \tag{2.26}$$

Here and forth x is an element (or configuration) in the space X which takes the role of the sample space, *e.g.* an element in a basis of a Hilbert space. The rationale behind this factorization is that one proposes a new configuration generated according to the trial probability and subsequently either accepts or rejects it according to the acceptance probability. The trial probability can be very general: usually it is chosen to reflect the physics of the system under investigation, but also has to deal with computational complexity; most importantly, a correct choice of the trial probability is needed in order for the system to be ergodic. In this context, the trial probability is chosen to be symmetric, that is $T(x'|x) = T(x|x')$: this choice leads to a simplification of the algorithm but it is not a necessary requirement. Then, the equation of detailed balance reduces to:

$$\frac{p_{eq}(x')}{p_{eq}(x)} = \frac{A(x'|x)}{A(x|x')}; \tag{2.27}$$

thus the acceptance probability is defined by the stationary probability. Equation (2.27) makes sense, since $p_{eq}(x)$ is non-zero for any configuration x, as required by equation (2.14). Taking into account that the acceptance probability is itself a probability distribution, it follows that:

$$A(x'|x) = \min \left[1, \frac{p_{eq}(x')}{p_{eq}(x)} \right]. \tag{2.28}$$

Operatively, one proceeds in the following way:

1. Starting from a configuration x, a trial configuration x' is randomly generated with probability $T(x'|x)$. For instance, in the Hubbard model a new configuration is generated by hopping an electron to a neighboring site.
2. If the trial configuration is more probable than the old, then $A(x'|x) = 1$, so x' is automatically accepted and it becomes the new configuration of the system.
3. If this is not the case, the trial configuration may still be accepted with a probability $\frac{p_{eq}(x')}{p_{eq}(x)}$.

The algorithm is then iterated until the probability distribution reaches equilibration and further on to sample $p_{eq}(x')$.

2.2.3 Decorrelation and errors

It is now important to take into account the statistical analysis which enables to compute a proper result (expectation value plus error) for any quantity. Suppose one can sample values of $O_V(x)$ according to the desired probability and compute the expectation value by means of equation (2.7), which can be done via the previously described Markov process. It is then clear that values of $O_V(x)$ sampled consecutively are not independent, since each configuration is generated by the previous according to the transition probability $K(x'|x)$.

The outcome of the sampling is a dataset of N values of the random variable O, which we denote by $\{O_i\}_{i=1}^N$. Expectation values of observables are computed according to (2.7), which is nothing but the mean value of the dataset. If each evaluation of an observable were independent from the others, the error could be estimated simply as the square root of the variance:

$$\sigma_{O_V} \equiv \sqrt{\langle (O - \langle O \rangle_V)^2 \rangle_V} = \sqrt{\frac{1}{N-1} \sum_i \left(O_i - \frac{1}{N} \sum_j O_j \right)^2}. \tag{2.29}$$

Unfortunately, this is not the case, since at any step the configuration is generated from the previous one; thus, subsequent configurations are correlated between themselves and so are the values O_i. To be more specific, the Markov process will have a correlation time within which configurations cannot be considered independent. In order to be able to apply equation (2.29), one has to take care of this correlation time. An immediate way to accomplish this is to sample a value of the observable with a frequency at most equal to the inverse of the correlation time, so that a dataset with independent measures is obtained. Another way to obtain a set of independent measures is to subdivide the original dataset into N_b long enough bins, each containing M values of the observable (with $N = M \cdot N_b$) and compute the mean value for each bin $O_l^{(b)}$. The expectation value is then computed as the mean over the bins of this quantity:

$$\langle O \rangle_V = \frac{1}{N_b} \sum_{l=1}^{N_b} O_l^{(b)} = \frac{1}{N} \sum_{i=1}^{N} O_i \,. \tag{2.30}$$

Turning to the evaluation of errors, the measures in the dataset $\{O_l^{(b)}\}_{l=1}^{N_b}$ will be effectively independent, provided that the bin length M is larger than the correlation time of the Markov chain. If this condition is satisfied, then:

$$\sigma_{O_V} = \sqrt{\frac{1}{N_b - 1} \sum_{l=1}^{N_b} \left(O_l^{(b)} - \langle O \rangle \right)^2}. \tag{2.31}$$

2.3 Stochastic Reconfiguration

Until now, we made no mention about the parametric dependence of the variational wave function: the whole discussion assumed that a particular wave function, possibly the best out of many, was given and dealt with computing expectation values on it. However, the variational wave function usually depends on a set of parameters $\{\alpha_i\}_{i=1}^{N_p}$, which may allow to represent states with various physical properties. In principle, this wave function could be as general as desired at the expenses of computational efficiency.

When a manifold of wave functions is given, the question is how to sort out the best one out of it. Here *best* means lowest in energy[5], which is the criterion used to determine how close a variational wave function is to the ground-state. It is possible to compute energy values on different trial wave functions, compare them and determine which one is the best. Alternatively, one can take advantage of the aleatory nature of the computation and introduce the so-called Stochastic Reconfiguration algorithm (SR), a generalization of the Steepest Descent method (SD), which was developed by S. Sorella and collaborators [49, 50]. Consider the variational wave function $|\Psi_V(\boldsymbol{\alpha})\rangle$ and make a small change of the parameters $\boldsymbol{\alpha} \to \boldsymbol{\alpha}' = \boldsymbol{\alpha} + \delta\boldsymbol{\alpha}$. At first order, one can represent the transformed variational state in terms of the original one as:

$$|\Psi_V(\boldsymbol{\alpha}')\rangle = |\Psi_V(\boldsymbol{\alpha})\rangle + \delta\boldsymbol{\alpha} \cdot \boldsymbol{\nabla}_{\boldsymbol{\alpha}} |\Psi_V(\boldsymbol{\alpha})\rangle , \tag{2.32}$$

which is just the first order Taylor expansion of $|\Psi_V(\boldsymbol{\alpha}')\rangle$ around $\boldsymbol{\alpha}'$. This equation can be rewritten as:

$$\begin{aligned} |\Psi_V(\boldsymbol{\alpha}')\rangle &= \delta\alpha_0 |\Psi_V(\boldsymbol{\alpha})\rangle + \sum_{i=1}^{N_p} \delta\alpha_i \frac{\partial}{\partial\alpha_i} |\Psi_V(\boldsymbol{\alpha})\rangle = \\ &= \delta\alpha_0 |\Psi_V(\boldsymbol{\alpha})\rangle + \sum_{i=1}^{N_p} \sum_x \delta\alpha_i \frac{\partial \langle x | \Psi_V(\boldsymbol{\alpha})\rangle}{\partial\alpha_i} |x\rangle = \\ &= \delta\alpha_0 |\Psi_V(\boldsymbol{\alpha})\rangle + \sum_{i=1}^{N_p} \sum_x \delta\alpha_i \frac{\partial \ln\left(\langle x | \Psi_V(\boldsymbol{\alpha})\rangle \right)}{\partial\alpha_i} |x\rangle \langle x | \Psi_V(\boldsymbol{\alpha})\rangle ; \end{aligned} \tag{2.33}$$

in principle the existence of a $\delta\alpha_0$ different from 1 is allowed, since an overall prefactor does not change the state, since it can be normalized. Equation (2.33) can be cast into the very

[5] Another criterion is to adopt the energy variance as the figure of merit which has to be minimized, since the ground-state has zero energy variance.

compact form:

$$\left|\Psi_V(\boldsymbol{\alpha}')\right\rangle = \delta\boldsymbol{\alpha}\cdot\hat{\boldsymbol{\Gamma}}_\Psi\left|\Psi_V(\boldsymbol{\alpha})\right\rangle, \tag{2.34}$$

where the vector operator of logarithmic derivatives $\hat{\boldsymbol{\Gamma}}_\Psi$ has been introduced:

$$\hat{\Gamma}_{\Psi i} = \begin{cases} \mathbb{1} & \text{if} \quad i=0; \\ \sum_x \dfrac{\partial \ln\left(\left\langle x\,|\,\Psi_V(\boldsymbol{\alpha})\right\rangle\right)}{\partial\alpha_i}\left|x\right\rangle\left\langle x\right| & \text{otherwise.} \end{cases} \tag{2.35}$$

Until now, only a few algebraic manipulations were done and the optimization of the parameters was not mentioned. In order to find a better variational wave function, the new state is set equal to the application of the operator of the power method on the original one, that is:

$$\left|\Psi_V(\boldsymbol{\alpha}')\right\rangle = \left(\Lambda\mathbb{1}-\hat{H}\right)\left|\Psi_V(\boldsymbol{\alpha})\right\rangle \tag{2.36}$$

and for a large enough $\Lambda\in\mathbb{R}^+$, this is known to improve the variational wave function[6]. Then, by combining equations (2.36) and (2.34) and projecting onto $\hat{\Gamma}_{\Psi i}\left|\Psi_V(\boldsymbol{\alpha})\right\rangle$, one obtains N_p+1 equations for the variation of the parameters:

$$\Lambda\left\langle\hat{\Gamma}^\dagger_{\Psi i}\right\rangle_V - \left\langle\hat{\Gamma}^\dagger_{\Psi i}\hat{H}\right\rangle_V = \sum_{j=0}^{N_p}\delta\alpha_j\left\langle\hat{\Gamma}^\dagger_{\Psi i}\hat{\Gamma}_{\Psi j}\right\rangle_V. \tag{2.37}$$

This is a linear system in $\delta\boldsymbol{\alpha}$. The arbitrary quantities Λ and $\delta\alpha_0$ still figure in the equations. One can get rid of Λ by considering the equation for $i=0$ and obtaining $\delta\alpha_0$,

$$\delta\alpha_0 = \Lambda - \left\langle\hat{H}\right\rangle_V - \sum_{j=1}^{N_p}\delta\alpha_j\left\langle\hat{\Gamma}_{\Psi j}\right\rangle_V, \tag{2.38}$$

which is then plugged in the remaining N_p equations for $i\geq 1$:

$$\Lambda\left\langle\hat{\Gamma}^\dagger_{\Psi i}\right\rangle_V - \left\langle\hat{\Gamma}^\dagger_{\Psi i}\hat{H}\right\rangle_V = \left(\Lambda - \left\langle\hat{H}\right\rangle_V - \sum_{j=1}^{N_p}\delta\alpha_j\left\langle\hat{\Gamma}_{\Psi j}\right\rangle_V\right)\left\langle\hat{\Gamma}^\dagger_{\Psi i}\right\rangle_V + \sum_{j=1}^{N_p}\delta\alpha_j\left\langle\hat{\Gamma}^\dagger_{\Psi i}\hat{\Gamma}_{\Psi j}\right\rangle_V. \tag{2.39}$$

At this point Λ disappears from these N_p equations:

$$\sum_{j=1}^{N_p}\delta\alpha_j\left(\left\langle\hat{\Gamma}^\dagger_{\Psi i}\hat{\Gamma}_{\Psi j}\right\rangle_V - \left\langle\hat{\Gamma}^\dagger_{\Psi i}\right\rangle_V\left\langle\hat{\Gamma}_{\Psi j}\right\rangle_V\right) = \left\langle\hat{H}\right\rangle_V\left\langle\hat{\Gamma}^\dagger_{\Psi i}\right\rangle_V - \left\langle\hat{\Gamma}^\dagger_{\Psi i}\hat{H}\right\rangle_V. \tag{2.40}$$

Finally, the $\delta\alpha_j$ have been cast as the solutions of a system of N_p linear equations, which can be represented easily as:

$$\boldsymbol{S}\cdot\delta\boldsymbol{\alpha} = \boldsymbol{f}, \tag{2.41}$$

[6] In the limit of infinite applications of the operator $\left(\Lambda\mathbb{1}-\hat{H}\right)$ it becomes the exact ground state.

where the $\boldsymbol{S}$ matrix and the *pseudo-force* vector $\boldsymbol{f}$ are defined by comparison with equation (2.40):

$$S_{ij} = \langle \hat{\Gamma}^{\dagger}_{\Psi i} \hat{\Gamma}_{\Psi j} \rangle_V - \langle \hat{\Gamma}^{\dagger}_{\Psi i} \rangle_V \langle \hat{\Gamma}_{\Psi j} \rangle_V ; \tag{2.42}$$

$$f_i = \langle \hat{H} \rangle_V \langle \hat{\Gamma}^{\dagger}_{\Psi i} \rangle_V - \langle \hat{\Gamma}^{\dagger}_{\Psi i} \hat{H} \rangle_V . \tag{2.43}$$

Note that $\boldsymbol{S}$ is the covariance matrix of the logarithmic derivatives, hence it is symmetric and positive-semidefinite. The expectation values are computed during the simulation and subsequently the equations are solved numerically. The parameters are changed according to the substitution:

$$\boldsymbol{\alpha}' = \boldsymbol{\alpha} + \delta\tau \, \boldsymbol{S}^{-1} \cdot \boldsymbol{f} \tag{2.44}$$

and the new wave function is obtained. It can be seen that the parameter $\delta\tau$ is related to the variation of $\delta\alpha_0$ and thus depends on Λ; this parameter may vary during the simulation, but it is known in analogy with the SD method that picking $\delta\tau$ small enough ensures convergence. $\delta\tau$ controls the rate of change of the parameters: small values of $\delta\tau$ will lead to a slow but steady convergence, an optimal value leads to rapid convergence, whereas too large values make the algorithm unstable. Given a small enough $\delta\tau$, the variation in energy is:

$$\begin{aligned} \delta E = \delta\boldsymbol{\alpha} \cdot \boldsymbol{\nabla}_{\boldsymbol{\alpha}} E = \delta\tau \sum_{i,k} f_i \left(S^{-1} \right)_{i,k} \frac{\partial E}{\partial \alpha_k} = -\delta\tau \sum_{i,k} f_i \left(S^{-1} \right)_{i,k} f_k = \\ = -\delta\tau \left(\boldsymbol{f}, \boldsymbol{S}^{-1} \cdot \boldsymbol{f} \right) . \end{aligned} \tag{2.45}$$

This quantity is either negative or zero, since it is a negative number times the quadratic form of a positive-semidefinite matrix. In order to cure cases with vanishing eigenvalues, which could happen if one or more parameters are not independent and would lead to zero variation in energy, a small c-number ϵ is added to the $\boldsymbol{S}$ matrix, so that equation (2.44) becomes:

$$\boldsymbol{\alpha}' = \boldsymbol{\alpha} + \delta\tau \left(\epsilon \, \mathbb{I} + \boldsymbol{S}^{-1} \right) \cdot \boldsymbol{f} . \tag{2.46}$$

It can be seen that the SR algorithm is nothing but a generalization of the SD method which takes into account the metric of the manifold of variational states [51]. The optimization method is stable whenever the new parameters are close to the old ones according to some distance. For the SD method, this distance is just the Cartesian distance between two sets of parameters, whilst in SR the metric is determined by $\boldsymbol{S}$ as $d(\boldsymbol{\alpha}, \boldsymbol{\alpha}') \equiv (\boldsymbol{\alpha}, \boldsymbol{S} \cdot \boldsymbol{\alpha}')$; the improvement is clear, since a small variation of some parameters in SD could give rise to a huge modification of the energy: this effect is corrected in SR by means of the matrix $\boldsymbol{S}$. On the other hand, the SD method is recovered in the $\epsilon = \infty$ limit, where the matrix $\boldsymbol{S}$ is disregarded entirely.

3

The Variational Monte Carlo method for the two-band model

The Variational Monte Carlo (VMC) method has been applied to many different physical systems in Condensed Matter Physics, ranging from the electron gas [52] to liquid helium [53]. In the field of strongly-correlated electron systems, VMC enables to gain insight on properties of many materials, be them quantum spin liquids [54] or cuprate superconductors [55]. One great strength of VMC for strongly-correlated systems is that it is a non-perturbative method and thus it is not restrained to some small value of the couplings (for instance the Hubbard U), but it is meaningful even away from the limiting behaviors. The first and most important issue in VMC is choosing the correct wave function for the system: the typical wave functions for this class of systems are strictly related to Anderson's Resonating Valence Bond (RVB) theory [27] and thus reflect the underlying physics related to the lack of spontaneous magnetization and presence of superconductivity. The properties of such wave functions are nicely reviewed in reference [30]. Here they are applied to the two-band Hubbard model, whose Hamiltonian is equation (1.12); a more technical discussion is found in reference [56], where the case of the bilayer Hubbard model is treated. The first section of the present chapter introduces the variational wave function itself along with its properties, while the remaining two sections deal with certain details of the algorithm for this specific wave function.

3.1 The trial wavefunction

The wave function should encode all the physics of the system. The goal is to describe a strongly-correlated system, in particular a two-band model on a square lattice with L sites, including its superconducting behavior: the ingredients to construct the wave function must contain correlations and superconductive pairing. This paradigm is achieved by starting from a *mean-field* state $|\Phi_{MF}\rangle$, whose many-body wave function is a Slater determinant, and by applying a generalized projector operator $\mathcal{P}_J$ which includes the correlations:

$$|\Psi_V\rangle = \mathcal{P}_J \, |\Phi_{MF}\rangle \, . \tag{3.1}$$

The mean-field state can be always taken as the ground-state of a mean-field Hamiltonian, which intrinsically describes non-interacting electrons and thus does not include any correlation. The mean-field Hamiltonian has the form:

$$\begin{aligned} H_{MF} = &- \sum_{i,j} \sum_{\alpha} \sum_{\sigma} \tilde{t}^{\alpha}_{i,j} \left(c^{\dagger}_{i,\alpha,\sigma} c_{j,\alpha,\sigma} + \text{h.c.} \right) + \sum_{\alpha} \mu^{\alpha} N_{\alpha} + \\ &+ \sum_{i,j} \sum_{\alpha} \left[\Delta^{\alpha}_{i,j} \left(c^{\dagger}_{i,\alpha,\uparrow} c^{\dagger}_{j,\alpha,\downarrow} + c^{\dagger}_{j,\alpha,\uparrow} c^{\dagger}_{i,\alpha,\downarrow} \right) + \text{h.c.} \right] + \\ &+ \sum_{i,j} \sum_{\substack{\alpha,\beta \\ \alpha \neq \beta}} \left[\Delta^{\perp}_{i,j} \left(c^{\dagger}_{i,\alpha,\uparrow} c^{\dagger}_{j,\beta,\downarrow} + c^{\dagger}_{j,\beta,\uparrow} c^{\dagger}_{i,\alpha,\downarrow} \right) + \text{h.c.} \right] , \end{aligned} \tag{3.2}$$

or its equivalent k-space representation:

$$\begin{aligned} H_{MF} = &\sum_{k} \sum_{\alpha} \sum_{\sigma} \xi^{\alpha}_{k} c^{\dagger}_{k,\alpha,\sigma} c_{k,\alpha,\sigma} + \sum_{k} \sum_{\alpha} \Delta^{\alpha}_{k} \left(c^{\dagger}_{k,\alpha,\uparrow} c^{\dagger}_{-k,\alpha,\downarrow} + \text{h.c.} \right) + \\ &+ \sum_{k} \sum_{\substack{\alpha,\beta \\ \alpha \neq \beta}} \Delta^{\perp}_{k} \left(c^{\dagger}_{k,\alpha,\uparrow} c^{\dagger}_{-k,\beta,\downarrow} + \text{h.c.} \right) . \end{aligned} \tag{3.3}$$

$c_{i,\alpha,\sigma}$ (or $c_{k,\alpha,\sigma}$ in k-space) are the usual one-particle annihilation operators at site i, band α and spin σ and N_{α} is the total number of electrons in band α. The sum over the band indices with $\alpha \neq \beta$ includes the cases $\alpha = 1, \beta = 2$ and $\alpha = 2, \beta = 1$. ξ^{α}_{k} is the free band dispersion energy for the band α, which is taken as the usual two-dimensional band dispersions with nearest-neighbor hopping matrix elements $\tilde{t}^{\alpha}$ and chemical potentials μ^{α}, namely:

$$\xi^{\alpha}_{k} = -2\tilde{t}^{\alpha} \left(\cos k_x + \cos k_y \right) + \mu^{\alpha} . \tag{3.4}$$

$\Delta^{\alpha}_{k}, \Delta^{\perp}_{k}$ are intra-band and inter-band BCS pairing parameters, which include the superconducting behavior in the wave function. In particular, in order to describe unconventional superconductivity, the intra-band BCS couplings should have $d_{x^2-y^2}$-wave symmetry, which for nearest-neighbor pairing only would imply the following k-space form:

$$\Delta^{\alpha}_{k} = 2\Delta^{\alpha} \left(\cos k_x - \cos k_y \right) . \tag{3.5}$$

The ground-state of (3.2) is general and can take into account different scenarios. Eventually, an on-site (s-wave) singlet pairing can be allowed as an alternative (this would correspond to a constant pairing in k-space). On the other hand, the inter-band coupling $\Delta_k^{\perp}$ may either have the same nearest-neighbor d-wave symmetry of the intra-band ones or an on-site s-wave symmetry which allows both for singlet and triplet couplings, the latter being obtained by modifying the inter-band term of the real space mean-field Hamiltonian as:

$$\sum_{i,j} \sum_{\substack{\alpha,\beta \\ \alpha \neq \beta}} \left[\Delta_{i,j}^{\perp} \left(c_{i,\alpha,\uparrow}^{\dagger} c_{j,\beta,\downarrow}^{\dagger} - c_{j,\beta,\uparrow}^{\dagger} c_{i,\alpha,\downarrow}^{\dagger} \right) + \text{h.c.} \right] . \tag{3.6}$$

As a matter of fact, the mean-field Hamiltonian is assumed to be characterized by few, short-range parameters, which reflect in a manifest way the physical properties encoded in the wave function.

In order to diagonalize the Hamiltonian and cast its ground-state wave function into a Slater determinant form, it is necessary to introduce the *particle-hole* canonical transformation, which allows the mean-field Hamiltonian to be written in a quadratic form. The particle-hole transformation is achieved by defining a new set of creation and annihilation operators in the following way:

$$d_{i,\alpha,\uparrow} \equiv c_{i,\alpha,\uparrow} \tag{3.7a}$$

$$d_{i,\alpha,\downarrow} \equiv c_{i,\alpha,\downarrow}^{\dagger} , \tag{3.7b}$$

which can be easily proved to provide a canonical transformation, *i.e.* the canonical anti-commutation relations are preserved. Moreover, this implies that the spectrum of H_{MF} will be preserved by the transformation as well. These new creation-annihilation operators describe a picture of spin-$\uparrow$ electrons and spin-$\downarrow$ holes. The new vacuum state is the state with all sites occupied by spin down electrons:

$$\left| \tilde{0} \right\rangle = \prod_{i,\alpha} c_{i,\alpha,\downarrow}^{\dagger} \left| 0 \right\rangle \tag{3.8}$$

and it can be verified that $\left| \tilde{0} \right\rangle$ is annihilated by applying any of the $d_{i,\alpha,\sigma}$ operators. Acting on this vacuum state with the new d-operators, the one-electron basis $\mathcal{B}_d = \left\{ d_{i,\alpha,\sigma}^{\dagger} \left| \tilde{0} \right\rangle \right\}$ is obtained. It can be seen that the local z-axis spin polarization and occupation numbers exchange roles when the transformation is applied: if the new densities $\tilde{n}_{i,\alpha,\sigma}$ are expressed as:

$$\tilde{n}_{i,\alpha,\uparrow} = n_{i,\alpha,\uparrow} \tag{3.9a}$$

$$\tilde{n}_{i,\alpha,\downarrow} = 1 - n_{i,\alpha,\downarrow} , \tag{3.9b}$$

the following relations between the original z-axis spin polarization (defined along with equation (1.3)) and densities $m_{i,\alpha}, n_{i,\alpha}$ and the particle-hole ones $\tilde{m}_{i,\alpha}, \tilde{n}_{i,\alpha}$ hold:

$$\tilde{n}_{i,\alpha} = m_{i,\alpha} + 1 \tag{3.10a}$$

$$\tilde{m}_{i,\alpha} \equiv n_{i,\alpha} - 1 . \tag{3.10b}$$

Note that the total numbers of spin-↑ particles stays the same after applying the transformation, but the total number of spin-↓ holes does not:

$$\tilde{N}_{\alpha,\uparrow} = N_{\alpha,\uparrow} \tag{3.11a}$$

$$\tilde{N}_{\alpha,\downarrow} = L - N_{\alpha,\downarrow}\,. \tag{3.11b}$$

The half-filling condition for electrons is represented by the very same condition for spin-↑ particles and spin-↓ holes. However, this is true only at half-filling: suppose one wants to remove $4k$ electrons in order to look for superconductivity out of the half-filled case (global hole-doping $x > 0$) and consider the simplest case of a non-polarized ($S^z = 0$) system with evenly doped bands. In this case, the total numbers of spin-σ electrons in the α-band satisfies $N_{\alpha,\sigma} = L/2 - k$. According to the particle-hole transformation, this maps into the particle-hole spin-polarized case with $\tilde{N}_{\alpha,\uparrow} = L/2 - k$ and $\tilde{N}_{\alpha,\downarrow} = L/2 + k$[1].

It is now useful to arrange the transformed creation operators into a $4L$-dimensional vector, where the creation operators of spin-↑ particles in the first band come first, then those of spin-↑ particles in the second band and finally those for spin-↓ holes in the first and second bands. Introducing the multi-index $I \equiv \{i, \alpha, \sigma\}$, which also allows to write the states of $\mathcal{B}_d$ as $d_I^\dagger \left|\tilde{0}\right\rangle \equiv \left|I\right\rangle$, the Hamiltonian can be cast as:

$$H_{MF} = \sum_{I,J} d_I^\dagger \left(\boldsymbol{H}_{MF}\right)_{IJ} d_J = \left(\boldsymbol{d}^\dagger, \boldsymbol{H}_{MF} \cdot \boldsymbol{d}\right)\,, \tag{3.12}$$

where, explicitly:

$$\boldsymbol{d}^\dagger \equiv \left(d_{1,1,\uparrow}^\dagger, \ldots, d_{L,1,\uparrow}^\dagger, d_{1,2,\uparrow}^\dagger, \ldots, d_{1,1,\downarrow}^\dagger, \ldots, d_{L,1,\downarrow}^\dagger, d_{1,2,\downarrow}^\dagger \ldots d_{L,2,\downarrow}^\dagger\right)\,; \tag{3.13}$$

$$\boldsymbol{H}_{MF} \equiv \left(\begin{array}{c|c} \boldsymbol{\xi} & \boldsymbol{\Delta} \\ \hline \boldsymbol{\Delta} & -\boldsymbol{\xi} \end{array}\right)\,. \tag{3.14}$$

The diagonal form of H_{MF} and its eigenvalues are easily obtained via a unitary transformation $\boldsymbol{U}$ of the annihilation operators, which are mapped into a new set according to $\gamma_\mu = \boldsymbol{U}_{\mu I} d_I$. The explicit form of the matrix $\boldsymbol{U}$ is:

$$\begin{pmatrix} \langle 1 \,|\, \phi_1\rangle & \ldots & \langle 1 \,|\, \phi_{4L}\rangle \\ \vdots & \ddots & \vdots \\ \langle 4L \,|\, \phi_1\rangle & \ldots & \langle 4L \,|\, \phi_{4L}\rangle \end{pmatrix}\,, \tag{3.15}$$

where the basis of eigenstates $\mathcal{B}_\gamma = \left\{\left|\phi_\mu\right\rangle = \gamma_\mu^\dagger \left|\tilde{0}\right\rangle\right\}$ has been introduced. The states $\left|\phi_\mu\right\rangle$ are assumed to be arranged according to their corresponding energy eigenvalues ϵ_μ, ϵ_1 being the lowest one. The diagonal form of the mean-field Hamiltonian is then:

$$H_{MF} = \sum_\mu \epsilon_\mu \gamma_\mu^\dagger \gamma_\mu. \tag{3.16}$$

[1] Indeed, this implies doping with holes.

If there are N_e electrons in the system, the mean-field state $|\Phi_{MF}\rangle$ is obtained by acting with the first N_p of the $\gamma_\mu^\dagger$ operators on the vacuum $|\tilde{0}\rangle$:

$$|\Phi_{MF}\rangle = \prod_{\mu=1}^{N_e} \gamma_\mu^\dagger |\tilde{0}\rangle . \tag{3.17}$$

Since $|\Phi_{MF}\rangle$ is a many-body state of independent fermions, its wave function is a Slater determinant. It must be stressed that $|\Phi_{MF}\rangle$ has a parametric dependence, which has been left implied, on the mean field couplings $\tilde{t}_{i,j}^{\alpha}$, $\Delta_{i,j}^{\alpha}$ and $\Delta_{i,j}^{\perp}$ introduced with equation (3.2).

Up to this point, no correlation has been included into the trial wave function and indeed the Hamiltonian has been cast in an independent-particle (quadratic) form. Correlations are included by means of the so-called Jastrow operator $\mathcal{P}_J$ acting on the determinantal state. Jastrow factors were introduced by R. Jastrow [57] as a means to improve the single-particle description of a system of interacting particles by adding pair contributions. The Jastrow operator can be easily described, as it is diagonal in the local occupation-number basis. Its form is:

$$\mathcal{P}_J \equiv \exp\left(\sum_{i\leq j}\sum_{\alpha,\alpha'} J_{i,j}^{\alpha,\alpha'} n_{i,\alpha} n_{j,\alpha'}\right). \tag{3.18}$$

The simplest Jastrow operator is that containing only on-site terms, which is a generalized Gutzwiller operator:

$$\mathcal{P}_G \equiv \exp\left(\sum_{i}\sum_{\alpha,\alpha'} g_i^{\alpha,\alpha'} n_{i,\alpha} n_{j,\alpha'}\right) \tag{3.19}$$

which reduces to the Gutzwiller projector defined in (1.7) in the limit of $g \to -\infty$. If one seeks to represent states which have traslational symmetry, the Jastrow factors must depend on the lattice sites only as a function of the lattice distance $|i-j| = |\boldsymbol{R}_i - \boldsymbol{R}_j|$. As a generalization of the simple Gutzwiller, a repulsive Jastrow operator, namely one with a negative $J_{i,j}^{\alpha,\alpha'}$, attenuates the configurations which have particles at site i of band α and at site j of band α', whereas the attractive case with positive $J_{i,j}^{\alpha,\alpha'}$ would imply that such configurations are promoted (which would be the case for the attractive Hubbard model). In practice, what one usually sees is that the Jastrow parameters $J_{i,j}^{\alpha,\alpha'}$ are all negative and that they decay with increasing $|i-j|$, which means that the effect of correlations becomes weaker at larger distances. Jastrow factors can be easily rewritten in terms of particle-hole operators via equation (3.10b):

$$\begin{aligned}
\mathcal{P}_J &= \exp\left(\sum_{i\leq j}\sum_{\alpha,\alpha'} J_{i,j}^{\alpha,\alpha'} n_{i,\alpha} n_{j,\alpha'}\right) = \\
&= \exp\left(\sum_{i\leq j}\sum_{\alpha,\alpha'} J_{i,j}^{\alpha,\alpha'} (\tilde{m}_{i,\alpha}+1)(\tilde{m}_{j,\alpha'}+1)\right) = \\
&= \exp\left(\sum_{i\leq j}\sum_{\alpha,\alpha'} J_{i,j}^{\alpha,\alpha'} \tilde{m}_{i,\alpha}\, \tilde{m}_{j,\alpha'}\right) \exp\left(\sum_{i\leq j}\sum_{\alpha,\alpha'} J_{i,j}^{\alpha,\alpha'} (\tilde{m}_{i,\alpha} + \tilde{m}_{j,\alpha'} + 1)\right).
\end{aligned} \tag{3.20}$$

In the last line of this equation only the first factor is relevant, since the second one can be shown to be independent from the configuration, once the number of particles has been fixed. There is an easy argument to see that Jastrow factors are a major improvement to the Gutzwiller: once a holon-doublon pair is produced in a one-band Hubbard model the energy cost for driving the single quasi-particles apart (*e.g.* via an external electric field) is zero. Thus, a wave function with a finite Gutzwiller operator alone cannot describe a Mott insulator properly. On the other hand, a long-ranged Jastrow projector can bind the holon-doublon pair [58].

Magnetic terms such as an antiferromagnetic coupling to an external field could have been included in the mean-field Hamiltonian as well. However, in the spirit of the RVB theory of Mott insulators, such terms have been neglected: the rationale behind this omission is that magnetic order may be suppressed by frustration, thereby making the RVB spin-liquid the favored state, but how such frustration arises is not directly taken into account for the sake of simplicity. At the same time, a small non-zero temperature destroys the magnetic order in two dimensions. Moreover, hole-doping the system also favors the RVB state: in the perspective of studying superconductivity, it is useful to study the pairing symmetry in the half-filled case. Possible developments of the free part of the wave function include backflow terms for correlations [59] and a pfaffian as an alternative and generalization to the determinantal part [60].

3.1.1 Physical properties of the wave function: correlation functions

From the wave function $|\Psi_V\rangle$ it is possible to extract any static[2] correlation function of interest, according to equation (2.3). One of the most important static correlation functions is the (intra-band) *density-density* correlation function:

$$N_\alpha(\boldsymbol{q}) \equiv \frac{\langle\Psi_V|\, n_\alpha(-\boldsymbol{q})\, n_\alpha(\boldsymbol{q})\, |\Psi_V\rangle}{\langle\Psi_V|\Psi_V\rangle}\,, \tag{3.21}$$

where $n_\alpha(\boldsymbol{q})$ is the Fourier transform of the occupation number operator of band α:

$$n_\alpha(\boldsymbol{q}) \equiv \frac{1}{\sqrt{L}} \sum_i \sum_\sigma \exp(i\boldsymbol{q}\cdot\boldsymbol{R}_i)\, n_{i,\alpha,\sigma}\,. \tag{3.22}$$

The correlation function $N_\alpha(\boldsymbol{q})$ quantitatively encodes the information about density fluctuations with momentum $\boldsymbol{q}$. If this correlation function diverges[3] for a value of momentum $\boldsymbol{q}^*$, then the density of the system experiences some kind of density ordering: for instance, if $\boldsymbol{q}^* = (\pi, \pi)$ the α-band displays a charge-density-wave, since charge density is proportional to

[2] Dynamical correlations cannot be accessed, because VMC is a ground-state method and time evolution is not considered. Nonetheless, many physical properties can be deduced from static correlation alone.

[3] In place of a divergence, the correlation function would attain a value proportional to the system size L. An actual divergence cannot occur on a finite-size system.

the number density, not unlike a divergence in the *spin-spin* correlation function at $\boldsymbol{q}^* = (\pi, \pi)$ signals antiferromagnetic ordering.

Superconductive pairing correlation functions are as important as the density-denstity ones. To define a pairing correlation function, introduce the singlet pair creation operator (with $\hat{\mu}, \hat{\nu}$ being lattice unit vectors):

$$B^\dagger_{i,\alpha,\hat{\mu}} \equiv c^\dagger_{i,\alpha,\uparrow} c^\dagger_{i+\hat{\mu},\alpha,\downarrow} - c^\dagger_{i,\alpha,\downarrow} c^\dagger_{i+\hat{\mu},\alpha,\uparrow} \,. \tag{3.23}$$

Then, the pairing correlation function is:

$$C^{\hat{\mu},\hat{\nu}}_\alpha(|i-j|) \equiv \frac{\langle \Psi_V | \, B_{j,\hat{\mu}} B^\dagger_{i,\hat{\nu}} \, | \Psi_V \rangle}{\langle \Psi_V \, | \, \Psi_V \rangle}. \tag{3.24}$$

Superconductive pairing is related to the presence of off-diagonal long-range order (ODLRO) in the corresponding correlation functions (or the two-body density matrix), that is, the limit of $C^{\hat{\mu},\hat{\nu}}_\alpha(|i-j|)$ for large $|i-j|$ is non-zero[4]. The square root of such value, name it Φ_α, is a measure of the *pseudo-condensate* fraction [61, 62]:

$$(\Phi_\alpha)^2 \equiv \lim_{|i-j|\to+\infty} C^{\hat{\mu},\hat{\nu}}_\alpha(|i-j|) \,. \tag{3.25}$$

3.1.2 Physical properties of the wave function: charge gap

When discussing metal-insulator transitions, the most relevant physical properties are those related to the conducting behavior of the material. Conductivity is a dynamical quantity and thus it cannot be accessed by VMC; therefore, an alternative way to discriminate between an insulator and a metal is needed. This is possible within the Single-Mode Approximation (SMA), which was first employed by R. Feynman to give a microscopic description of excitations in ^{4}He [63] and was later extended to fermionic systems [64]. The rationale behind the SMA is that an excited state can be obtained by applying an excitation operator to the ground-state wave function. More preciesly, Feynman showed that, being ϕ the ground-state wave function for a system of identical bosons, $\psi = F\phi$ is the best variational excited state for $F = \sum_i \exp(i\boldsymbol{q} \cdot \boldsymbol{r}_i)$, $\boldsymbol{r}_i$ being the position of the i-th particle. This state depends indeed upon the momentum $\boldsymbol{q}$ and indeed it is an eigenstate of the total momentum operator, indicating that it gives a non-zero density current. From these grounds, the SMA gives a qualitative atomic description of experimentally observed features of ^{4}He, such as the structure factor and the spectrum of elementary excitations. In this context, the main assumption is that this argument would work for a system of strongly interacting electrons on a lattice as well. Moreover, the same structure assumed by Feynman, namely an excited state obtained by acting on the exact ground state, is believed to be present if one builds the excited state starting from the variational trial state; in the case of a functional form (3.1) with long-ranged Jastrow factors,

[4] In a d-wave superconductor, in case of $\hat{\mu}$ and $\hat{\nu}$ being two different unit vectors, the absolute value of $C^{\hat{\mu},\hat{\nu}}_\alpha(|i-j|)$ must be taken, since it is a negative quantity according to the pairing symmtery.

the Feynman construction was shown to be consistent [65]. Under these assumptions, it is possible to verify whether a state is insulating or metallic [59]. To see how this is possible, it is useful to cast Feynman's excitation operator in second quantization form, namely as the Fourier transform of the density operator (equation (3.22)). Acting on the trial wave function, the excited state $|\Psi_V(\boldsymbol{q},\alpha)\rangle \equiv n_\alpha(\boldsymbol{q})\,|\Psi_V\rangle$ is obtained. The excitation energy of such a state is defined as the energy difference between lowest variational energy E_V and the expectation value of the excited state, mathematically:

$$\mathcal{E}(\boldsymbol{q}) \equiv \frac{\langle\Psi_V(\boldsymbol{q},\alpha)|\,H\,|\Psi_V(\boldsymbol{q},\alpha)\rangle}{\langle\Psi_V(\boldsymbol{q},\alpha)\,|\,\Psi_V(\boldsymbol{q},\alpha)\rangle} - E_V\,. \tag{3.26}$$

With a little algebra it can be seen that if E_V is the variational minimum of the energy, the excitation energy turns out to be (see supplementary material to reference [65]):

$$\mathcal{E}(\boldsymbol{q}) \equiv \frac{1}{2\,N_\alpha(\boldsymbol{q})}\,\frac{\langle\Psi_V|\,\big[n_\alpha(-\boldsymbol{q}),[H,n_\alpha(\boldsymbol{q})]\,\big]\,|\Psi_V\rangle}{\langle\Psi_V\,|\,\Psi_V\rangle}\,, \tag{3.27}$$

having expressed the denominator in terms of the density-density correlation function $N_\alpha(\boldsymbol{q})$. The microscopic quantity that encodes the conducting properties of the system is the charge-excitation gap, which is the limit for low momentum of the excitation energy. Just as in the case of ^{4}He, the excited state has a non-zero momentum, which is now associated with a charge current; if the charge-excitation gap vanishes, a conducting state can be established by paying an infinitesimal amount of energy, which is the case of a conductor. It can be seen that the charge gap for a given band turns out to be proportional to:

$$\mathcal{E}_g = \lim_{|\boldsymbol{q}|\to 0}\mathcal{E}(\boldsymbol{q}) \propto \lim_{|\boldsymbol{q}|\to 0}\frac{|\boldsymbol{q}|^2}{N_\alpha(\boldsymbol{q})}\,. \tag{3.28}$$

In other words, the insulating or metallic behavior of the wave function can be discriminated by looking at the low-momentum behavior of the density-density correlation function: a quadratic behavior of $N_\alpha(\boldsymbol{q})$ implies that charge excitations are gapped, *i.e.* an insulating behavior; on the contrary, a linearly vanishing $N_\alpha(\boldsymbol{q})$ reflects the presence of gapless excitations, meaning that the wave function is metallic. The presence of $|\boldsymbol{q}|^2$ at the numerator of the last term reflects the fact that $n_\alpha(\boldsymbol{q})$ commutes with H at $|\boldsymbol{q}| = 0$.

The behavior of the density-density correlation function is strictly related to that of the Jastrow parameters inside the wave function. This relation has been studied for the one-dimensional [66] and two-dimensional [67] Hubbard model, and for its bosonic counterpart as well [68]. Once again, studies on ^{4}He give insight in the case of strongly-correlated electron systems: the structure factor (here the density-density correlation function) of a wave function of the form (3.1) can be written in terms of the structure factor $N^0_\alpha(\boldsymbol{q})$ of the free part alone according to the relation [69]:

$$N_\alpha(\boldsymbol{q}) = \frac{N^0_\alpha(\boldsymbol{q})}{1 + 2J^{\alpha,\alpha}(\boldsymbol{q})\,N^0_\alpha(\boldsymbol{q})}\,, \tag{3.29}$$

$J_{\alpha,\alpha}(\boldsymbol{q})$ being the Fourier transform of the Jastrow factors. This RPA-like formula, which holds in the two-band model for non-hybridized bands, allows one to deduce the conducting behavior of the wave function just by looking at $J^{\alpha,\alpha}(\boldsymbol{q})$: since $N^0_\alpha(\boldsymbol{q})$ approaches a constant value at small $|\boldsymbol{q}|$, if $J^{\alpha,\alpha}(\boldsymbol{q})$ diverges as $|\boldsymbol{q}|^{-2}$, $N_\alpha(\boldsymbol{q})$ will vanish quadratically[5], whereas if $J^{\alpha,\alpha}(\boldsymbol{q})$ diverges as $|\boldsymbol{q}|^{-1}$, $N_\alpha(\boldsymbol{q})$ will be linear.

3.2 Evaluation of acceptances and local values

In order to generate new configurations via the Metropolis algorithm, it is necessary to compute the ratio between the stationary probabilities of equation (2.28), which amounts to the square modulus of the ratio between the overlaps of the trial wave function with the new configuration and the old configuration:

$$\frac{p(x')}{p(x)} = \left| \frac{\langle x' \,|\, \Psi_V \rangle}{\langle x \,|\, \Psi_V \rangle} \right|^2 = \left| \frac{\mathcal{P}_J(x')}{\mathcal{P}_J(x)} \, \frac{\langle x' \,|\, \Phi_{MF} \rangle}{\langle x \,|\, \Phi_{MF} \rangle} \right|^2 . \tag{3.30}$$

Here $|x\rangle = |x_1, \ldots, x_{N_e}\rangle$ is the configuration of N_e electrons occupying the levels $|x_i\rangle$, each of them belonging to the base $\mathcal{B}_d$ and $\mathcal{P}(x)$ is the Jastrow potential evaluated on $|x\rangle$. Electrons are fermions, thus each x_i can appear only once in any configuration $|x\rangle$. The new configuration $|x'\rangle$ is generated from the previous one by moving one or more electrons according to processes enabled by the Hamiltonian, for example by hopping it to a nearest neighbor site. However, it is possible to allow moves which are not present in the Hamiltonian, such as the spin-flip of two neighboring electrons, in order to drive the system towards ergodicity. Similarly, another process that may be allowed is the on-site hopping to the other band; having contracted the site and band degrees of freedom into the same index, this kind of process can be treated just as a standard hopping, because of the similarity with a bilayer Hubbard model. Another process which is allowed in the case of finite Hund's coupling is the spin-flip of an on-site singlet of electrons between the two bands, as well as the pair hopping of a singlet of electrons from one band to the other; these processes can be obtained by applying twice the hopping described previously.

Just as the wave function is made of the correlated and uncorrelated parts, this ratio factorizes into ratios between Jastrow factors and between Slater determinants, which take into account the antisymmetrization of the wave function. The former is easy to be computed in principle, since it is just a difference in the exponent, while the latter is a ratio between determinants, which is not as straightforward. The most important issue is that of computational complexity, because in principle the wave function $\langle x' \,|\, \Psi_V \rangle$ has to be computed each time a new configuration $|x'\rangle$ is generated. In practice this is not the case, because one can take advantage of some tricks which arise from the fact that consecutive configurations are correlated.

[5] For a 2-dimensional insulating state there are logarithmic corrections to the quadratic behavior of the Jastrow factors in momentum space, which diverges as $\frac{\log(|\boldsymbol{q}|)}{|\boldsymbol{q}|^2}$ [67].

This leads to a significative reduction in complexity and makes it practically possible to carry out the simulation.

Consider the mean-field part $\langle x \,|\, \Phi_{MF} \rangle$ first [58]. As anticipated, this overlap is a Slater determinant: indeed, by applying the explicit form of $|\Phi_{MF}\rangle$, equation (3.17), the wave function can be cast as the determinant of a $(N_e \times N_e)$-dimensional matrix $\boldsymbol{D}$:

$$\begin{aligned} \langle x \,|\, \Phi_{MF} \rangle = \langle x_1, \dots, x_{N_e} \,|\, \Phi_{MF} \rangle = \langle x_1, \dots, x_{N_e} | \, \gamma_1^\dagger \dots \gamma_{N_e}^\dagger \, |\tilde{0}\rangle = \\ = \det \begin{pmatrix} \langle x_1 \,|\, \phi_1 \rangle & \dots & \langle x_1 \,|\, \phi_{N_e} \rangle \\ \vdots & \ddots & \vdots \\ \langle x_{N_e} \,|\, \phi_1 \rangle & \dots & \langle x_{N_e} \,|\, \phi_{N_e} \rangle \end{pmatrix} = \det \boldsymbol{D} \end{aligned} \tag{3.31}$$

where $|\phi_1\rangle \dots |\phi_{N_e}\rangle$, defined with equation (3.15), are the occupied eigenstates; if $\boldsymbol{D'}$ is the same matrix for the primed $|x'\rangle$ configuration, then what has to be computed is the ratio:

$$\frac{\det \boldsymbol{D'}}{\det \boldsymbol{D}} \,. \tag{3.32}$$

It is useful to introduce the auxiliary $(4L \times N_e)$ rectangular matrix $\boldsymbol{M}$, defined by retaining only the N_e columns of the matrix $\boldsymbol{U}$ relative to the N_e lowest-energy states:

$$\boldsymbol{M} \equiv \begin{pmatrix} \langle 1 \,|\, \phi_1 \rangle & \dots & \langle 1 \,|\, \phi_{N_e} \rangle \\ \vdots & \ddots & \vdots \\ \langle 4L \,|\, \phi_1 \rangle & \dots & \langle 4L \,|\, \phi_{N_e} \rangle \end{pmatrix} . \tag{3.33}$$

Suppose now that the configuration $|x'\rangle$ is obtained from $|x\rangle$ by changing the configuration number x_β to the l-th level, which is the easiest possible move. The new matrix is obtained by specifically substituting a row of $\boldsymbol{D}$ with another from $\boldsymbol{M}$, specifically element-wise:

$$D'_{\alpha k} = D_{\alpha k} + \delta_{\alpha\beta} \left(M_{lk} - D_{\alpha k} \right) = D_{\alpha k} + \delta_{\alpha\beta} \, v_k^{l\beta} \,, \tag{3.34}$$

where $v_k^{l\beta}$ is defined by comparison in the last equation and represents the amount which has to be added to the β-th row of $\boldsymbol{D}$ in order to update it correctly; equivalently, the β-th row of $\boldsymbol{D}$ has to be replaced with the l-th row of $\boldsymbol{M}$. The previous equation can be manipulated in the following way:

$$D'_{\alpha k} = D_{\alpha k} + \delta_{\alpha\beta} \, v_k^{l\beta} = \sum_m D_{\alpha m} \left(\delta_{mk} + D^{-1}_{m\beta} v_k^{l\beta} \right) = \sum_m D_{\alpha m} K_{mk} \,, \tag{3.35}$$

where a new matrix $\boldsymbol{K}$ is defined by comparison in the last passage. This matrix is such that $\boldsymbol{D} \cdot \boldsymbol{K} = \boldsymbol{D'}$ and therefore using the property of the determinant of the product allows to write the determinant ratio of equation (3.32) as a determinant itself. However, $\boldsymbol{K}$ has a peculiar form: it is the identity matrix plus the tensor product of two vectors, and in this case the

determinant can be shown to satisfy $\det\left(\mathbb{I} + \boldsymbol{u} \otimes \boldsymbol{v}^T\right) = 1 + \boldsymbol{v} \cdot \boldsymbol{u}$. Thereby, understanding $D^{-1}_{m\beta}$ and $v^{l\beta}_k$ respectively as the m-th and k-th components of two vectors $\boldsymbol{D}^{-1}_{\beta}$ and $\boldsymbol{v}^{l\beta}$:

$$\begin{aligned}\frac{\det \boldsymbol{D'}}{\det \boldsymbol{D}} = \det \boldsymbol{K} = \det\left(\mathbb{I} + \boldsymbol{D}^{-1}_{\beta} \otimes \boldsymbol{v}^{l\beta}\right) = 1 + \sum_m D^{-1}_{m\beta} v^{l\beta}_m = \\ = 1 + \sum_m D^{-1}_{m\beta}\left(M_{lm} - D_{\alpha m}\right) = \sum_m D^{-1}_{m\beta} M_{lm}.\end{aligned} \tag{3.36}$$

Thus, the ratio of the determinants of two consecutive configurations such as $|x'\rangle$ and $|x\rangle$ is reduced to the entries of the matrix:

$$\boldsymbol{W} = \boldsymbol{M} \cdot \boldsymbol{D}^{-1}. \tag{3.37}$$

Having to look at the entries of $\boldsymbol{W}$ instead of computing each of the determinants and then taking their ratio may seem a great step towards efficiency. However up to this point this is essentially useless, because $\boldsymbol{W}$ needs to be computed each time the new configuration is accepted and this process involves the inversion of $\boldsymbol{D}$, which has to be achieved by solving a linear system[6]. The key point is that $\boldsymbol{W}$ has not to be computed from scratch, again because of the fact that $|x'\rangle$ resembles $|x\rangle$. Indeed, from the definition of $\boldsymbol{K}$, it can be seen that:

$$D'^{-1}_{k\alpha} = \sum_m K^{-1}_{km} D^{-1}_{m\alpha} \tag{3.38}$$

and with the help of the Sherman-Morrison formula it can be seen that the inverse of $\boldsymbol{K}$ involves the matrix $\boldsymbol{W}$:

$$K^{-1}_{km} = \delta_{km} - \frac{D^{-1}_{k\beta} v^{l\beta}_m}{1 + \sum_i D^{-1}_{i\beta} v^{l\beta}_i} = \delta_{km} - \frac{D^{-1}_{k\beta} v^{l\beta}_m}{W_{l\beta}}. \tag{3.39}$$

At this point, it is possible to write the expression for $\boldsymbol{W'} = \boldsymbol{M} \cdot \boldsymbol{D'}^{-1}$ just by applying the definition:

$$\begin{aligned}W'_{h\alpha} &= \sum_k M_{hk} D'^{-1}_{k\alpha} = \sum_k M_{hk} \left(\sum_m K^{-1}_{km} D^{-1}_{m\alpha}\right) \\ &= \sum_k M_{hk} D^{-1}_{k\alpha} - \sum_k M_{hk} \frac{D^{-1}_{k\beta}}{W_{l\beta}} \sum_m D^{-1}_{m\alpha} v^{l\beta}_m = \\ &= W_{h\alpha} - \frac{W_{h\beta}}{W_{l\beta}} \sum_m D^{-1}_{m\alpha} v^{l\beta}_m = W_{h\alpha} - \frac{W_{h\beta}}{W_{l\beta}} \sum_m D^{-1}_{m\alpha} \left(M_{lm} - D_{\beta m}\right) = \\ &= W_{h\alpha} - \frac{W_{h\beta}}{W_{l\beta}} \left(W_{l\alpha} - \delta_{\alpha\beta}\right).\end{aligned} \tag{3.40}$$

[6]Both solving a linear system and computing a determinant with standard methods scale as the third power of the order of the matrix.

If the move has been accepted, the matrix $\boldsymbol{W}$ is then updated with little effort[7]: it is therefore more efficient to store the matrix $\boldsymbol{W}$ alone, instead of doing any cumbersome calculation. Once in a while it is cautious to re-compute $\boldsymbol{W}$ from scratch in order to prevent the accumulation of round-off errors.

So far, only the Slater determinants has been treated. The remaining part of the ratio between the probabilities involves the Jastrow factors, which are easier to be dealt with. Expanding the Jastrow factors in terms of the particle-hole operators, their ratio becomes:

$$\begin{aligned}\frac{\mathcal{P}_J(x')}{\mathcal{P}_J(x)} &= \frac{\exp\left(\sum_{i\leq j}\sum_{\alpha,\alpha'} J_{i,j}^{\alpha,\alpha'}\,\tilde{m}_{i,\alpha}(x')\,\tilde{m}_{j,\alpha'}(x')\right)}{\exp\left(\sum_{i\leq j}\sum_{\alpha,\alpha'} J_{i,j}^{\alpha,\alpha'}\,\tilde{m}_{i,\alpha}(x)\,\tilde{m}_{j,\alpha'}(x)\right)} = \\ &= \exp\left(\sum_{i\leq j}\sum_{\alpha,\alpha'} J_{i,j}^{\alpha,\alpha'}\left(\tilde{m}_{i,\alpha}(x')\,\tilde{m}_{j,\alpha'}(x') - \tilde{m}_{i,\alpha}(x)\,\tilde{m}_{j,\alpha'}(x)\right)\right), \end{aligned} \tag{3.41}$$

where the magnetization operators $\tilde{m}_{i,\alpha}$ have been replaced by their values over the configurations and the last term of equation (3.20) has been dropped. Recall now that $|x'\rangle$ is obtained from configuration $|x\rangle$ by changing the configuration number from site k and band κ to site l and band λ: being the magnetization a local quantity, its change is limited only to the sites involved in the hopping process. Depending on the spin of the particle, the new magnetization is written in terms of the old one as:

$$\tilde{m}_{i,\alpha}(x') = \tilde{m}_{i,\alpha}(x) + \sigma\left(\delta_{il}\delta_{\alpha\lambda} - \delta_{ik}\delta_{\alpha\kappa}\right), \tag{3.42}$$

where $\sigma = \pm 1/2$ for spin-$\uparrow$ or spin-$\downarrow$ respectively. Plugging this formula into equation (3.41) gives:

$$\frac{\mathcal{P}_J(x')}{\mathcal{P}_J(x)} = \exp\left(\frac{1}{2}\,J_{l,l}^{\lambda,\lambda} + \frac{1}{2}\,J_{k,k}^{\kappa,\kappa} - J_{l,k}^{\lambda,\kappa} + \sum_{\alpha}\left(\sigma\,T_l^{\alpha,\lambda}(x) - \sigma\,T_k^{\alpha,\kappa}(x)\delta_{\alpha'\kappa}\right)\right), \tag{3.43}$$

with:

$$T_k^{\alpha,\alpha'}(x) = \sum_i J_{i,k}^{\alpha,\alpha'}\,\tilde{m}_{i,\alpha}(x). \tag{3.44}$$

The vector $\boldsymbol{T}^{\alpha,\alpha'}(x)$ plays the same role of the matrix $\boldsymbol{W}$ used for the determinantal part; the advantage is the same, it can be stored in place of other quantities and, upon acceptance, its update is computationally least expensive:

$$\begin{aligned} T_j^{\alpha,\alpha'}(x') &= \sum_i J_{i,j}^{\alpha,\alpha'}\,\tilde{m}_{i,\alpha}(x') = \sum_i J_{i,j}^{\alpha,\alpha'}\left(\tilde{m}_{i,\alpha}(x) + \sigma(\delta_{il}\delta_{\alpha\lambda} - \delta_{ik}\delta_{\alpha\kappa})\right) = \\ &= T_j^{\alpha,\alpha'}(x) + \sigma\left(J_{l,j}^{\alpha,\alpha'}\,\delta_{\alpha\lambda} - J_{k,j}^{\alpha,\alpha'}\,\delta_{\alpha\kappa}\right). \end{aligned} \tag{3.45}$$

[7] This operation scales as N_e^2, which is the number of elements of $\boldsymbol{W}$.

3.3 Computation of logarithmic derivatives

The last important step needed for the simulation is implementing the SR, which implies the calculation of the logarithmic derivatives of equation (2.35), where now the vector of the parameters $\boldsymbol{\alpha}$ is realized by the Jastrow parameters $J_{i,j}^{\alpha,\alpha'}$ and any other parameter included in $|\Phi_{MF}\rangle$. Again, the task splits between what concerns the Jastrow factors and the mean-field wave function. The logarithmic derivatives with respect to the Jastrow parameters $J_{i,j}^{\alpha,\alpha'}$ are straightforwardly computed, thanks to the fact that they are exponentiated. Taking into account that the Jastrow parameters are indeed a function only of $|i-j|$:

$$\begin{aligned}\frac{\partial \ln\left(\langle x\,|\,\Psi_V(\boldsymbol{\alpha})\rangle\right)}{\partial J_{|i-j|}^{\alpha,\alpha'}} &= \frac{\partial \ln\left(\langle x|\,\mathcal{P}_J\,|\Phi_{MF}\rangle\right)}{\partial J_{|i-j|}^{\alpha,\alpha'}} = \frac{\partial \ln\left(\mathcal{P}_J(x)\;\langle x\,|\,\Phi_{MF}\rangle\right)}{\partial J_{|i-j|}^{\alpha,\alpha'}} = \\ &= \langle x\,|\,\Phi_{MF}\rangle\;\frac{\partial}{\partial J_{|i-j|}^{\alpha,\alpha'}}\left(\sum_{k\leq l}\sum_{\gamma,\gamma'} J_{|k-l|}^{\gamma,\gamma'}\tilde{m}_{k,\gamma}(x)\,\tilde{m}_{l,\gamma'}(x)\right) = \\ &= \langle x\,|\,\Phi_{MF}\rangle\sum_{k\leq l}\sum_{\gamma,\gamma'}\delta_{|i-j|\,|k-l|}\,\delta_{\alpha\gamma}\,\delta_{\alpha'\gamma'}\,\tilde{m}_{k,\gamma}(x)\,\tilde{m}_{l,\gamma'}(x)\,,\end{aligned} \tag{3.46}$$

which corresponds to computing the sum of the products of the magnetization for all couples of sites at the fixed distance $|i-j|$.

To compute the derivative with respect to one of the mean-field parameters α_k, consider a small change $\alpha_k \to \alpha_k' = \alpha_k + \delta\alpha_k$ [51]. Since the mean-field Hamiltonian is linear in α_k, its change can be written as:

$$H_{MF} \to H'_{MF} = H_{MF} + \sum_{k=1}^{p}\delta\alpha_k\,V_k \tag{3.47}$$

where $V_k = \sum_{I,J}(\boldsymbol{V}_k)_{IJ}\,d_I^\dagger d_J$ is the operator associated to the parameter α_k in the mean-field Hamiltonian, *e.g.* if α_k is the nearest-neighbor hopping parameter $\tilde{t}^\alpha$, then the operator V_k will be the kinetic energy term of (3.2). If corrections to the mean-field state are considered up to first order in the variation $\delta\alpha_k$, the modified mean-field state is:

$$|\Phi'_{MF}\rangle = \left[1 + \sum_{k=1}^{p}\delta\alpha_k\sum_{\mu,\nu}(\boldsymbol{Q}_k)_{\mu\nu}\,\gamma_\mu^\dagger\gamma_\nu\right]|\Phi_{MF}\rangle\,, \tag{3.48}$$

where the matrix $\boldsymbol{Q}_k$ is defined by:

$$(\boldsymbol{Q}_k)_{\mu\nu} = \begin{cases}\dfrac{\left(\boldsymbol{U}^\dagger\,\boldsymbol{V}_k\,\boldsymbol{U}\right)_{\mu\nu}}{\epsilon_\mu - \epsilon_\nu} & \text{for } \mu > N_e \text{ , and } \nu \leq N_e \\ 0 & \text{otherwise,}\end{cases} \tag{3.49}$$

where $\boldsymbol{U}$ is the unitary matrix that diagonalizes $\boldsymbol{H}_{MF}$. By comparison with equation (2.34), it can be seen that the operator inside the brackets in equation (3.48) is nothing but the loga-

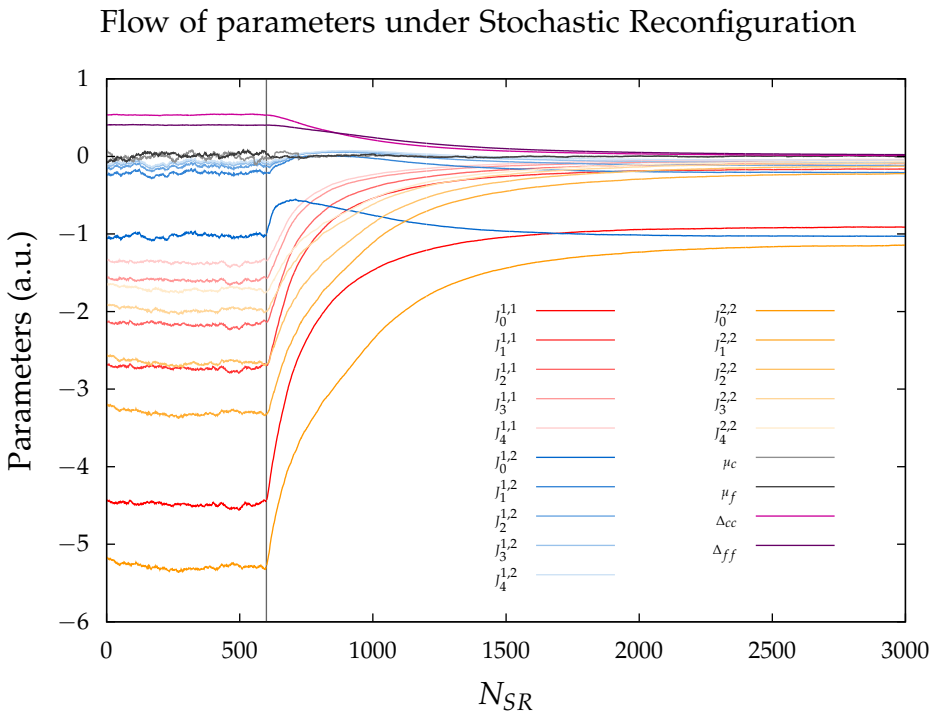


Figure 3.1: A practical implementation of the SR algorithm to a typical wave function used to describe a two-band model. Here $J_r^{\alpha,\alpha'}$ is the Jastrow factor between bands α and α' at the r-th shortest distance ($r = 0$ refers to the Gutzwiller factor, $r = 1$ to the nearest-neighbor Jastrow factor and so on), Δ_{cc} and Δ_{ff} are d-wave nearest-neighbor intra-band pairings. On the x-axis the number of SR steps is represented, while the y-axis represents the value of the parameters. The system starts at $N_{SR} = 0$ from a Mott insulating phase $U = 12, R = 0.7$ and the parameters are already optimized; here, intra-band Jastrow factors are large and could be seen to diverge as $|q|^{-2}$ and nearest-neighbor intra-band d-wave pairings are non-zero. At $N_{SR} = 600$ (grey vertical line) the two-band Hamiltonian is quenched into the metallic phase $U = 6, R = 0.7$: the insulating wave function does not describe properly the system anymore and, accordingly, the parameters flow to a set of values which corresponds to metallic behavior.

rithmic derivative operator $\hat{\mathbf{\Gamma}}_\Phi$[8]. The local value of this operator, which is needed to obtain its average, is easily computed by writing it in terms of the d operators:

$$\frac{\langle x|\,\hat{\mathbf{\Gamma}}_{\Phi k}\,|\Phi_{MF}\rangle}{\langle x\,|\,\Phi_{MF}\rangle} = \sum_{I,J} (\boldsymbol{R}_k)_{IJ}\,\frac{\langle x|\,d_I^\dagger d_J\,|\Phi_{MF}\rangle}{\langle x\,|\,\Phi_{MF}\rangle}\,, \tag{3.50}$$

where the matrix $\boldsymbol{R}_k$ is obtained by rotating the matrix $\boldsymbol{Q}_k$ back to the basis $\mathcal{B}_d$:

$$\boldsymbol{R}_k = \boldsymbol{U}\,\boldsymbol{Q}_k\,\boldsymbol{U}^\dagger. \tag{3.51}$$

[8]Here $\hat{\mathbf{\Gamma}}_\Phi$ is the same operator with the subscript Ψ of equation (2.34) restricted to the mean-field part of the wave function, therefore the subscript has been modified to Φ.

Equation (3.50) is extremely advantageous, because the local value of the bilinear term $d_I^\dagger d_J$ is already available in the $\boldsymbol{W}$ matrix introduced in the previous section. As a consequence, the local value of the logarithmic derivative operator is expressed as a matrix product between $\boldsymbol{W}$ and $\boldsymbol{R}$, the latter being computed only once at the beginning of the simulation.

This concludes the discussion about the application of the VMC method to the two-band model. It should be noted that this method can be easily generalized to fit different physical circumstances, even with a slight modification of the parameters in the wave function (3.1): choosing opposite signs for t^1 and t^2 would reflect the physics of a two-band model with bands centered on different k-space points, while allowing complex couplings could be suitable for a description of systems with non-trivial topological structure. Moreover, the foregoing discussion can be straightforwardly generalized to multi-band models with more than two bands.

4

Metal-insulator transitions and superconductivity in the two-band model

The two-band Hubbard model was introduced in chapter 1. It was then shown in chapter 3 that the VMC method can be employed to compute its ground-state properties, which indeed is the purpose of this thesis. The present chapter describes the outcomes of the simulations. To illustrate the application of the VMC method, we start with a brief detour on the single-band Hubbard model. Subsequently, we present the results on the two-band models, starting from the phase diagrams at global half-filling with and without Hund's coupling, comparing them with known results. The superconducting properties of the two-band model are then presented.

All error bars have been omitted from the graphics: if not explicitly stated otherwise (that is, in the case of phase diagrams), this is due to the fact that error bars are comprised within the size of the marks.

4.1 Single-band Hubbard model

While the properties of multi-band models are not well-established, those of the single-band model have been investigated in depth, although some details still have to be clarified. Nevertheless, it is useful to make a survey of some results obtained with VMC, which will eventually be used as a reference for a consistent comparison to the results of the two-band model; indeed, certain findings on the two-band model can be interpreted in terms of those of the

single-band model.

For the one-band model, we use the non-magnetic wave function described in section 3.1 (adequately restricted to a one band), which we remind for clarity:

$$|\Psi_V\rangle = \mathcal{P}_J \, |\Phi_{MF}\rangle \, . \tag{4.1}$$

As expected from the theory of cuprate superconductors, the superconductive pairing has $d_{x^2-y^2}$ symmetry and is characterized only by its nearest-neighbor value as in equation (3.5); this parameter will be referred to as Δ as long as the focus is on the single-band model. Other than that, the BCS part of the trial wave function, which corresponds to $|\Phi_{MF}\rangle$, is characterized by the nearest-neighbor renormalized hopping parameter $\tilde{t}$, so that the free band dispersion has the form (3.4); the chemical potential enters the wave function only out of half-filling. Correlations are implemented by means of Jastrow factors.

All simulations of this section refer to a square supercell of $L = 242$ sites and periodic boundary conditions. Such supercell belongs to a class of 45°-tilted elementary cells (as the one represented in figure 4.1), which have the property that the projected Fermi sea is non-degenerate at half-filling [30], in other words eigenstates of the mean-field Hamiltonian form a closed shell, provided that the linear size of the cell n_x is odd.

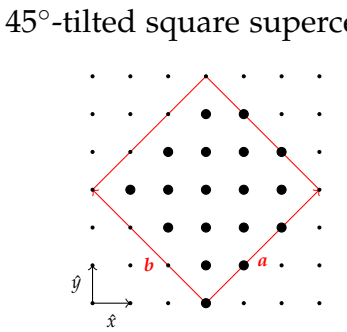


Figure 4.1: A 45°-tilted square supercell of $L = 18$ total cluster sites. $\hat{x}$ and $\hat{y}$ are the lattice basis vectors. For general rectangular cells, the defining vectors are given by the equations $\boldsymbol{a} = n_x\hat{x} + n_y\hat{y}$ and $\boldsymbol{b} = -n_y\hat{x} + n_x\hat{y}$, the coefficients n_x and n_y being integers; for square cells $n_x = n_y$, in this particular case $n_x = n_y = 3$. The total number of sites is $L = 2(n_x)^2$, that is 18.

4.1.1 The Mott Transition at half-filling

The first issue to be addressed is the phase diagram of the model, which is determined by marking the location of the Mott transition. As previously proved, this can be accomplished by computing the density-density correlation function $N(\boldsymbol{q})$ for different values of U/t and observing where the change from linear to quadratic behavior occurs. It is found that the

critical value is $(U/t)_c = 7.5 \pm 0.5$[1], as can be seen by figure 4.2. By increasing the ratio U/t, the system undergoes a continuous second-order phase transition. This continuous behavior can be seen from the change of the pairing of the wave function Δ as a function of U/t: this starts from a negligible value at small U/t, it increases rapidly when approaching the critical value $(U/t)_c$ and finally stabilizes at a plateau.

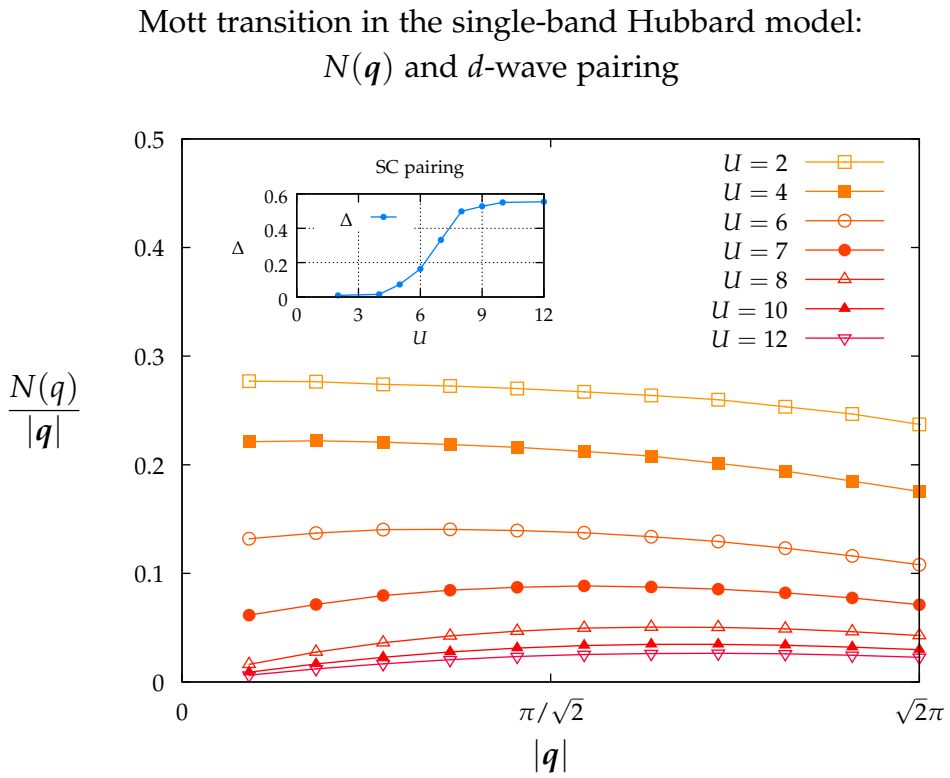


Figure 4.2: The main frame shows the correlation function $N(\boldsymbol{q})$ divided by $|\boldsymbol{q}|$ computed along the line $\Gamma - M$ in the Brillouin zone. The hopping matrix element t is set equal to 1 and thus the Coulomb interaction U defines the situation. Approaching from the metallic region at low U and increasing it, the value of $\frac{N(\boldsymbol{q})}{|\boldsymbol{q}|}$ at small $|\boldsymbol{q}|$ becomes smaller until $U = 7$; at $U = 8$ it becomes vanishing. This marks the onset of the Mott insulating phase at $U_c = 7.5$. The inset depicts the evolution of the superconductive d-wave pairing Δ over different values of U.

[1] Using a trial wave function without superconductive pairing gives a slightly larger critical value $(U/t)_c = 8.5 \pm 0.5$ [67].

4.1.2 Superconductivity

The presence of superconductive pairing in the wave function of the Hubbard model at half-filling is a hint of the presence of superconductivity. At half-filling pairing does not induce superconductivity, because of the presence of the long-ranged Jastrow factors, which suppresses charge fluctuations and thus the *preformed* pairs described by $|\Phi_{MF}\rangle$ cannot conduct; upon hole-doping, electrons are no more frozen in the Mott insulating state and superconductivity arises. To see that the trial wave function is indeed capable of describing this picture, the pairing correlation functions can be computed for a moderately large value of the Coulomb interaction, for instance $U/t = 12$.

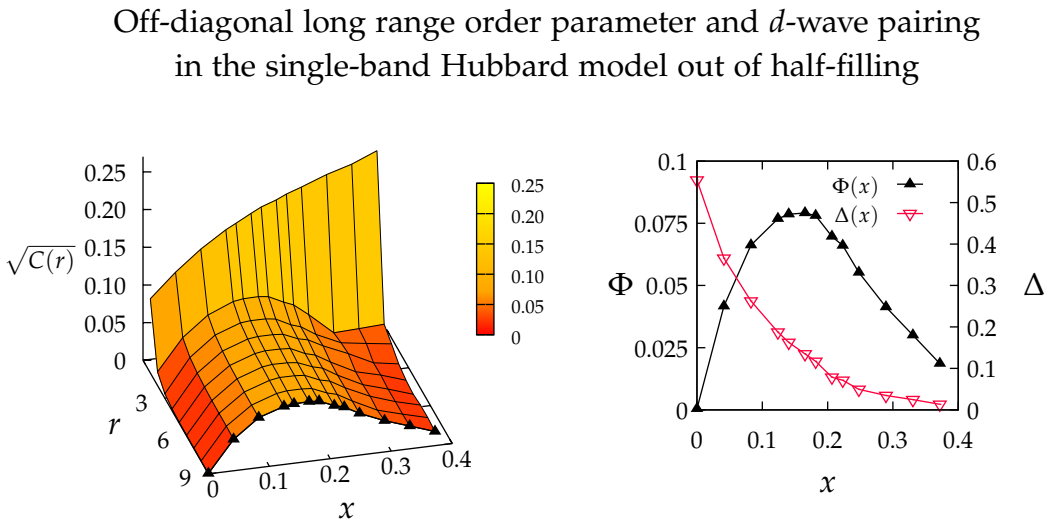


Figure 4.3: The square root of the superconducting pairing correlation (defined in equation (3.24)) as a function of the distance r and the hole doping concentration x, for a fixed value of the Coulomb interaction $U/t = 12$. For the largest value of r allowed by the adopted supercell, it approximates the ODLRO parameter $\Phi(x)$, which is represented as a function of the doping by the bold black line plus triangles; the same function is shown in the right panel for clarity (left scale), together with the d-wave pairing parameter $\Delta(x)$, red and empty triangles (right scale).

Figure 4.3 shows that the data reproduce the physical picture predicted by the RVB theory: at half-filling the system is a fully-developed Mott insulator. Adding holes to the system lowers the magnitude of the d-wave pairing, but the off-diagonal long-range order parameter (ODLRO) Φ (defined in equation (3.25)) becomes non-zero, reaches its maximum around an optimal doping $\bar{x} \simeq 0.15$ and drops off if x is increased further beyond that value. This behavior recalls the superconducting dome[2] displayed in the typical phase diagram of figure

[2]The superconducting dome in this framework does not refer to the behavior of the critical temperature $T_c(x)$

1.2. These outcomes are in agreement with what is already known from other calculations, for example see reference [32].

4.2 Phase Diagram of the two-band model at half-filling

Turn now to the main problem, the two-band degenerate Hubbard model on a square lattice, whose Hamiltonian is written in equation (1.12). Let the parameters satisfy $U' = U - 2J$ and $J' = J$ and consider at first the model at half-filling. Under these conditions, the question that has to be answered is to what extent the VMC method can reproduce the phase diagram of the model. This question stands as a consistency test before trying to understand the properties out of half-filling, which are less well known. The goal is to qualitatively reproduce the phase diagrams obtained with other numerical methods (see figure 1.8). At the same time, it is crucial to understand the features of the trial wave functions that represent these phase diagrams and to interpret their physical meaning. For instance, in the presence of Hund's coupling the wave function may effectively describe strongly-correlated electrons in ruthenium oxides.

All computations on the phase diagram of the two-band model are based on a square supercell of $L = 98$ sites, which belongs to the same class of cells employed in the previous section (see figure 4.1).

4.2.1 Orbital Selective Mott insulator

Let the Hund's coupling be zero. In this case, the trial wave function is characterized by nearest-neighbor intra-band d-wave pairings Δ_{cc} and Δ_{ff} (the inter-band pairing can be neglected) and by intra-band and inter-band Jastrow factors. The nearest-neighbor hopping parameter has been chosen to be equal to the hopping matrix element of the Hamiltonian $\tilde{t}^{\alpha} = t^{\alpha}$. As already done for the single-band model, the conducting properties of the wave function can be deduced by computing the density-density correlation functions $N^{\alpha}(\boldsymbol{q})$ of both bands. The analysis of the phase diagram reduces to computing $N^{\alpha}(\boldsymbol{q})$ on a thick enough grid of values of R and, at fixed R, for different values of U/t_1.

The outcomes are shown in figure 4.4: the separation lines are determined by fixing R and observing the behavior of the best (lowest in energy) trial state while varying U/t_1, which in practice is done by looking at how $N^{\alpha}(\boldsymbol{q})/|\boldsymbol{q}|$ changes (see figure 4.5). For a small enough value of R such as 0.3, the evolution as a function of U is continuous; explicitly, the behavior of $N^{\alpha}(\boldsymbol{q})/|\boldsymbol{q}|$ evolves continuously from a fully metallic one ($U/t_1 = 4$, where it approaches a finite value at $|\boldsymbol{q}| = 0$ for both $\alpha = c, f$), to an orbital selective one ($U/t_1 = 5$ and $U/t_1 = 7$, where it approaches a finite value at $|\boldsymbol{q}| = 0$ for $\alpha = c$ and zero for $\alpha = f$) to an insulating one ($U/t_1 = 8$, where it vanishes for both bands). This behavior depicts what happens for $R <$

as a function of the hole doping, but of the ODLRO parameter $\Phi(x)$, which is the physical quantity available for a variational method at $T = 0$. The two behaviors are to some extent related, as it can be thought that the tougher the zero-temperature ODLRO parameter is, the higher the critical temperature required to break the order will be.

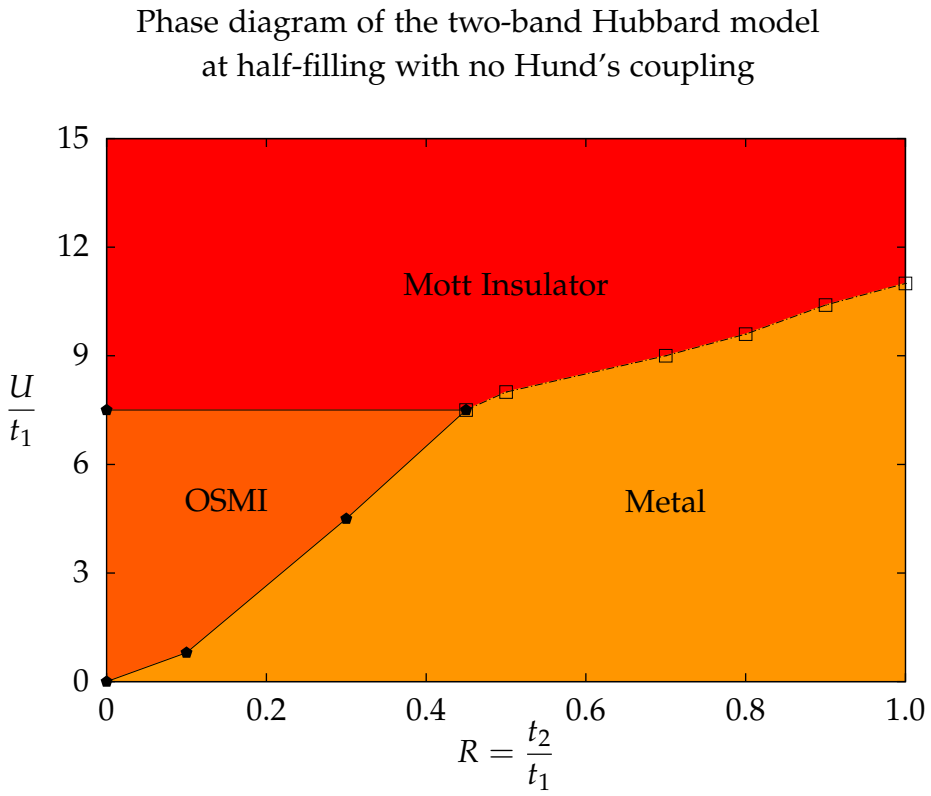


Figure 4.4: Phase diagram of the two-band Hubbard model at half-filling with no Hund's coupling ($J = 0$). Solid lines plus pentagons indicate a second-order phase transition, whereas dotted-dashed lines plus empty squares a first-order one. Second order phase transition marks have an uncertainty of about 0.5, while first order transition marks of about 0.1.

0.45, where it is found that by increasing U/t_1 at fixed R, the system undergoes two continuous second-order phase transitions: each band undergoes a Mott transition at its own critical value of U/t_α, which for the f-band is significantly modified from its value for independent bands, but the continuous character of each transition resembles that of the single-band model. The OSMP is reduced with respect to the case of $U' = 0$ (compare figure 1.7) and is present up to the bandwidth ratio threshold of $R = 0.45$ and for intermediate couplings up to $U/t_1 = 7.5$. For $R > 0.45$ there is no OSMP between the metallic and insulating phases and a different transition shows up: both bands go from a metallic to an insulating behavior at the same value of U/t_1 in a discontinuous manner, a feature typical of a first-order phase transition. This can be seen from the behaviors of $N^\alpha(\boldsymbol{q})/|\boldsymbol{q}|$ at a bandwidth ratio $R = 0.7$: the lines of $N^\alpha(\boldsymbol{q})/|\boldsymbol{q}|$ are well-separated into two behaviors: for $U/t_1 = 8$ and $U/t_1 = 9$ both bands are conducting, while at $U/t_1 = 10$ and $U/t_1 = 11$ they are both insulating and there is no intermediate behavior. At $R = 1.0$ the physical picture closely resembles the case with $R = 0.7$,

with the only difference that the correlation functions of the two bands are identical, a trivial consequence of the fact that $t_1 = t_2$. In this case, it should be noted that the large peak at the M point attained by $N^{\alpha}(\boldsymbol{q})/|\boldsymbol{q}|$ in the metallic phase denotes a tendency toward developing an intra-band charge-density-wave state. The first-order behavior reflects itself also in the fact that we find it possible to stabilize the insulating state in the region where the metallic one is favored and viceversa.

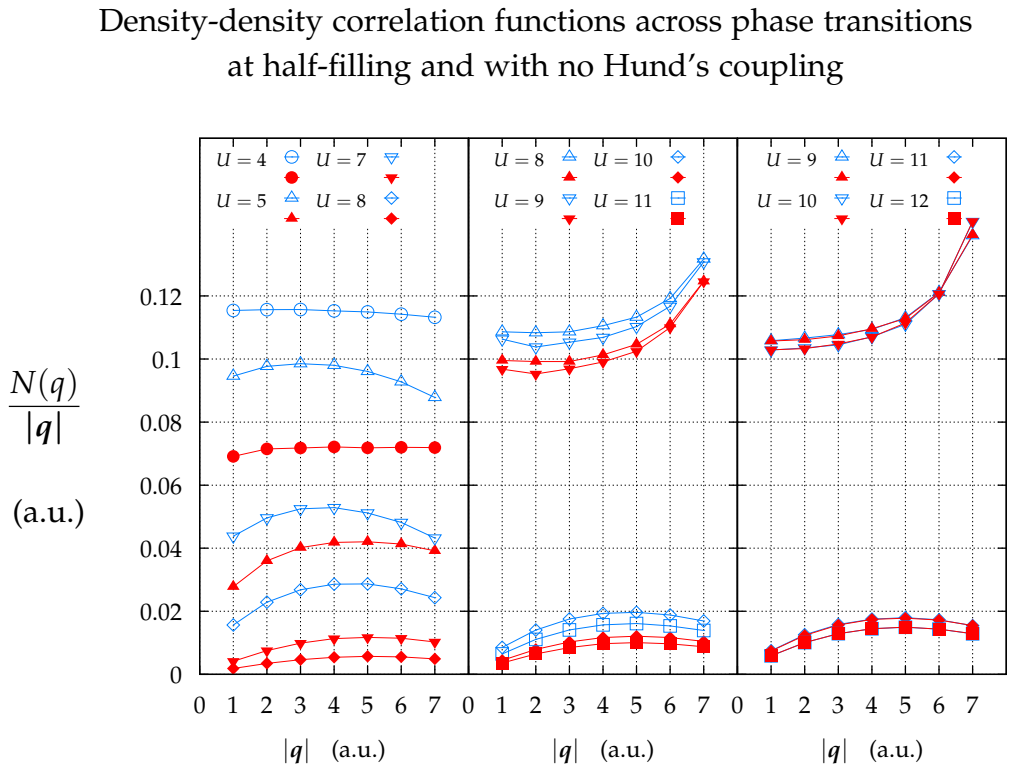


Figure 4.5: The density-density correlation functions divided by the momentum $N^{\alpha}(\boldsymbol{q})/|\boldsymbol{q}|$ computed for different values of the bandwidths ratio $R = 0.3$ (left panel), $R = 0.7$ (middle panel) and $R = 1.0$ (right panel) along the $\Gamma - M$ line in the first Brillouin zone in the case with no Hund's coupling. The blue empty figures refer to the c-band; the red filled ones refer to the f-band. In the figure t_1 is supposed to be 1, so that U is the only parameter to be given.

While for $R < 0.45$ the OSMP-insulator transition line is essentially flat (independent on R), in the right half of the phase diagram the metal-insulator transition line increases from about $U/t_1 \simeq 7.5$ up to $U/t_1 \simeq 11$. This enhanced value is not unexpected and was qualitatively and quantitatively determined by Gunnarsson *et al.* [70] for a general Hubbard model with N-fold orbital degeneracy: the qualitative argument behind the enhancement is that for large

U the hopping of one electron allows an energy gain of $t\sqrt{N}$ rather than only t. One can then write the energy gap as $\mathcal{E}_g = U - \sqrt{N}W$, W being the bandwidth. Extrapolating this relation to smaller values of U, predicts the location of the Mott transition to be at a value which scales with $\sqrt{N}$. By considering the fact that the single-band Mott transition is placed at $U/t = 7.5$, the argument would predict the transition in the two-band model at $R = 1$ to occur at $U/t_1 \simeq \sqrt{2} \cdot 7.5 \simeq 10.6$, which is convincingly close to 11.

The broad metallic region is described by a wave function with intra-band Jastrow factors displaying a $|q|^{-1}$ divergence. On the other hand, the OSMP is described by a $|q|^{-2}$-divergent intra-band Jastrow potential for the f-band only, while the insulating phase by $|q|^{-2}$-divergent intra-band Jastrow potential for both bands. Inter-band Jastrow factors are present, but their behavior is less singular compared to the intra-band ones. d-wave superconductive pairings have a continuous increase by moving from the metal into the OSMP and then into the insulating phase, while abruptly become non-zero if transitioning directly from the metal into the insulator, which is just another evidence for the existence of different kinds of transition: it can be seen from figure 4.6 that the growth of the pairing parameters is faster compared to the single-band for both $R = 0.5$ and $R = 1.0$, being extremely sharp in the latter case. Considering the limiting values of the d-wave pairings, it can be seen that for large U/t_1 their values approach those of the corresponding single-band model.

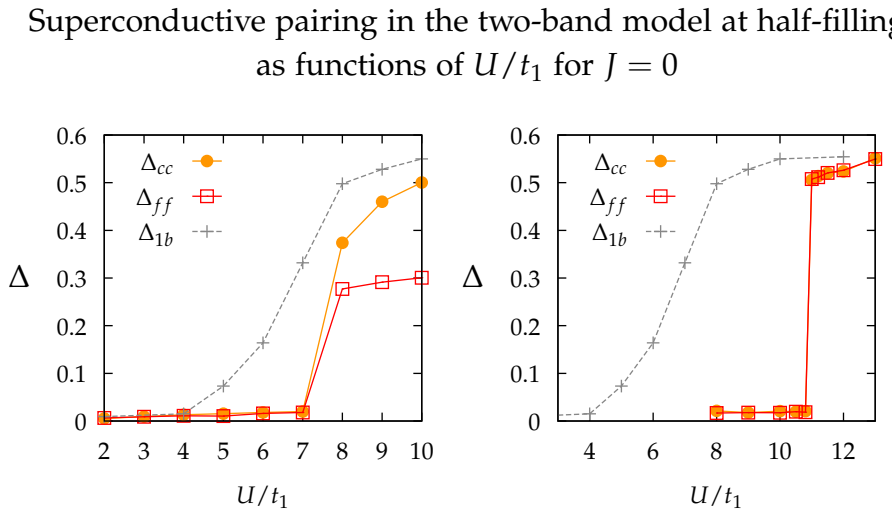


Figure 4.6: Intra-band superconductive pairings Δ_{cc} (amber filled circles) and Δ_{ff} (red empty squares) as functions of the ratio U/t_1 for $R = 0.5$ (left panel) and $R = 1$ (right panel), compared to the superconductive pairing of the one-band Δ_{1b} (grey crosses).

4.2.2 Effects of Hund's J

Turning on the Hund's coupling J, which favors the on-site ferromagnetic alignment of electrons, calls for the trial wave function to be refined beyond the simple projected d-wave paired BCS state for each band. Indeed, the physics of spin-triplet superconductors, which arises for instance in Sr_2RuO_4 [71] (although this compound would actually be described by a three-band model), suggests that the introduction of an on-site inter-band s-wave spin-triplet coupling might be consistent. It is found that this is the case: introducing such a spin-triplet Δ_{cf} significantly lowers the variational energy of the wave function, while the intra-band pairings Δ_{cc} and Δ_{ff} are suppressed to a lower (but still significantly non-zero, of the same order of magnitude) value than for the $J = 0$ case.

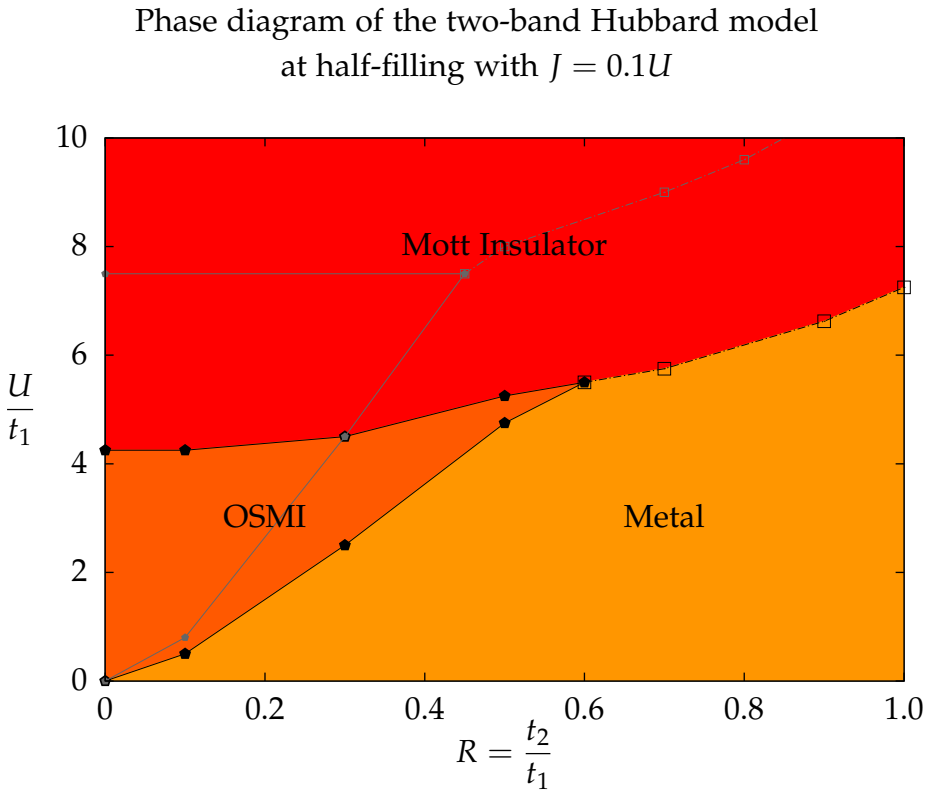


Figure 4.7: Phase diagram of the two-band Hubbard model at half-filling with a finite Hund's coupling $J = 0.1U$. For comparison we overlay the $J = 0$ phase diagram of figure 4.4 as grey lines; solid lines plus pentagons indicate a second-order phase transition, dotted-dashed lines plus empty squares a first-order one. All marks have an uncertainty of about 0.2, except for the metal-OSMI transition point at $R = 0.1$, which is known with an uncertainty of about 0.5.

Figure 4.7 shows the phase diagram for a physically reasonable value of $J = 0.1U$ (this

quantifies the supposed order of magnitude of J); phase transition lines are determined once again from the behavior of $N^{\alpha}(\boldsymbol{q})$, which is shown in figure 4.8. As in the case with no Hund's coupling (compare figure 4.5), a small enough value of R gives a continuous behavior of $N^{\alpha}(\boldsymbol{q})/|\boldsymbol{q}|$ while moving across the metallic, orbital selective and insulating phases. Like in the absence of Hund's coupling, the OSMP exists up to a threshold value of R, which is increased compared to the value $R \simeq 0.45$ obtained for $J = 0$ to $R \simeq 0.6$; this enhancement reflects an overall expansion of the OSMP in the direction of larger bandwidth ratios. Again, at the larger bandwidth ratio $R = 0.7$, the transition becomes first-order, although with a less sharp character with respect to the $J = 0$ case, revealing that the OSMP extends closer to $R = 0.7$. At $R = 1.0$ the two bands display the same behavior, as expected. The peaks at the M point observed in the insulating phase with no Hund's coupling have disappeared, signaling that a non-zero J suppresses the intra-band charge-density-wave state. While the OSMP is favored with respect to the metallic phase, we find a significant lowering of the critical value of the Coulomb interaction that marks the onset of the insulating phase (from $U/t_1 = 7.5$ to 4.5 at $R = 0$ and from $U/t_1 = 11$ to 7.25 at $R = 1$). This calculation provides a clear picture of the effect of the Hund's coupling in a degenerate Hubbard model with orbital degeneracy at global half-filling: our results are in qualitative agreement with previous studies [46].

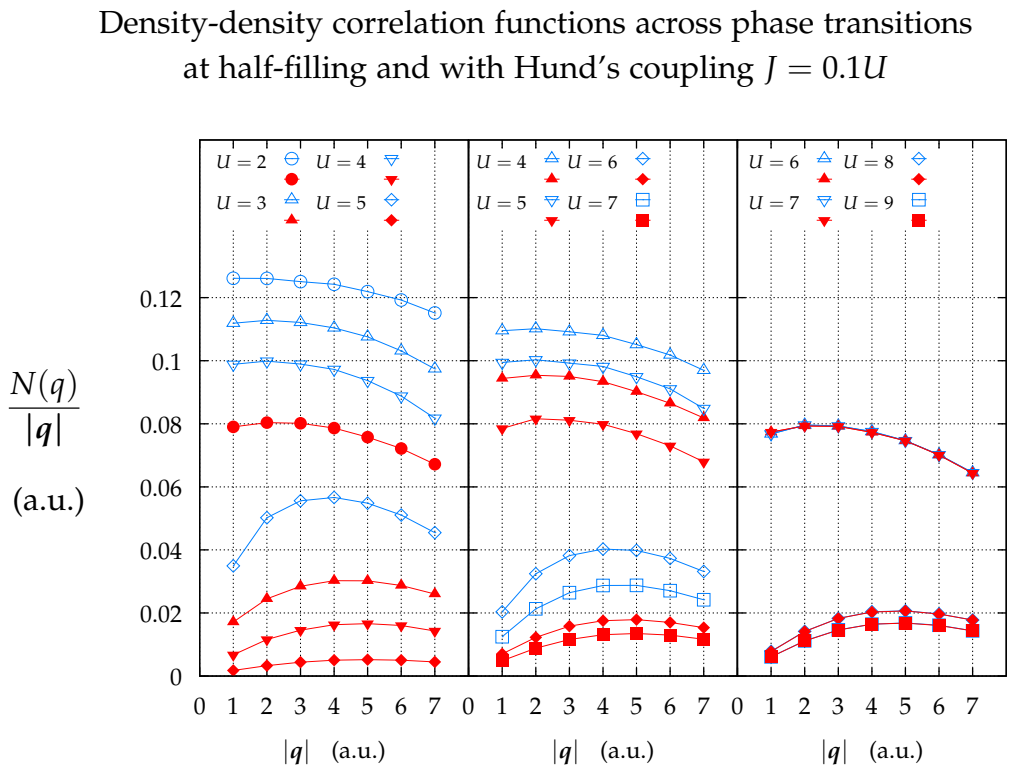


Figure 4.8: Same as figure 4.5, but with a finite Hund's coupling $J = 0.1U$.

4.3 Superconductivity

So far, only the half-filling case has been discussed. In the single-band model, in the presence of a finite hole-doping concentration, a non-zero ODLRO parameter appears, signaling the presence of superconductivity. This section explores superconductivity in the two-band model with no Hund's coupling: this particular case has no direct physical application, but rather answers the general question of whether superconductivity is enhanced or not by orbital degeneracy. In order to do this, starting from the parameters of the best trial state at half-filling, the wave function is optimized for different values of the global hole-doping x (a number of holes is introduced in the model and are left free to rearrange between the bands).

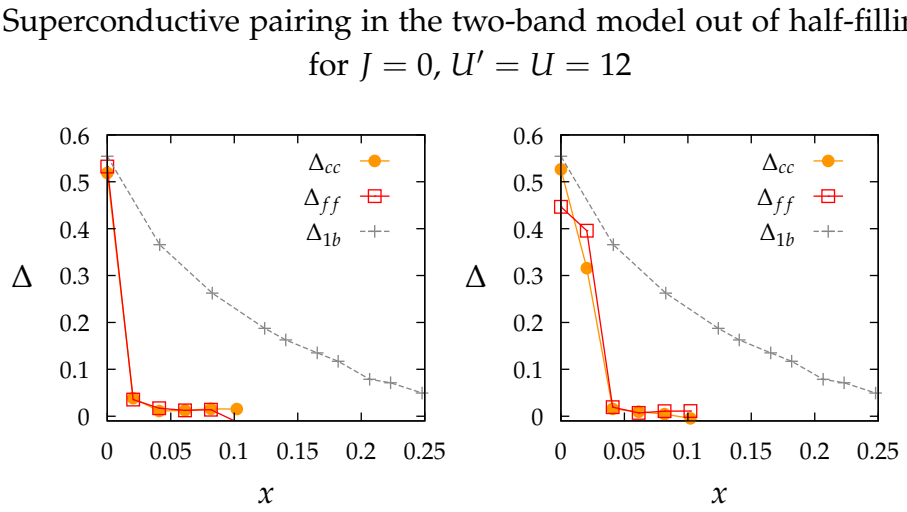


Figure 4.9: Intra-band superconductive pairings Δ_{cc} (amber filled circles) and Δ_{ff} (red empty squares) as functions of the hole doping concentration x for $R = 1$ (left panel) and for $R = 0.8$ (right panel) in the insulating phase at $U' = U = 12$, compared with the corresponding value of the single-band model (grey crosses). The state with lowest energy loses pairing as soon as x is non-zero, while the reduction of the pairing magnitude in the one-band model occurs over a significantly wider doping scale.

By doping an insulating phase at $R = 1.0$, it is seen that superconductive pairing is disrupted as soon as holes are introduced in the system. The variational parameters Δ_{cc} and Δ_{ff}, which retain a finite value at half-filling, vanish as soon as $x > 0$. In this particular case ($R = 1$ and $J = 0$) the model has the enhanced $SU(4)$ symmetry, but the the best trial state at half-filling has finite Δ_{cc} and Δ_{ff}, which explicitly break such symmtery[3]; moreover, its inter-band Jastrow factors differ from the intra-band ones, another condition which breaks the

[3]Unlike in the single-band model, where pairing in the wave function preserves the $SU(2)$-spin symmetry.

symmetry. At half-filling, this state is very close in energy to a $SU(4)$-symmetric state, which has equal intra-band and inter-band Jastrow factors and lacks any pairing parameter. Out of half-filling, this $SU(4)$-symmetric state immediately becomes lowest in energy and thus Δ_{cc} and Δ_{ff} vanish, indicating that no superconductivity occurs. Breaking the $SU(4)$ symmetry by decreasing R does not modify this behavior.

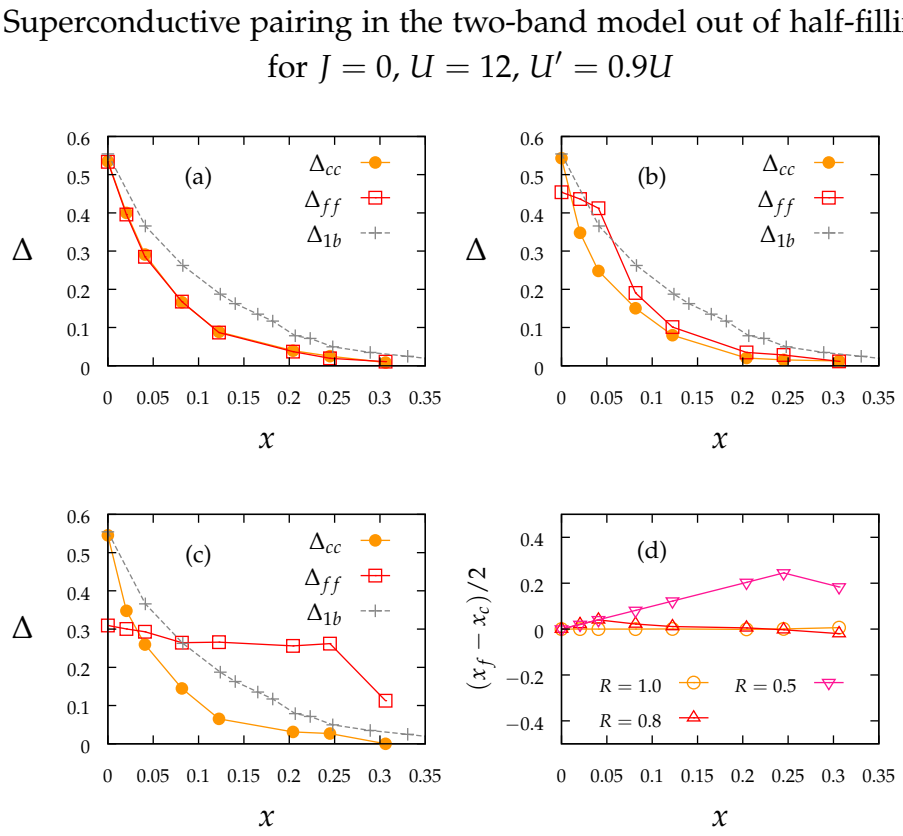


Figure 4.10: Intra-band superconductive pairings Δ_{cc} (amber filled circles) and Δ_{ff} (red empty squares) as functions of the hole-doping x for $R = 1$ (upper-left panel), $R = 0.8$ (upper-right panel) and $R = 0.5$ (lower-left panel) in the insulating phase at $U = 12$ and $U' = 0.9U$, compared with the corresponding value of the single-band model (grey crosses). Contrarily to the $U' = U$ case shown in figure 4.9, the pairing parameters at $R = 1.0$ behave similarly to the one of the single-band model. Panel (d) shows the discrepancy between the hole density in the c-band $x_c \equiv 1 - n_c = (L - N_c)/L$ and in the f-band $x_f \equiv 1 - n_f = (L - N_f)/L$ divided by 2: if this quantity equals the global hole concentration x, there are no holes in the f-band, while if it vanishes, bands are evenly doped.

Indeed, superconductivity exists in the limit $U' = 0$, where the system coincides with two

single-band models, thereby for some value of $U' < U$ superconductivity should be restored. Accordingly, lowering U' should penalize the symmetric state. For $U' = 0.9U$, the hole-doping behavior of the pairing parameters looks much alike that of the single-band models in terms of magnitude (this is seen in figure 4.10). However, their behavior becomes asymmetric if R is moved away from the value 1: for $t_2 < t_1$ the bands are unevenly populated with holes and up to a certain value of the doping concentration $\tilde{x}$ all holes go into the less correlated c-band. If $x < \tilde{x}$ the pairing of the c-band Δ_{cc} decreases as expected, while the one of the f-band Δ_{ff} remains roughly constant. For $x > \tilde{x}$, holes begin to populate the f-band as well and the parameter Δ_{ff} starts decaying. $\tilde{x}$ amounts to about 0.05 for $R = 0.8$ and to about 0.25 to $R = 0.5$.

Off-diagonal long-range order parameters of the two-band model for $J = 0$, $U = 12$, $U' = 0.9U$ and different values of R

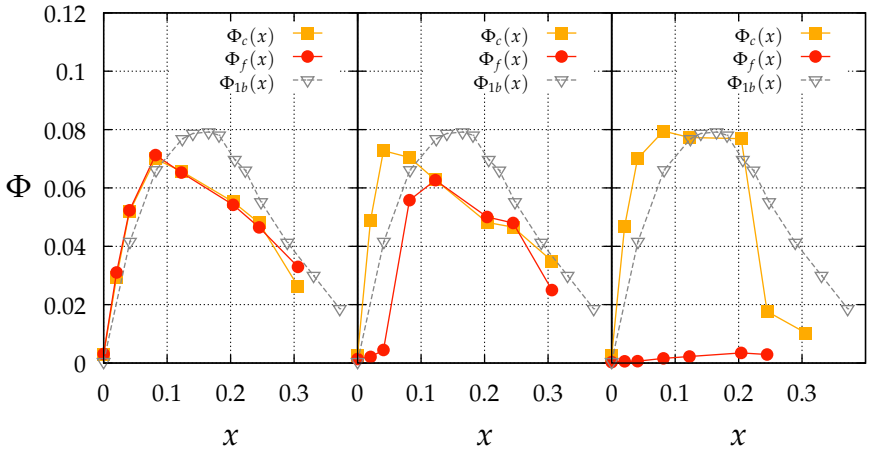


Figure 4.11: The intra-band off-diagonal long-range order parameters Φ_c (amber) and Φ_f (red) for $J = 0$, $U = 12$, $U' = 0.9U$ as functions of the hole-doping x for $R = 1$ (left), $R = 0.8$ (center) and $R = 0.5$ (right) starting from the half-filling insulating phase at $U = 12$ and $U' = 0.9U$, compared with the corresponding value of the single-band model (grey).

At the same time, the superconducting ODLRO parameters of the two bands at $U' = 0.9U$ exhibit a behavior which reminds of the one-band model, starting from zero at half-filling, attaining a maximum value and then decreasing, as shown by figure 4.11. In the case of different bandwidths, the *domes* arise in different ranges of doping (most strikingly, for $R = 0.5$ there is no ODLRO parameter of the f-band up to the value where that of the c-band has become negligible). In any case, however, the maximum magnitude of Φ_α is at most comparable with that of the one-band model. Although we expect that at $R = 0.5$ Φ_f would become large for a global doping percentage of 0.3, this would be an artifact as the constant

value of Δ_{ff} as a function of x is an artifact as well: all holes populate the c-band and the f-band remains at half-filling and thus insulating. In conclusion, for $U' = 0.9U$ the two-band model does not display an enhancement of superconductivity compared to what observed in the single-band model: this statement can be reasonably extrapolated for any value of $U' < U$, since the limiting case $U' = 0$ reproduces the single-band model itself.

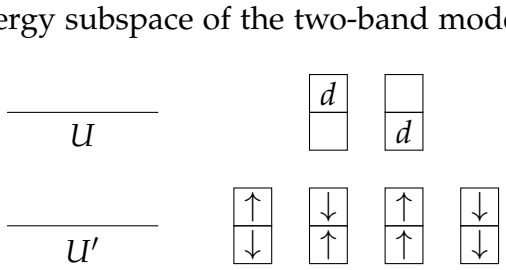


Figure 4.12: The on-site low-energy subspace of the two-band model at global half-filling. In this case $U' < U$ and the two doubly occupied states lay higher in energy with respect to the four singly occupied ones. If $U = U'$ all states are equivalent in energy and the degeneracy is six.

In the light of this, the suppression of superconductivity in the two-band model is probably due to U' hindering d-wave intra-band superconductive pairing in the wave function rather than to the presence of the enhanced symmetry, since lifting such symmetry by choosing $R < 1$ does not favor superconductivity. In terms of on-site low-energy states at global half-filling, the two-band model with $U' < U$ resembles the case of two single-band models (see figure 4.12): if $U' < U$ the low-energy subspace contains the 4 states with singly occupied orbitals. At $U' = U$ the degeneracy is increased to 6 and the states with doubly occupied orbitals may oppose d-wave intra-band superconductive pairing: this probably shows up out of half-filling because holes generate empty orbitals. Indeed, choosing the opposite situation $U' > U$ stabilizes an intra-band s-wave on-site pairing, which is supported by configurations with doubly occupied orbitals.

Conclusions

In the present work, we investigate the two-band degenerate Hubbard model by using the Variational Monte Carlo approach and employing optimized Jastrow-enhanced BCS wave function as trial wave functions, in the spirit of the Resonating Valence Bond theory. In the absence of Hund's coupling, we predict the existence of an Orbital Selective Mott Phase. We reproduce the whole half-filling phase diagram, which is found to agree with previous calculations done with different methods, especially DMFT. Contrarily to DMFT results which are correct in infinite dimensions, the present work fully deals with the two-dimensional system. A finite Hund's coupling modifies the phase diagram by an overall lowering of the critical value of the insulating phase and a broadening of the OSMP to the detriment of the metallic phase: this feature confirms previous studies too. Furthermore, the Hund's coupling stabilizes a best trial wave function with an on-site triplet pairing. We then conclude that the half-filling phase diagram of the two-band model can be effectively described by means of the Variational Monte Carlo method. We also consider the hole-doped region of the model, in the case with no Hund's coupling. It turns out that no enhancement of the superconductivity was possible compared to the single-band model. On the contrary, we find that superconductivity is suppressed for $U' = U$, while it can be restored by enforcing $U' < U$: in the latter case, the system behaves similarly to two separate single-band models, even for fairly small deviations of U' from U.

The outcomes of this work discusse the issue of superconductivity in the two-band model in a very preliminary way: the physically most interesting case of spin-triplet superconductivity in the presence of Hund's coupling has not been addressed. This work also gives pre-

liminary answers on whether the Variational Monte Carlo method can be applied to Hubbard models with orbital degeneracy. In this sense, it can be regarded as a benchmark of the accuracy of the method on a system whose properties are known, before tackling less known cases: in particular, three bands can give a more realistic picture of some materials, the iridium oxides for instance, but also the ruthenates or the cuprates themselves. Certain quantum spin liquid systems involve orbital degeneracy as well. These are only a few of the possibile developments that can derive from the present work.

Appendices

A

Derivation of the Hamiltonian

In this appendix the two-band model H_{2B} of equation 1.12 is derived. The starting point is the many-body Hamiltonian for interacting electrons in a periodic potential, equation 1.10. The first step is to introduce a basis set of localized wave functions, such as Wannier states, which are defined by:

$$\chi_{i,\alpha,\sigma}(\boldsymbol{x}) = \chi_{\alpha,\sigma}(\boldsymbol{x} - \boldsymbol{R}_i) \equiv \frac{1}{\sqrt{L}} \sum_{k \in BZ} e^{-i\boldsymbol{k}\cdot\boldsymbol{R}_i}\, \psi_{\boldsymbol{k},\alpha,\sigma}(\boldsymbol{x}) \tag{A.1}$$

in terms of the Bloch wave functions $\psi_{\boldsymbol{k},\alpha,\sigma}(\boldsymbol{r})$. As a basis, the Wannier wave functions satisfy orthogonality and completeness relations:

$$\int d^3\boldsymbol{x}\, \chi_{i_1,\alpha_1,\sigma_1}(\boldsymbol{x})^*\, \chi_{i_2,\alpha_2,\sigma_2}(\boldsymbol{x}) = \delta_{i_1,i_2}\, \delta_{\alpha_1,\alpha_2} \delta_{\sigma_1,\sigma_2}\,; \tag{A.2a}$$

$$\sum_{i,\alpha,\sigma} \chi_{i,\alpha,\sigma}(\boldsymbol{x})^*\, \chi_{i,\alpha,\sigma}(\boldsymbol{y}) = \delta(\boldsymbol{x} - \boldsymbol{y})\,. \tag{A.2b}$$

Changing of the basis of the Hilbert space corresponds to a canonical transformation of the annihilation and creation operators, the wave functions being the matrix elements of the transformation as in equation:

$$\hat{\psi}_\sigma(\boldsymbol{x}) = \sum_i \sum_\alpha \chi_{i,\alpha,\sigma}(\boldsymbol{x})\, b_{i,\alpha,\sigma}\,. \tag{A.3}$$

The canonical anti-commutation relations of the electron field operators map into those of the new basis $\{b_{i,\alpha,\sigma}\}_{i,\alpha,\sigma}$:

$$\left\{b_{i,\alpha,\sigma}, b^{\dagger}_{i,\alpha,\sigma}\right\}_{i,\alpha,\sigma} = \delta_{i,j}\,\delta_{\alpha_1,\alpha_2}\,\delta_{\sigma_1,\sigma_2}\,; \tag{A.4a}$$

$$\left\{b_{i,\alpha,\sigma}, b_{i,\alpha,\sigma}\right\}_{i,\alpha,\sigma} = 0\,. \tag{A.4b}$$

Mapping the Hamiltonian consists in applying the transformation A.3 to the field operators in the *ab initio* expression. The single-particle term (kinetic energy plus the periodic potential generated by the ions) becomes:

$$\begin{aligned} H_0 &= \sum_{\sigma}\int d^3\boldsymbol{x}\,\hat{\psi}^{\dagger}_{\sigma}(\boldsymbol{x})\left(-\frac{\hbar^2}{2m_e}\,\boldsymbol{\nabla}^2+\sum_l V(\boldsymbol{x}-\boldsymbol{R}_l)\right)\hat{\psi}_{\sigma}(\boldsymbol{x}) = \\ &= \sum_{\sigma}\sum_{\alpha,\alpha'}\sum_{i,j}\int d^3\boldsymbol{x}\,\chi_{i,\alpha,\sigma}(\boldsymbol{x})^*\,b^{\dagger}_{i,\alpha,\sigma}\left(-\frac{\hbar^2}{2m_e}\,\boldsymbol{\nabla}^2+\sum_l V(\boldsymbol{x}-\boldsymbol{R}_l)\right)\chi_{j,\alpha',\sigma}(\boldsymbol{x})\,b_{j,\alpha',\sigma} = \\ &= \sum_{\sigma}\sum_{\alpha,\alpha'}\sum_{i,j} t^{\alpha,\alpha'}_{i,j}\;b^{\dagger}_{i,\alpha,\sigma}b_{j,\alpha',\sigma}\,, \end{aligned} \tag{A.5}$$

where the hopping matrix elements $t^{\alpha,\alpha'}_{i,j}$ correspond to:

$$\begin{aligned} t^{\alpha,\alpha'}_{i,j} &= \int d^3\boldsymbol{x}\,\chi_{i,\alpha,\sigma}(\boldsymbol{x})^*\left(-\frac{\hbar^2}{2m_e}\,\boldsymbol{\nabla}^2+\sum_l V(\boldsymbol{x}-\boldsymbol{R}_l)\right)\chi_{j,\alpha',\sigma}(\boldsymbol{x}) \\ &= \int d^3\boldsymbol{x}\left(\frac{\hbar^2}{2m_e}\left(\boldsymbol{\nabla}\chi_{i,\alpha,\sigma}(\boldsymbol{x})^*\right)\cdot\boldsymbol{\nabla}\chi_{j,\alpha',\sigma}(\boldsymbol{x})+\sum_l V(\boldsymbol{x}-\boldsymbol{R}_l)\,\chi_{i,\alpha,\sigma}(\boldsymbol{x})^*\,\chi_{j,\alpha',\sigma}(\boldsymbol{x})\right). \end{aligned} \tag{A.6}$$

Usually, due to the fact that Wannier orbitals are localized inside the unit cell, the hopping amplitude is not long-ranged and considering only the nearest-neighbor terms is a valid approximation; indeed, the hopping matrix can be shown to be translationally invariant. It can be seen that the $t^{\alpha,\alpha'}_{i,j}$ form a hermitian matrix in terms of the orbital indices. This allows to rewrite them in a diagonal form by a unitary transformation (in orbital space) of the Wannier functions (and for the operators as well), namely:

$$\phi_{i,\alpha,\sigma}(\boldsymbol{x}) = \phi_{\alpha,\sigma}(\boldsymbol{x}-\boldsymbol{R}_i) \equiv \frac{1}{L}\sum_{k\in BZ}\sum_j\sum_{\beta} e^{-i\boldsymbol{k}\cdot(\boldsymbol{R}_i-\boldsymbol{R}_j)}\left(U(\boldsymbol{k})\right)_{\alpha,\beta}\chi_{j,\beta,\sigma}(\boldsymbol{x})\,. \tag{A.7}$$

In terms of the new operators, the kinetic energy becomes:

$$H_0 = \sum_{i,j}\sum_{\alpha}\sum_{\sigma} t^{\alpha}_{i,j}\,c^{\dagger}_{i,\alpha,\sigma}\,c_{j,\alpha,\sigma}\,. \tag{A.8}$$

The interacting Hamiltonian is treated similarly: substituting the field operators with the new Wannier functions gives:

$$H_{int} = \frac{1}{2}\,U^{\alpha_1,\alpha_2,\alpha_3,\alpha_4}_{i_1,i_2,i_3,i_4}\,c^{\dagger}_{i_1,\alpha_1,\sigma}\,c^{\dagger}_{i_2,\alpha_2,\sigma'}\,c_{i_3,\alpha_3,\sigma'}\,c_{i_4,\alpha_4,\sigma}\,, \tag{A.9}$$

with the interaction terms given by:

$$U^{\alpha_1,\alpha_2,\alpha_3,\alpha_4}_{i_1,i_2,i_3,i_4} = \int d^3\boldsymbol{x}\, d^3\boldsymbol{y}\, \phi_{i_1,\alpha_1,\sigma}(\boldsymbol{x})^*\, \phi_{i_2,\alpha_2,\sigma}(\boldsymbol{y})^*\, U(\boldsymbol{x}-\boldsymbol{y})\, \phi_{i_3,\alpha_3,\sigma'}(\boldsymbol{y})\, \phi_{i_4,\alpha_4,\sigma}(\boldsymbol{x}). \tag{A.10}$$

Up to this point the discussion is general, the number of bands was not specified and also the orbital wave functions depend on the spin index (this would be required in presence of a spin-orbit interaction). By considering a case with two bands, the kinetic term becomes the one in 1.12. As for the interacting part, it is necessary to split it into contributions due to single channels. Consider the case where all the site indices correspond to the same site $i = i_1 = i_2 = i_3 = i_4$.

- The easiest channel possible is the on-site interaction between electrons in the same orbital, which is nothing but the contribution to the interaction in the Hubbard model. In this case $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha$. The interaction matrix element is:

$$U^{\alpha}_i = \frac{1}{2}\int d^3\boldsymbol{x}\, d^3\boldsymbol{y}\, |\phi_{i,\alpha}(\boldsymbol{x})|^2\; U(\boldsymbol{x}-\boldsymbol{y})\; |\phi_{i,\alpha}(\boldsymbol{y})|^2\,; \tag{A.11}$$

  because of the properties of Wannier wave functions, U^{α}_i is actually independent on the site index i. The corresponding contribution to the Hamiltonian is:

$$\begin{aligned} H^U_{int} &= \frac{1}{2}\sum_i \sum_\alpha \sum_{\sigma,\sigma'} U^\alpha c^\dagger_{i,\alpha,\sigma}\, c^\dagger_{i,\alpha,\sigma'}\, c_{i,\alpha,\sigma'}\, c_{i,\alpha,\sigma} = \\ &= \frac{1}{2}\sum_i \sum_\alpha \sum_{\sigma,\sigma'} U^\alpha \left(c^\dagger_{i,\alpha,\sigma}\, c_{i,\alpha,\sigma}\, c^\dagger_{i,\alpha,\sigma'}\, c_{i,\alpha,\sigma'} - \delta_{\sigma,\sigma'} c^\dagger_{i,\alpha,\sigma}\, c_{i,\alpha,\sigma'} \right) = \\ &= \sum_\alpha U^\alpha \sum_i n_{i,\alpha,\uparrow}\, n_{i,\alpha,\downarrow} = \\ &= U \sum_i \sum_\alpha n_{i,\alpha,\uparrow} n_{i,\alpha,\downarrow}. \end{aligned} \tag{A.12}$$

  It is the usual intra-band Hubbard interaction for each of the two bands; if the Coulomb integrals of equation A.11 are the same for the two bands, this term reduces to the intra-band contribution of 1.12.

- The second contribution comes from the on-site interaction between electrons in different bands $\alpha_1 = \alpha_4 = \alpha \neq \beta = \alpha_2 = \alpha_3$. The interaction integral is (again, it does not depend on the site index i):

$$U'^{\alpha,\beta} = \frac{1}{2}\int d^3\boldsymbol{x}\, d^3\boldsymbol{y}\, |\phi_\alpha(\boldsymbol{x})|^2\; U(\boldsymbol{x}-\boldsymbol{y})\; |\phi_\beta(\boldsymbol{y})|^2\,; \tag{A.13}$$

and the corresponding Hamiltonian is:

$$\begin{aligned} H_{int}^{U'} &= \frac{1}{2}\sum_i \sum_{\alpha,\beta} \sum_{\sigma,\sigma'} U'^{\alpha,\beta} c^\dagger_{i,\alpha,\sigma}\, c^\dagger_{i,\beta,\sigma'}\, c_{i,\beta,\sigma'}\, c_{i,\alpha,\sigma} = \\ &= \frac{1}{2}\sum_i \sum_{\alpha,\beta} \sum_{\sigma,\sigma'} U'^{\alpha,\beta} n_{i,\alpha,\sigma}\, n_{i,\beta,\sigma'} = \\ &= \frac{1}{2}\sum_i \sum_{\sigma} \sum_{\alpha,\beta} U'^{\alpha,\beta} \left(n_{i,\alpha,\sigma}\, n_{i,\beta,\sigma} + n_{i,\alpha,\sigma}\, n_{i,\beta,-\sigma}\right) = \\ &= U' \sum_i \sum_{\sigma,\sigma'} n_{i,1,\sigma} n_{i,2,\sigma'}; \end{aligned} \tag{A.14}$$

in writing the second line the condition $\alpha \neq \beta$ was exploited to anti-commute the operators. This originates the on-site inter-band repulsion.

- Channels which are not diagonal in the occupation number are allowed as well; one of the two is given by the case with $\alpha_1 = \alpha_3 = \alpha \neq \beta = \alpha_2 = \alpha_4$. In this case, the *exchange* is given by:

$$J^{\alpha,\beta} = \frac{1}{2}\int d^3\boldsymbol{x}\, d^3\boldsymbol{y}\, \phi_\alpha(\boldsymbol{x})^*\, \phi_\beta(\boldsymbol{y})^*\, U(\boldsymbol{x}-\boldsymbol{y})\, \phi_\alpha(\boldsymbol{y})\, \phi_\beta(\boldsymbol{x}), \tag{A.15}$$

which, unlike the Coulomb integrals A.11 and A.13 is not expressed in terms of the square moduli of the wave functions. The corresponding Hamiltonian is easily written as the sum required to reproduce 1.12 by expanding the sums on the spin and band indices and noticing that terms sum up in pairs:

$$\begin{aligned} H_{int}^{J} &= \frac{1}{2}\sum_i \sum_{\alpha,\beta} \sum_{\sigma,\sigma'} J^{\alpha,\beta} c^\dagger_{i,\alpha,\sigma}\, c^\dagger_{i,\beta,\sigma'}\, c_{i,\alpha,\sigma'}\, c_{i,\beta,\sigma} = \\ &= J \sum_i \sum_{\sigma,\sigma'} c^\dagger_{i,1,\sigma} c^\dagger_{i,2,\sigma'} c_{i,1,\sigma'} c_{i,2,\sigma}. \end{aligned} \tag{A.16}$$

- The last of the on-site contributions is the pair-hopping term, which arises for $\alpha_1 = \alpha_2 = \alpha \neq \beta = \alpha_3 = \alpha_4$:

$$J'^{\alpha,\beta} = \frac{1}{2}\int d^3\boldsymbol{x}\, d^3\boldsymbol{y}\, \phi_\alpha(\boldsymbol{x})^*\, \phi_\alpha(\boldsymbol{y})^*\, U(\boldsymbol{x}-\boldsymbol{y})\, \phi_\beta(\boldsymbol{y})\, \phi_\beta(\boldsymbol{x}), \tag{A.17}$$

along with its contribution (note that in order to pair-hop, electrons must satistfy the additional condition $\sigma \neq \sigma'$):

$$\begin{aligned} H_{int}^{J'} &= \frac{1}{2}\sum_i \sum_{\alpha,\beta} \sum_{\sigma,\sigma'} J'^{\alpha,\beta} c^\dagger_{i,\alpha,\sigma}\, c^\dagger_{i,\alpha,\sigma'}\, c_{i,\beta,\sigma'}\, c_{i,\beta,\sigma} = \\ &= \frac{1}{2}\sum_i \sum_{\alpha,\beta} J'^{\alpha,\beta} \left(c^\dagger_{i,\alpha,\uparrow} c^\dagger_{i,\alpha,\downarrow} c_{i,\beta,\downarrow} c_{i,\beta,\uparrow} + c^\dagger_{i,\alpha,\downarrow} c^\dagger_{i,\alpha,\uparrow} c_{i,\beta,\uparrow} c_{i,\beta,\downarrow}\right) = \\ &= -J' \sum_i \left(c^\dagger_{i,1,\uparrow} c_{i,1,\downarrow} c^\dagger_{i,2,\downarrow} c_{i,2,\uparrow} + \text{h.c.}\right). \end{aligned} \tag{A.18}$$

These 4 contributions sum up the on-site interaction for the two-band model and together with the kinetic energy give the form 1.12. In principle, one could derive an extended Hubbard model with off-site interactions. However, the Hamiltonian becomes denser and the model too intricate; then again, multi-band Hubbard models are usually studied by considering only on-site interactions, an approximation which is indeed supported by the fact that off-site interactions are orders of magnitude smaller than the on-site ones.

Consider the case where the relevant orbitals are two out of three orbitals which have a t_{2g} symmetry, as would be the case for a transition metal in an octahedral crystal field. This situation corresponds to describing the ruthenate compounds in terms of a simplified two-orbital model, while at the same time retaining the correct orbital symmetry. Since tesseral harmonics are *real*, it immediately follows that $J = J'$. In order to show that $U' = U - 2J$, first observe that all the on-site interaction integrals are rotationally invariant, then write them in terms of the transformed basis and compare the two forms, taking into account the reality of the wave functions [38].

Bibliography

[1] **F. Bloch**. Über die Quantenmechanik der Elektronen in Kristallgittern. *Zeitschrift für Physik* 52 (1929), p. 555.

[2] **A. H. Wilson**. The Theory of Electronic Semi-Conductors. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 133 (1931), p. 458.

[3] **L. D. Landau**. The Theory of a Fermi Liquid. *Journal of Experimental and Theoretical Physics (U.S.S.R.)* 30 (1956), p. 1058.

[4] **J. H. de Boer and E. J. W. Verwey**. Semi-conductors with partially and with completely filled 3*d*-lattice bands. *Proceedings of the Physical Society* 49 (1937), p. 59.

[5] **N. F. Mott**. The Basis of the Electron Theory of Metals, with Special Reference to the Transition Metals. *Proceedings of the Physical Society, Section A* 62 (1949), p. 416.

[6] **M. C. Gutzwiller**. Effect of Correlation on the Ferromagnetism of Transition Metals. *Physical Review Letters* 10 (1963), p. 159.

[7] **J. Hubbard**. Electron Correlations in Narrow Energy Bands. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 276 (1963), p. 238.

[8] **J. Kanamori**. Electron Correlation and Ferromagnetism of Transition Metals. *Progress of Theoretical Physics* 30 (1963), p. 275.

[9] **J. Bardeen, L. N. Cooper and J. R. Schrieffer**. Theory of Superconductivity. *Physical Review* 108 (1957), p. 1175.

[10] **F. Steglich, J. Aarts, C. D. Bredl, W. Lieke, D. Meschede, W. Franz and H. Schäfer**. Superconductivity in the Presence of Strong Pauli Paramagnetism: $CeCu_2Si_2$. *Physical Review Letters* 43 (1979), p. 1892.

[11] **B. T. Matthias, T. H. Geballe and V. B. Compton**. Superconductivity. *Reviews of Modern Physics* 35 (1963), p. 1.

[12] **J. Nagamatsu, N. Nakagawa, T. Muranaka, Y. Zenitani and J. Akimitsu**. Superconductivity at $39°K$ in magnesium diboride. *Nature* 410 (2001), p. 63.

[13] **J. G. Bednorz and K. A. Müller**. Possible high T_c superconductivity in the BaLaCuO system. *Zeitschrift für Physik B Condensed Matter* 64 (1986), p. 189.

[14] **D. A. Wollman, D. J. Van Harlingen, W. C. Lee, D. M. Ginsberg and A. J. Leggett**. Experimental determination of the superconducting pairing state in YBCO from the phase coherence of $YBCO - Pb$ dc SQUIDs. *Physical Review Letters* 71 (1993), p. 2134.

[15] **J. R. Kirtley, C. C. Tsuei, A. Ariando, C. J. M. Verwijs, S. Harkema and H. Hilgenkamp**. Angle-resolved phase-sensitive determination of the in-plane gap symmetry in $YBa_2Cu_3O_{7-\delta}$. *Nature Physics* 2 (2006), p. 190.

[16] **F. C. Zhang and T. M. Rice**. Effective Hamiltonian for the superconducting Cu oxides. *Physical Review B* 37 (1988), p. 3759.

[17] **V. I. Anisimov, I. A. Nekrasov, D. E. Kondakov, T. M. Rice and M. Sigrist**. Orbital-selective Mott-insulator transition in $Ca_{2-x}Sr_xRuO_4$. *European Physics Journal B* 25 (2002), p. 191.

[18] **B. J. Kim et al.** Novel $J_{\text{eff}} = 1/2$ Mott State Induced by Relativistic Spin-Orbit Coupling in Sr_2IrO_4. *Physical Review Letters* 101 (2008), p. 076402.

[19] **K. Haule, J. H. Shim and G. Kotliar**. Correlated Electronic Structure of $LaO_{1-x}F_xFeAs$. *Physical Review Letters* 100 (2008), p. 226402.

[20] **M. Imada, A. Fujimori and Y. Tokura**. Metal-insulator transitions. *Reviews of Modern Physics* 70 (1998), p. 1039.

[21] **E. H. Lieb and F. Y. Wu**. Absence of Mott Transition in an Exact Solution of the Short-Range, One-Band Model in One Dimension. *Physical Review Letters* 20 (1968), p. 1445.

[22] **M. R. Norman**. The Challenge of Unconventional Superconductivity. *Science* 332 (2011), p. 196.

[23] **M. R. Norman et al.** Destruction of the Fermi surface in underdoped high-T_c superconductors. *Nature* 392 (1998), p. 157.

[24] **P. Fazekas**. Lecture notes on correlation and magnetism. Ed. by I. Dzyaloshinski, S. Lundqvist and Y. Lu. Vol. 5. Series in Modern Condensed Matter Physics. World Scientific Publishing, 1999.

[25] **P. Fazekas and P. W. Anderson**. On the ground state properties of the anisotropic triangular antiferromagnet. *Philosophical Magazine* 30 (1974), p. 423.

[26] **H. Bethe**. Zur Theorie der Metalle. *Zeitschrift für Physik* 71 (1931), p. 205.

[27] **P. W. Anderson**. The Resonating Valence Bond State in La_2CuO_4 and Superconductivity. *Science* 235 (1987), p. 1196.

[28] **D. Poilblanc, P. Corboz, N. Schuch and J. I. Cirac**. Resonating-valence-bond superconductors with fermionic projected entangled pair states. *Physical Review B* 89 (2014), p. 241106.

[29] **G. Kotliar and J. Liu**. Superexchange mechanism and *d*-wave superconductivity. *Physical Review B* 38 (1988), p. 5142.

[30] **C. Gros**. Physics of projected wavefunctions. *Annals of Physics* 189 (1989), p. 53.

[31] **B. Edegger, V. N. Muthukumar and C. Gros**. Gutzwiller-RVB theory of high-temperature superconductivity: Results from renormalized mean-field theory and variational Monte Carlo calculations. *Advances in Physics* 56 (2007), p. 927.

[32] **P. W. Anderson**. Present status of the theory of the high-T_c cuprates. *Low Temperature Physics* 32 (2006), p. 282.

[33] **L. de'Medici, A. Georges and S. Biermann**. Orbital-selective Mott transition in multiband systems: Slave-spin representation and dynamical mean-field theory. *Physical Review B* 72 (2005), p. 205124.

[34] **K. Inaba and A. Koga**. Phase diagrams of the two-orbital Hubbard model with different bandwidths. *Physical Review B* 73 (2006), p. 155106.

[35] **M. J. Rozenberg**. Integer-filling metal-insulator transitions in the degenerate Hubbard model. *Physical Review B* 55 (1997), R4855.

[36] **N. Manini, G. E. Santoro, A. Dal Corso and E. Tosatti**. Sensitivity of the Mott transition to noncubic splitting of the orbital degeneracy: Application to $NH_3K_3C_{60}$. *Physical Review B* 66 (2002), p. 115107.

[37] **P. W. Anderson**. Localized Magnetic States in Metals. *Physical Review* 124 (1961), p. 41.

[38] **C. Castellani, C. R. Natoli and J. Ranninger**. Magnetic structure of V_2O_3 in the insulating phase. *Physical Review B* 18 (1978), p. 4945.

[39] **Y. Maeno, H. Hashimoto, K. Yoshida, S. Nishizaki, T. Fujita, J. G. Bednorz and F. Lichtenberg**. Superconductivity in a layered perovskite without copper. *Nature* 372 (1994), p. 532.

[40] **Y. Maeno, S. Kittaka, T. Nomura, S. Yonezawa and K. Ishida**. Evaluation of Spin-Triplet Superconductivity in Sr_2RuO_4. *Journal of the Physical Society of Japan* 81 (2012), p. 011009.

[41] **M. Yi et al.** Observation of Temperature-Induced Crossover to an Orbital-Selective Mott Phase in $A_xFe_{2-y}Se_2$ (A=K, Rb) Superconductors. *Physical Review Letters* 110 (2013), p. 067003.

[42] **A. Liebsch**. Absence of orbital-dependent Mott transition in $Ca_{2-x}Sr_xRuO_4$. *Europhysics Letters* 63 (2003), p. 97.

[43] **A. Liebsch**. Single Mott transition in the multiorbital Hubbard model. *Physical Review B* 70 (2004), p. 165103.

[44] **T. A. Costi and A. Liebsch**. Quantum Phase Transition in the Two-Band Hubbard Model. *Physical Review Letters* 99 (2007), p. 236404.

[45] **M. Ferrero, F. Becca, M. Fabrizio and M. Capone**. Dynamical behavior across the Mott transition of two bands with different bandwidths. *Physical Review B* 72 (2005), p. 205126.

[46] **L. de'Medici**. Hund's coupling and its key role in tuning multiorbital correlations. *Phys. Rev. B* 83 (2011), p. 205112.

[47] **L. de'Medici, J. Mravlje and A. Georges**. Janus-Faced Influence of Hund's Rule Coupling in Strongly Correlated Materials. *Physical Review Letters* 107 (2011), p. 256401.

[48] **N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller**. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics* 21 (1953), p. 1087.

[49] **S. Sorella**. Generalized Lanczos algorithm for variational quantum Monte Carlo. *Physical Review B* 64 (2001), p. 024512.

[50] **S. Sorella**. Wave function optimization in the variational Monte Carlo method. *Physical Review B* 71 (2005), p. 241103.

[51] **S. Yunoki and S. Sorella**. Two spin liquid phases in the spatially anisotropic triangular Heisenberg model. *Physical Review B* 74 (2006), p. 014408.

[52] **D. M. Ceperley and B. J. Alder**. Ground State of the Electron Gas by a Stochastic Method. *Physical Review Letters* 45 (1980), p. 566.

[53] **W. L. McMillan**. Ground State of Liquid He^4. *Physical Review* 138 (1965), A442.

[54] **L. Capriotti, F. Becca, A. Parola and S. Sorella**. Resonating Valence Bond Wave Functions for Strongly Frustrated Spin Systems. *Physical Review Letters* 87 (2001), p. 097201.

[55] **S. Sorella, G. B. Martins, F. Becca, C. Gazza, L. Capriotti, A. Parola and E. Dagotto**. Superconductivity in the Two-Dimensional $t - J$ Model. *Physical Review Letters* 88 (2002), p. 117002.

[56] **R. Rüger**. Implementation of the Variational Monte Carlo method for the Hubbard model. MA thesis. Institut für Theoretische Physik - Johann Wolfgang Goethe-Universität, 2013.

[57] **R. Jastrow**. Many-Body Problem with Strong Forces. *Physical Review* 98 (1955), p. 1479.

[58] **M. Capello**. Variational description of Mott insulators. PhD thesis. Scuola Internazionale Superiore di Studi Avanzati (ISAS), 2006.

[59] **L. F. Tocchio, F. Becca and C. Gros**. Backflow correlations in the Hubbard model: An efficient tool for the study of the metal-insulator transition and the large-U limit. *Physical Review B* 83 (2011), p. 195138.

[60] **M. Lugas, L. Spanu, F. Becca and S. Sorella**. Finite compressibility in the low-doping region of the two-dimensional t-J model. *Physical Review B* 74 (2006), p. 165122.

[61] **A. J. Leggett**. Quantum Liquids. Oxford University Press, 2011.

[62] **A. Paramekanti, M. Randeria and N. Trivedi**. High-T_c superconductors: A variational theory of the superconducting state. *Physical Review B* 70 (2004), p. 054504.

[63] **R. P. Feynman**. Atomic Theory of the Two-Fluid Model of Liquid Helium. *Physical Review* 94 (1954), p. 262.

[64] **A. W. Overhauser**. Simplified Theory of Electron Correlations in Metals. *Physical Review B* 3 (1971), p. 1888.

[65] **L. F. Tocchio, C. Gros, X.-F. Zhang and S. Eggert**. Phase Diagram of the Triangular Extended Hubbard Model. *Physical Review Letters* 113 (2014), p. 246405.

[66] **M. Capello, F. Becca, M. Fabrizio, S. Sorella and E. Tosatti**. Variational Description of Mott Insulators. *Physical Review Letters* 94 (2005), p. 026406.

[67] **M. Capello, F. Becca, S. Yunoki and S. Sorella**. Unconventional metal-insulator transition in two dimensions. *Physical Review B* 73 (2006), p. 245116.

[68] **M. Capello, F. Becca, M. Fabrizio and S. Sorella**. Superfluid to Mott-Insulator Transition in Bose-Hubbard Models. *Physical Review Letters* 99 (2007), p. 056402.

[69] **L. Reatto and G. V. Chester**. Phonons and the Properties of a Bose System. *Physical Review* 155 (1967), p. 88.

[70] **O. Gunnarsson, E. Koch and R. M. Martin**. Mott transition in degenerate Hubbard models: Application to doped fullerenes. *Physical Review B* 54 (1996), R11026.

[71] **A. P. Mackenzie and Y. Maeno**. The superconductivity of Sr_2RuO_4 and the physics of spin-triplet pairing. *Reviews of Modern Physics* 75 (2003), p. 657.