



COMPUTER SIMULATION OF PHYSICAL SYSTEMS I SEMESTER PROJECT

Gutzwiller Projection Monte Carlo for a $SU(4)$ Symmetric Heisenberg Model in a Honeycomb Lattice

Francisco Monteiro de Oliveira Brito

December 20, 2017

course taught by
Prof. Alfredo PASQUARELLO

supervised by
Prof. Frédéric MILA¹

cosupervised by
Dr. Miklós LAJKÓ ¹

¹Condensed Matter Theory Group, Institute of Physics, EPFL Rte de la Sorge, CH-1015 Lausanne, Switzerland

Abstract

We use a variational Monte Carlo method to study correlations and average energy in a generalized $SU(4)$ symmetric Heisenberg Model on the honeycomb lattice. This model can be regarded as the $t/U \ll 1$ limit of a corresponding $SU(4)$ symmetric Hubbard model. The study of the model is motivated by the fact that it hosts a quantum spin-orbital liquid ground state.

We propose a novel variational ground state based on adding next nearest neighbor complex hopping to the tight-binding free fermion variational hamiltonian, and investigate if it constitutes an improvement over the nearest neighbor case. In particular, we argue that the variational state we propose might capture a topological insulator phase within the model. The approach is inspired by the work of Haldane [5]. He showed that a topologically insulating phase can arise due to time-reversal symmetry breaking on the honeycomb lattice. In Haldane's model this comes from adding next nearest neighbor complex hopping to the usual nearest neighbor tight-binding model. This may be realized in practice, for example, by applying a staggered magnetic field.

We measure first, second, and third nearest neighbor Heisenberg exchange energies for varying variational parameters t_2 , and ϕ_t , specifying, respectively, the hopping strength and a phase introduced in the next nearest neighbor hopping. A phase diagram is obtained, suggesting the existence of a topological insulator transition within this model.

Contents

I	Introduction	3
1	Mott insulators	3
2	Magnetic excitations	3
3	Orbital degeneracy	4
4	Quantum Spin Liquids	4
II	Motivation	5
5	From the usual spin- $\frac{1}{2}$ $SU(2)$ symmetry to $SU(N)$	9
6	$SU(4)$ Heisenberg Model	9
7	$SU(N)$ spins	10
8	Fermionic Mean Field and Gutzwiller Projection	11
III	Method	12
9	Color-color correlations	13
10	Markov Chain. Detailed Balance. Metropolis Algorithm	14
11	Efficient determinant update	15
12	Measurements	18
13	Bond energies	18
14	MC algorithm outline and code flowchart	21
IV	Benchmarks	23
15	Naïve calculation of determinant ratios vs. Efficient update	23
16	Warm-up time	26

17	Auto-correlation time	26
18	Correlation function measurements	26
V	Discussion	27
19	Introducing a novel variational state	27
20	Correlation measurements	30
21	Phase Diagram	31
VI	Outlook	32

Part I

Introduction

1 Mott insulators

One of the most important problems in solid state theory is that of describing electrons in a solid. A highly successful theory in many respects is the familiar band theory. However, this theory is fundamentally flawed. The approximation of independent electrons breaks down for a class of materials known as Mott insulators.

The simple picture of electron bands of allowed energies separated by forbidden gaps may be formulated in two ways. A possible approach is to consider the heavy nuclei at rest, while the much more mobile valence electrons move under a potential created by the ions. The quadratic free electron dispersion relation is distorted due to this periodic potential. The resulting dispersion relation is characterized by the opening of gaps. Alternatively, we can imagine that valence electrons hop between the valence shells of neighboring atoms - tight-binding model. It is natural that the probability of hopping is given by the overlap of the neighboring atomic wave functions. The electronic bands are now not formed through the opening of gaps, but they appear instead due to the broadening of atomic levels.

The ground state is formed by filling the doubly degenerate energy levels successively until we run out of available electrons, at which point we say that the Fermi energy is reached. The nature of this so called Fermi sea state allows us to classify materials as insulators or conductors. Let us analyze the predictions of band theory. Generally, for the case of an even number of electrons in the unit cell, we have an insulator. The Fermi energy is located between the completely filled band of highest energy and the empty band of lowest energy. This implies that electron excitations are only possible by overcoming the gap between the fully occupied and the unoccupied band, and we have an insulator. Only in the particular case where this gap vanishes, that is the bands cross, do we have a conductor. For the case of an odd number of electrons per unit cell, a partially filled band necessarily exists, and the material is a conductor.

Some materials, among which we find many transition-metal oxides have been found to contradict the predictions of band theory. Although predicted to be conductors by band theory, they are actually (Mott) insulators. A minimal model that encapsulates the effect of electron-electron Coulomb repulsion leading to this effect is the Hubbard model. The idea is to consider a tight-binding hamiltonian plus a term that penalizes the existence of two opposite-spin electrons on the same site.

$$\mathcal{H} = -t \sum_{\langle i,j \rangle, \sigma} \left(c_{i,\sigma}^\dagger c_{j,\sigma} + c_{j,\sigma}^\dagger c_{i,\sigma} \right) + \sum_i U n_{i,\uparrow} n_{i,\downarrow} \quad (1)$$

,where a $c_{i,\sigma}^{(\dagger)}$ -operator annihilates(creates) an electron at site i with spin σ . Here, t is the hopping integral, while U is the Coulomb energy of two opposite-spin electrons on the same orbital.

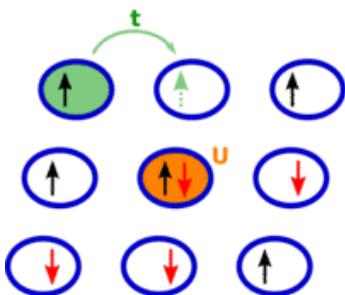


Figure 1: An illustration of the Hubbard model taken from https://www.ifp.tuwien.ac.at/cms/research/many-body_quantum_field_theory/

If $U = 0$, the probability that an electron occupying a given orbital has spin up is the same as that of having spin down. Thus, considering one electron per unit cell, a site of the Fermi sea ground state is singly occupied with probability $\frac{1}{2}$ and doubly occupied or empty with equal probabilities $\frac{1}{4}$. As the Coulomb repulsion U increases, doubly occupied and empty sites are suppressed. Electrons tend to become localized, and the material may eventually make a transition to a Mott insulating phase induced by electron correlations.

2 Magnetic excitations

In the $t/U \ll 1$ limit, a ground state configuration with one electron per site is favored. Applying second order degenerate perturbation theory to this 2^N -fold degenerate ground state, we find that an antiparallel spin configuration is preferred. Intuitively, this may be regarded as being due to a process of an electron hopping

to a neighboring site to lower the energy, that is only possible in such a configuration. The Pauli exclusion principle forbids this process for a parallel spin configuration.

The effective low-energy hamiltonian describing the lifting of the 2^N -fold degeneracy of the ground state is

$$\mathcal{H}_{eff} = -\frac{4t^2}{U} \sum_{\langle i,j \rangle} \left(\underbrace{1 - \frac{(\mathbf{S}_i \cdot \mathbf{S}_j)^2}{2}}_{(*)} \right) \quad (2)$$

where $(*)$ is 1 if the spins on i and j form a singlet, and 0 if they form a triplet. Noting that

$$\frac{(\mathbf{S}_i + \mathbf{S}_j)^2}{2} = \mathbf{S}_i \cdot \mathbf{S}_j + \frac{3}{4} \quad (3)$$

we recast the effective hamiltonian as an antiferromagnetic Heisenberg model.

As usual, the picture is not so pretty in most compounds. Magnetic sites hosting unpaired electrons are often connected through non magnetic ions with closed orbitals. This means the effective exchange interaction may stem from consecutive virtual hoppings. The effective low energy model will then contain higher order spin terms, and can even become ferromagnetic in certain situations.

3 Orbital degeneracy

In general each magnetic site has more than one orbital in which the electrons may be placed. Consequently, we will have orbital-dependent hopping amplitudes, and a varying Coulomb repulsion, depending on whether the electrons on the same site occupy the same orbital with opposite spins, or different orbitals.

For the case of 2-fold orbital degeneracy, pseudo spin- $\frac{1}{2}$ operators may be introduced, acting on the orbital states $|a\rangle$ and $|b\rangle$, which we conveniently rename $|a\rangle = |+\rangle$ and $|b\rangle = |-\rangle$, so that

$$\tau^z |\pm\rangle = \pm \frac{1}{2} |\pm\rangle \quad (4)$$

,while ladder operators τ^\pm connect the orbitals.

A simple example of a spin-orbital hamiltonian is the Kugel-Khomskii model [13]:

$$\mathcal{H}_{KK} = \sum_{\langle i,j \rangle} \left[u + \mathbf{S}_i \cdot \mathbf{S}_j \right] \otimes \left[v + \alpha(\tau_i^+ \tau_j^- + \tau_i^- \tau_j^+) + J'_z \tau_i^z \tau_j^z \right] \quad (5)$$

for some constants u and v , and interaction parameters α and J'_z . The spin part is isotropic, that is invariant under spin rotation. If we forbid hoppings of the type $|a\rangle \rightarrow |b\rangle$, the orbital part is $U(1)$ symmetric. However, in general if those hoppings are allowed, $\tau_i^+ \tau_j^+$, $\tau_i^- \tau_j^-$, or even single τ_i^+ , τ_i^- terms are possible. Many perturbative hopping processes exist in the low energy limit due to spin and orbital degeneracy.

A special case arises if one takes $\alpha = J'_z/2$: the orbital part is made $SU(2)$ symmetric, like the spin part. An even more special case of particular relevance to us is the one for which \mathcal{H}_{KK} is $SU(4)$ symmetric, at which point the spin and orbital states become equivalent, and the interaction corresponds to an exchange of a single spin-orbital degree of freedom.

4 Quantum Spin Liquids

A relevant question is whether the ground state of a given model shows any kind of ordering, or symmetry breaking, and if so, whether we can find an order parameter describing the transition between the ordered and disordered phases.

Frustrated interactions and quantum fluctuations may deem these questions very difficult to answer satisfactorily. This is because a very large number of possible orderings and ground state configurations exist in general.

Surprisingly enough, a semi-classical approach generally provides invaluable insight. It consists of neglecting the quantum nature of the problem, that is the entanglement of the spins. This is particularly true for ferromagnetic interactions: the classical ground state coincides with the quantum mechanical one, for which fluctuations ΔS vanish in the ground state. Then, low energy excitations are simply treated in the framework of linear spin wave theory.

On the other hand, if an antiferromagnet has soft enough excitation spectrum, quantum fluctuations will destroy the classical Néel long range order. Thus, it becomes natural to ask what is the structure of the ground state and of the excitations above it. There are a number of situations in which quantum fluctuations can stabilize a ground state which is essentially a product of wave functions of local singlets. The order parameter for an antiferromagnetic system is the staggered magnetization density. Since $[m_{st}, H_{Heis}] \neq 0$, we must define

$$m_{st} = \frac{1}{L} \left\langle \left(\sum_{j \in A} S_j^z - \sum_{m \in B} S_m^z \right) \right\rangle_{GS} \quad (6)$$

where we consider a bipartite lattice, constituted by lattices A and B , and a linear sample dimension L .

We may define an effective spin reduction, such that

$$m_{st} = S - \Delta S \quad (7)$$

The effect of ΔS will then depend on the dimensionality D . For example, for $D = 1$, a simple argument allows us to conclude that it diverges logarithmically. A divergent spin reduction constitutes a contradiction. This is due to the assumption of antiferromagnetic long-range order in the ground state being wrong in the first place: for a 1D Heisenberg model, quantum fluctuations destroy the long-range order for any finite spin S . We emphasize that $\Delta S = \infty$ cannot be taken literally. A more careful argument gives $\Delta S = S$ [4].

In spite of the approximate reasoning, the conclusion above is strictly valid. The total spin of the ground state must be $S_{tot} = 0$. While $D = 1$ isotropic AF Heisenberg magnets do not order at all, we can still find a limit of extreme quantum AF with arbitrarily small but finite m_{st} , for example for quasi 1D systems[4].

In low-dimensional, small spin systems, quantum entanglement is relevant, with the classical approach being unable to capture the relevant physics. The formation of antisymmetric singlet states of two spins- $\frac{1}{2}$, called valence bonds, leads precisely to the instance that is more favorable energetically than classical ordering.

A $S_{tot}^z = 0$ ground state is built from valence bond states, where two electron spins form a singlet due to the antiferromagnetic interaction. A valence bond solid corresponds to the case in which the spins are in a configuration of localized static bonds. Consequently, the VBS has lattice symmetry breaking, when the bonds order in a particular way. While the two spins forming the bond are maximally entangled, they are not entangled at all with all the other spins: a VBS ground state has no long-range entanglement. The picture changes, however, if we allow quantum fluctuations of the valence bonds. In that case, the obtained ground state is a superposition of distinct partitionings of the spins into valence bonds.

Resonating valence bond ground states are those for which long valence bonds are present. They are found by allowing valence bonds to jump anywhere, and switch from vertical to horizontal directions, in a resonating scheme (see figure below). These ground states consist of a tensor product of directed singlet dimer states of spins. RVB ground states are ubiquitous, having been found to arise for numerous models. The $S = 1/2$ triangular lattice has been found to have a RVB ground state by Anderson in 1973 [1], constituting the first example of a quantum spin liquid (QSL). Since then, the system has been extensively studied, but there is still no consensus on whether some partial long range order actually exists or not. Many more QSL have been found ever since, leading to a growing interest within the scientific community. Their main characteristic is that no lattice or spin symmetry breaking is found, making their experimental identification challenging.

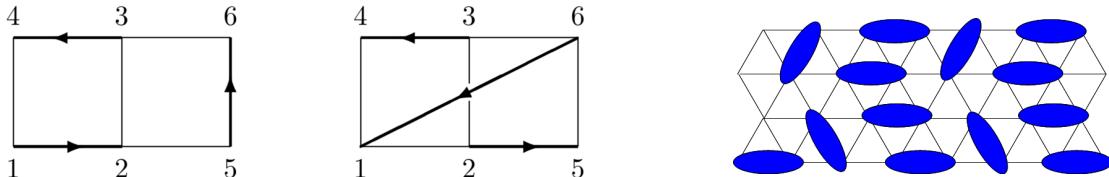


Figure 2: On the left a Rumer diagram illustrating resonating valence bonds, and on the right resonating valence bonds on the triangular lattice, showing lattice symmetry breaking.

The aim of this project is to use a numerical method to identify and characterize a model which has a quantum spin liquid ground state: the $SU(4)$ Heisenberg model on the honeycomb lattice.

Part II

Motivation

In a particularly symmetry instance, the Kugel-Khomskii model turns out to be a straightforward generalization of the $SU(2)$ symmetric spin- $\frac{1}{2}$ Heisenberg model[3]. The resulting $SU(4)$ symmetric Heisenberg model can thus be regarded as a low energy limit model of the Mott insulating phase of certain materials showing relevant spin-orbit coupling. Examples of these materials have a high degree of symmetry in both the spin and orbital degrees of freedom, leading to a spin-orbit coupling that is effectively approximated by an $SU(4)$ Heisenberg model. On the other hand, the Mott insulating phase of alkaline-earth atoms trapped on an optical lattice can also

be described by this model. In this system, the only relevant degree of freedom is the nuclear spin F . Each of these spins has $N = 2F + 1$ possible states. The interaction between nuclear spins may be modeled by a $SU(N)$ Heisenberg model[2]. One of the recent advances in the understanding of this highly symmetric version of the Kugel-Khomskii model has been the finding that it allows a unique phase in the honeycomb lattice: a spin-orbital liquid phase with algebraically decaying correlations[9].

Mott insulators allow low-energy spin fluctuations. Moreover, the partial filling of crystal-field levels may lead Mott insulators to host orbital degrees of freedom. In general, long-range order is established for both spin and orbital degrees of freedom. However, the ground state can also correspond to a spin-orbital quantum liquid. This refers to a system in which neither lattice nor $SU(N)$ symmetry breaking occurs. Corboz et al.[3] found a particular case of this quantum liquid ground state. It was encountered by studying the Kugel-Khomskii model[7], a minimal model of spin-orbit coupling in Mott insulators on the honeycomb lattice.

It has been shown that this spin liquid phase is prone to the spontaneous formation of $SU(4)$ singlet structures, named plaquettes, in a process called tetramerization. Using the same Variational Monte Carlo method to be used in this project, it was found that the algebraic liquid resists this unstable tendency towards tetramerization, requiring a small but finite next-nearest neighbor hopping exchange to induce it.

While early investigations of the Kugel-Khomskii model on the simple lattices, like the square and the triangular lattices, have shown the importance of four-site plaquettes, these do not describe the ground state on the honeycomb lattice. These plaquettes are the natural unit to form $SU(4)$ singlets, or tetramers. The possibility of plaquette long-range order was found for the square lattice, and a variational subspace constructed in terms of singlet plaquettes is relevant in describing the ground states properties of the $SU(4)$ Heisenberg model in the simple triangular and square lattices [3]. However, we will see, in the honeycomb lattice a unique quantum liquid ground state phase prevails instead.

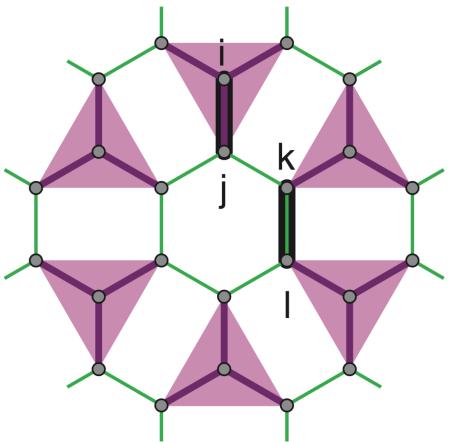


Figure 3: Some insight is gained by constructing a hamiltonian \mathcal{H}_Q for which a $SU(4)$ singlet-product is an exact ground state on the honeycomb lattice. The colored triangles correspond to such singlets, constituted by four spins connected by purple bonds. These singlet tetramers are all decoupled, and their stabilization requires a modulation of bond energies on the honeycomb lattice. Along with the green bonds between spins belonging to different tetramers, there must be two distinct bonds (thicker, in black) in each cell for the tetramer configuration to correspond to the ground state of \mathcal{H}_Q (taken from [9]).

Let us now turn to orbital degrees of freedom. Usually their energy scale is quite different from that of their spin counterparts. As a real world example of the use of this characteristic, we mention the experimental investigation of orbital fluctuations by resonant inelastic x-ray scattering in transition-metal oxides. It turns out that it has recently been boosted under specific conditions relating to this energy difference. If the crystal-field splitting is strong enough, a single orbital configuration will correspond to the ground state and the remainder are "repelled" away, corresponding to much higher energy configurations. This effectively separates them from the low-lying magnetic excitations. However, it is also possible that orbital fluctuations and magnetic excitations have comparable energy. It so happens that for a large class of systems orbital fluctuations in the ground state are suppressed. This happens due to a cooperative Jahn-Teller distortion. The interaction between localized orbital electronic states and the crystal lattice leads to a lifting of the orbital degeneracy. The spatially degenerate state distorts, seeking a lower energy equilibrium. Equilibrium is then reached, yielding a lower symmetry state. The Jahn-Teller distortion leads to orbital order with gapped excitations.

In contrast, there are situations in which fluctuating orbitals survive in the ground state. This is a currently active area of research. One of the materials that was thought to allow this peculiar characteristic was $Ba_3CuSb_2O_9$ [8]. There seemed to be no orbital order on the decorated honeycomb lattice in which the Cu^{++} ions in this compound were thought to condense.

The Jahn-Teller distortion leading to orbital order seems to be absent in $Ba_3CuSb_2O_9$. X-ray scattering studies show a persistence of orbital fluctuations down to very low temperatures and no magnetic long range ordering appears upon measuring the magnetic susceptibility. This is the reason why this compound was thought to be a good candidate to realize the spin-orbital liquid. While this result provided some motivation to study the model at stake here, which led to valuable further investigation, it was recently shown that the honeycomb lattice is not actually realized in this material [14]. Instead, $\alpha - ZrCl_3$ seems to be a stronger candidate to host a quantum spin liquid state, by extending the spin space symmetry from $SU(2)$ to $SU(4)$. Yamada et al.

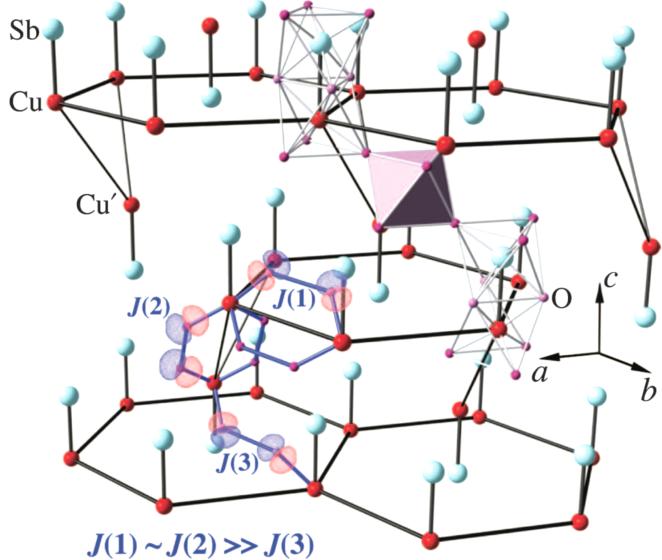


Figure 4: Crystal structure of $Ba_3CuSb_2O_9$. The center of the touching octahedra are alternately occupied by Cu and Sb atoms. Three sublattice ferrielectric order induced by these Cu - Sb dumbbells (that in turn form a triangular lattice) was thought to imply that the Cu^{++} ions lived in a honeycomb lattice; the ferrielectric order being due to the electric dipole moment carried by the dumbbells. It was thought that the Cu^{++} formed honeycomb lattices with weak interlayer coupling (taken from [8]).

recently proposed a mechanism in which $SU(4)$ symmetry emerges in the limit of strong spin-orbit coupling[11]. In some d^1 transition metal compounds, strong spin-orbit coupling leads to a strongly bond-dependent hopping between specific states, which on the $\alpha-ZrCl_3$ honeycomb structure may be shown to map to a $SU(4)$ Hubbard model through a gauge transformation. The model can in turn be generalized to other 3D lattices affording the proposal of crystalline spin-orbital liquids due to $SU(4)$ and space group symmetries.

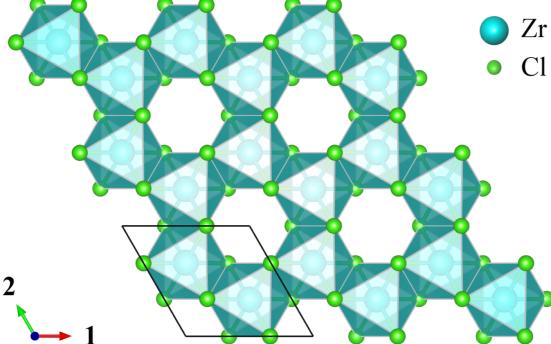


Figure 5: Honeycomb lattice geometric structure in $\alpha - ZrCl_3$. Crystallographic axes shown as 1, 2 (taken from [11]).

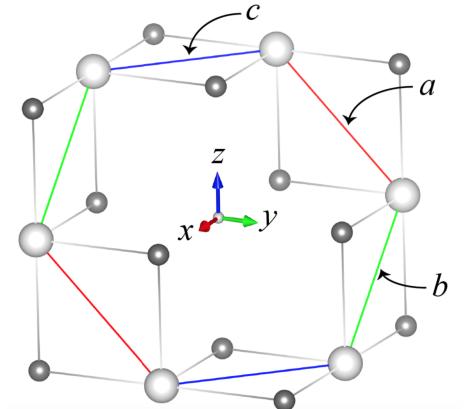


Figure 6: In $\alpha - ZrCl_3$ three different types of bonds are at play, shown in red, green, and blue. The bigger spheres correspond to Zr , while the smaller ones correspond to Cl (taken from [11]).

We aim to propose a novel variational state to be used in Variational Monte Carlo simulations of the $SU(4)$ symmetric Heisenberg model on the honeycomb lattice. The motivation for this approach comes from the work of Haldane [5]. He studied a model that includes not only nearest neighbor hopping, but also complex nearest neighbor hopping on a honeycomb lattice. In principle this configuration can be realized due to a *local* staggered magnetic field.

Solving the tight-binding hamiltonian on a honeycomb lattice yields a band degeneracy at the corners of the first Brillouin zone. The two non equivalent points of the lattice where the bands intersect, K and K' , are called band crossing points. Near them, the dispersion relation is approximately linear, corresponding to two fermions with opposite chirality, that is a massless Dirac fermion. The Dirac point corresponding to the band

crossing points in the honeycomb lattice is stable if space inversion and time reversal symmetry are preserved. Note that introducing longer range hoppings does not change this picture. Only if the hoppings are complex does the picture change.

In the presence of a magnetic field or spin-orbit coupling, the Aharonov-Bohm effect occurs. A particle moving around a closed contour acquires a phase difference between its final and initial states proportional to the magnetic flux enclosed by the contour $\phi = \frac{e}{\hbar} \iint \mathbf{B} \cdot d\mathbf{S} = \frac{e}{\hbar} \oint \mathbf{A} \cdot d\mathbf{l}$. We consider a discrete version on a lattice. Consider 3 sites a, b, c, with respective hopping strengths t_{ab}, t_{bc}, t_{ca} . A particle hops from a to b, from b to c, and then back to a. The hopping around the loop is $|t_{ab}t_{bc}t_{ca}|e^{i(\phi_{ab}+\phi_{bc}+\phi_{ca})}$, corresponding to the particle picking up a phase $\phi_{ab} + \phi_{bc} + \phi_{ca} = \frac{e}{\hbar} \iint \mathbf{B} \cdot d\mathbf{S}$. The individual phases have no physical meaning; in fact they are not gauge invariant. The total phase around a loop, however, the magnetic flux, is gauge invariant. The magnetic fields corresponding to this complex hopping break the time-reversal symmetry because $\mathbf{B} \rightarrow -\mathbf{B}$ under a time reversal transformation (the sign of all phases needs to be flipped upon time-reversal).

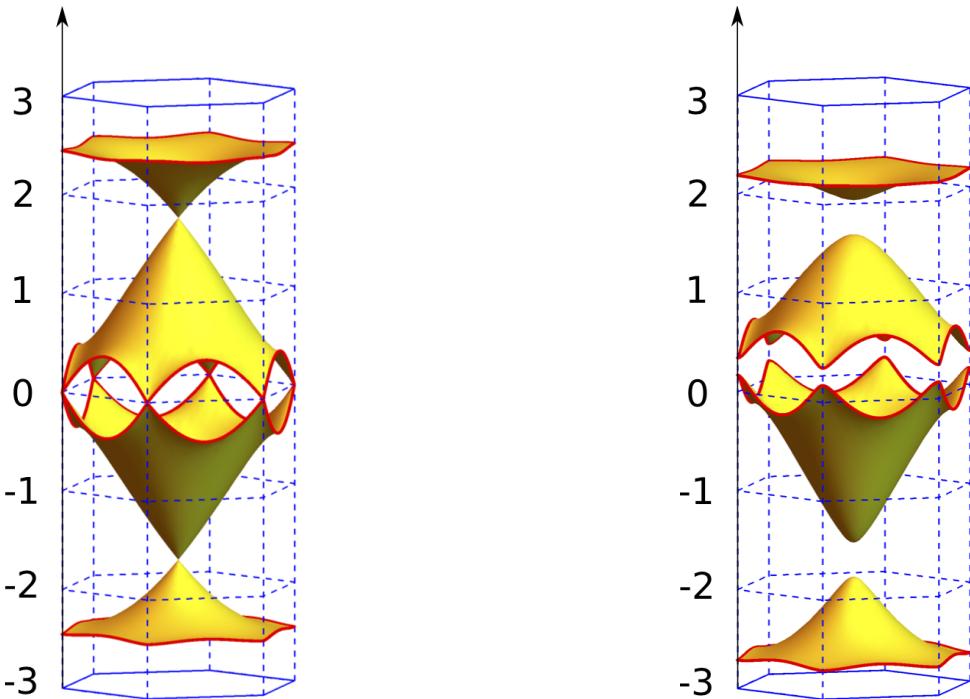


Figure 7: The dispersion relation computed for the model we analyze, showing the opening of a gap at the Dirac point when we turn on second nearest neighbor hopping (energy is on the vertical axis and position on the lattice on the horizontal plane).

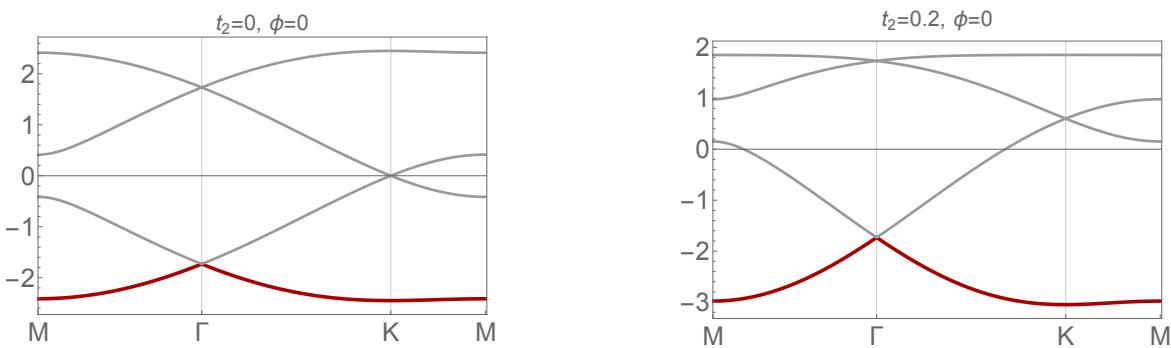


Figure 8: No opening of the gap at the K point arises if the next nearest neighbor hopping is real (the parameter specifying the phase we judiciously introduce in these complex hoppings is ϕ).

The breaking of time-reversal symmetry arising in this model is associated with a transition to a topological insulator phase. A gap in the dispersion relation is related to this transition. At the topological transition point, the gap closes, we have a band crossing again, and the system is not an insulator anymore. Topological insulators are generally associated with quantum Hall effect. In the case of the Haldane model we are in the presence of anomalous quantum Hall effect. A finite quantized Hall conductivity appears, and edge states are

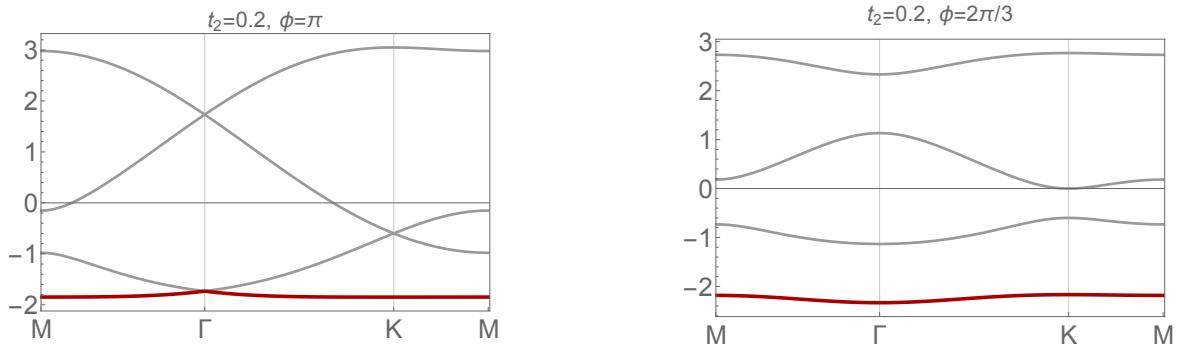


Figure 9: A phase of π still corresponds to real NN hopping, but with opposite sign. However if $\phi_t = 2\pi/3$, a gap appears, and the model allows a topological insulator phase.

present, just like in the case of regular quantum Hall effect. The crucial differences are that in this case the global magnetic field is zero, and while QHE requires a weak lattice potential, here the phenomenon arises precisely due to a very strong lattice background. Systems that have these characteristics appearing in Haldane's model are called topological Chern insulators. The model proposed by Haldane is the first example of a topological insulator beyond QHE, showing that the concept of a topological insulator is broad, and that they appear in cases other than quantum Hall effect.

Model

5 From the usual spin- $\frac{1}{2}$ $SU(2)$ symmetry to $SU(N)$

A spin- $\frac{1}{2}$ can be in one of two states $|\uparrow\rangle$ or $|\downarrow\rangle$. This is what deems the $S = \frac{1}{2}$ Heisenberg model $SU(2)$ symmetric. One can rewrite

$$\mathbf{S}_i \cdot \mathbf{S}_j = \frac{1}{2} P_{ij}^{(2)} - \frac{1}{4} \quad (8)$$

where P_{ij} is an exchange operator, "bringing" the spin on site j to site i and vice-versa.

Analogously, if the spins on each site have N states, we may introduce a $SU(N)$ symmetric Heisenberg model. Consider, for example, a $S = 1$ Heisenberg model. The three allowed spin states on each site are labeled $S^z = \pm 1, 0$. However, the model is not $SU(3)$ symmetric anymore because the interaction distinguishes between the spin states. They are not equivalent, in the sense that a transition between the states $S^z = 1$ and $S^z = -1$ is not allowed by an interaction term of the simple form $\mathbf{S}_i \cdot \mathbf{S}_j$. However, by introducing a biquadratic interaction, we recover the $SU(3)$ symmetry. Similarly to what was done for the $S = \frac{1}{2}$ case, we may introduce exchange operators

$$P_{ij}^{(3)} = (\mathbf{S}_i \cdot \mathbf{S}_j) + (\mathbf{S}_i \cdot \mathbf{S}_j)^2 - 1 \quad (9)$$

which now correspond to a $SU(3)$ symmetric model.

In general, if the model contains an interaction which may be written in terms of exchange operators P_{ij}^N , it is $SU(N)$ symmetric. The interaction is then invariant under application of a global transformation to the spins on each site.

6 $SU(4)$ Heisenberg Model

As we have seen, transition-metal oxides that preserve orbital degeneracy show relevant spin-orbit coupling. Their Mott-insulating phase can then be described by the Kugel-Khomskii model. The spatial extension of the degenerate orbitals makes the model anisotropic in the orbital part. Substantial simplification arises from considering a particular highly symmetric version of the model, actually showing $SU(4)$ symmetry, defined by

$$\mathcal{H} = \sum_{\langle i,j \rangle} \left(2\mathbf{S}_i \cdot \mathbf{S}_j + \frac{1}{2} \right) \left(2\mathbf{T}_i \cdot \mathbf{T}_j + \frac{1}{2} \right) \quad (10)$$

where $\langle i,j \rangle$ denotes a sum over nearest neighbors, \mathbf{S}_i is a spin- $\frac{1}{2}$ operator, and \mathbf{T}_i is a pseudospin- $\frac{1}{2}$ operator that takes the fluctuations of a doubly degenerate orbital into account. We assign the letters a and b to each

of these orbitals. We may define a basis in terms of redefined degrees of freedom that we name colors, say red (R), green (G), blue (B), and yellow (Y), such that

$$\begin{aligned} |R\rangle &= |\uparrow a\rangle \\ |G\rangle &= |\downarrow a\rangle \\ |B\rangle &= |\uparrow b\rangle \\ |Y\rangle &= |\downarrow b\rangle \end{aligned} \quad (11)$$

The hamiltonian is $SU(4)$ symmetric, which allows the introduction of exchange operators such that

$$\mathcal{H} = \sum_{\langle i,j \rangle} P_{i,j}^{(4)} \quad (12)$$

from where we can see that the model is a natural generalization of the Heisenberg model.

The $SU(4)$ Heisenberg model may be regarded as an effective low energy hamiltonian of the $SU(4)$ Hubbard model. For generality, we consider complex hoppings. The Hubbard hamiltonian is

$$\mathcal{H} = - \sum_{\langle i,j \rangle, \alpha} \left(t_{ij} c_{i,\alpha}^\dagger c_{j,\alpha} + t_{ij}^* c_{j,\alpha}^\dagger c_{i,\alpha} \right) + \sum_{i,\alpha < \beta} U n_{i,\alpha} n_{i,\beta} = H_t + H_U \quad (13)$$

where α and β are color indices.

We will set up a perturbative procedure. Performing a canonical transformation and projecting onto the sector with single fermion site occupancy, we obtain a second order effective hamiltonian in the limit $|t_{ij}| \ll U$.

$$\mathcal{H}_{eff} = P_s e^\eta \mathcal{H} e^{-\eta} P_s \quad (14)$$

where we introduced the anti-unitary operator $\eta \equiv (P_d H_t P_s - P_s H_t P_d)/U$, obeying $[\eta, H_U] = -H_t + P_d H_t P_d$. Here, P_s is the projector onto the subspace of single occupancy and $P_d = 1 - P_s$. Expanding the two-site permutation operator $P_{ij} = \sum_{\alpha\beta} c_{i\alpha}^\dagger c_{i\beta} c_{j\beta}^\dagger c_{j\alpha}$ to second order we obtain

$$H_{eff} = H_U - \frac{1}{U} P_s H_t P_d H_t P_s \quad (15)$$

Substituting H_t , the second order effective hamiltonian thus becomes

$$\begin{aligned} H_2 &= -\frac{1}{U} P_s \left(\sum_{\langle i,j \rangle} t_{ij} t_{ij}^* c_{i\alpha}^\dagger c_{j\alpha} c_{j\beta}^\dagger c_{i\beta} + t_{ij}^* t_{ij} c_{j\alpha}^\dagger c_{i\alpha} c_{i\beta}^\dagger c_{j\beta} \right) P_s \\ &= -\frac{1}{U} P_s \left(\sum_{\langle i,j \rangle} |t_{ij}|^2 (1 - P_{ij}) + |t_{ij}|^2 (1 - P_{ji}) \right) P_s \\ &= \frac{2}{U} P_s \left(\sum_{\langle i,j \rangle} |t_{ij}|^2 (P_{ij} - 1) \right) P_s \end{aligned} \quad (16)$$

This process can be generalized to yield higher order terms. Due to its symmetry, on the honeycomb lattice only even order terms contribute, so that in general we will have terms of the order $t^2/U, t^4/U^3, t^6/U^5 \dots$, with successively smaller contributions in the Mott insulating limit.

7 $SU(N)$ spins

$SU(N)$ spins have N possible colors, denoted by α . A general spin state is then written $|\psi\rangle = \sum_\alpha c_\alpha |\alpha\rangle$. A particularly important representation of the $SU(N)$ group is constructed in terms of $N \times N$ unitary matrices corresponding to transformations on these spins.

The operators that change the color of each spin $S_\beta^\alpha = |\alpha\rangle \langle \beta|$ form a set of infinitesimal generators of these transformations, obeying the commutation relations

$$[S_\beta^\alpha, S_\delta^\gamma] = \delta_\beta^\gamma S_\delta^\gamma - \delta_\delta^\alpha S_\beta^\gamma \quad \text{where } \alpha, \beta, \gamma, \delta = 1, \dots, N \quad (17)$$

The condition that $\sum_\alpha S_\alpha^\alpha = Id$ implies that the number of independent generators is $N^2 - 1$. Note that the operators S_β^α are the analogous of the ladder operators, familiar for the $SU(2)$ case. The addition of the more general $SU(N)$ spins is systematized in terms of Young-tableaux. These allow us to gather all the possible irreducible representations of the resulting spin. Going further into group theory is out of the scope of this report,

but we emphasize a key point. It is familiar from the $SU(2)$ case that two spins can form an antisymmetric pair, or singlet, invariant under spin rotation. This generalizes to the $SU(N)$ case. For concreteness, we give the example of the $SU(3)$ case. 3 spins have to be added to yield a totally antisymmetric singlet state. This state is then invariant under rotation of all spins by the same angle in spin space. The idea of spins forming singlets in $SU(N)$ acquires particular relevance here, in analogy with the formation of valence bonds in the $SU(2)$ case.

For models that consider only N fundamental $SU(N)$ spins on each site, representations in terms of either bosons or fermions exist, where the spin operators are recast in terms of fermionic annihilation and creation operators: $S_\beta^\alpha = f_\alpha^\dagger(i) f_\beta(i)$, where $f_\alpha^{(\dagger)}(i)$ annihilates/creates a fermion with color α at site i .

8 Fermionic Mean Field and Gutzwiller Projection

A $SU(N)$ symmetric hamiltonian is invariant under simultaneous rotation of all the spins, that is it commutes with all possible operators of the form $\sum_i S_\beta^\alpha(i)$. Operators that commute with the $\sum_i S_\beta^\alpha(i)$ are called Casimir operators. Taking an arbitrary subset of the spins Λ , we find the simplest non trivial such operator, the so called first Casimir operator:

$$C(\Lambda) = \sum_{\alpha, \beta} S(i)_\beta^\alpha S(j)_\alpha^\beta = \frac{1}{2} \left(C(\{i, j\}) - C(\{i\}) - C(\{j\}) \right) \quad (18)$$

where $\{\dots, k, \dots\}$ represents a subset of the spins. The first Casimir operator in $SU(2)$ is $C(\Lambda) = 2\mathbf{S}(\Lambda)^2 + \frac{1}{2}$, which is consistent with the familiar idea that $[\mathbf{S}^2, \mathbf{S}^z] = 0$. The $SU(2)$ Heisenberg interaction thus takes on the simple form

$$\mathbf{S}_i \cdot \mathbf{S}_j = \frac{1}{2} \left((\mathbf{S}_i + \mathbf{S}_j)^2 - \mathbf{S}_i^2 - \mathbf{S}_j^2 \right) \mapsto \frac{1}{2} S_{tot}(S_{tot} + 1) - S(S + 1) \quad (19)$$

where S_{tot} is the total spin on sites i and j .

Another relevant aspect is that the semi-classical approach that leads to linear spin wave theory can be generalized to $SU(N)$. In particular, if the wave function is site factorizable, there is no entanglement between the spins, and the picture is close to that of the classical case. Generically, as we add more colors, the number of possible configurations of the lattice increases. Therefore, it becomes more and more likely that two neighboring sites have different colors.

In the case of the $SU(2)$ Heisenberg model, the classical ground state manifold is degenerate. This degeneracy can be regarded as being lifted by linear spin wave theory. Similarly, linear flavor wave theory lifts the degeneracy of the $SU(4)$ ground state manifold by considering fluctuations in the bosonic representation of the ground state. These corrections favor orthogonal - differently colored - next-nearest neighbor "spins".

Analogously, for $SU(4)$, the honeycomb lattice can be filled with differently colored nearest and next-nearest neighbors. For a particularly symmetric configuration, the changing of the color along a stripe does not alter the energy. This affords a rough symmetry argument which allows one to see that the degeneracy, for a system of linear size L , is $\mathcal{O}(2^L)$. The exponentially large degeneracy suggests that the honeycomb lattice might be a good candidate to host a $SU(4)$ QSL ground state.

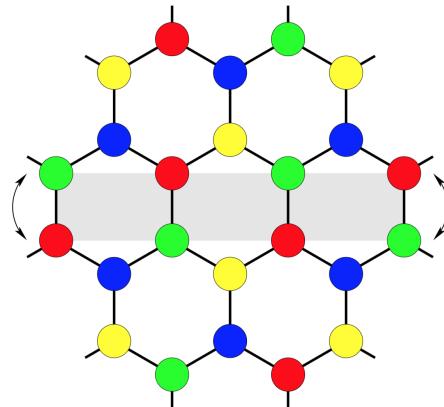


Figure 10: The most symmetric case of $SU(4)$ honeycomb lattice color filling (taken from [8])

Let us now turn to a fermionic color representation. Consider N free fermions corresponding to N different colors on a lattice, with only one fermion per site. The $SU(N)$ Heisenberg hamiltonian is then

$$\mathcal{H} = \sum_{\langle i,j \rangle} P_{i,j}^{(N)} = \sum_{\langle i,j \rangle} f_{i,\alpha}^\dagger f_{i,\beta} f_{j,\beta}^\dagger f_{j,\alpha} \quad (20)$$

, where a sum over repeated color indices is implied.

The usual fermionic anticommutation relations yield

$$\mathcal{H} = zN_f - \sum_{\langle i,j \rangle} f_{j,\beta}^\dagger f_{i,\beta} f_{i,\alpha}^\dagger f_{j,\alpha} \quad (21)$$

where z is the coordination number of the lattice, and N_f is the total number of fermions, i.e., the number of sites.

Now we apply a mean field approximation $f_{i,\beta}^\dagger f_{j,\beta} \mapsto \chi_{i,j} \equiv \langle f_{i,\beta}^\dagger f_{j,\beta} \rangle$, and keep only terms up to order one in the fluctuations.

$$\mathcal{H}_{MF} = zN_f + \sum_{\langle i,j \rangle} \left[|\chi_{i,j}|^2 - (\chi_{i,j} f_{j,\beta}^\dagger f_{i,\beta} + \chi_{i,j}^* f_{i,\beta}^\dagger f_{j,\beta}) \right] \quad (22)$$

The Gutzwiller projection method consists in projecting out multiply occupied sites from the free fermion Fermi-sea wave function. This method has been shown to accurately predict many quantities of interest, namely the exponents of correlation functions for 1D $SU(4)$ spin chains [6].

By solving the free fermionic hamiltonian self-consistently for $\chi_{i,j}$ in the MF approach, we often find that the total phase $\prod_{\langle i,j \rangle \in \text{plaq}} \chi_{i,j}$ around an elementary plaquette is non vanishing. For example, a π -flux state (corresponding to a hopping configuration giving a total phase of π at each hexagonal cell) yields the minimum energy solution on a square lattice.

Here, we shall use the Gutzwiller projection method to calculate bond energies and correlation functions on the honeycomb lattice $SU(4)$ symmetric Heisenberg model by use of Variational Monte Carlo.

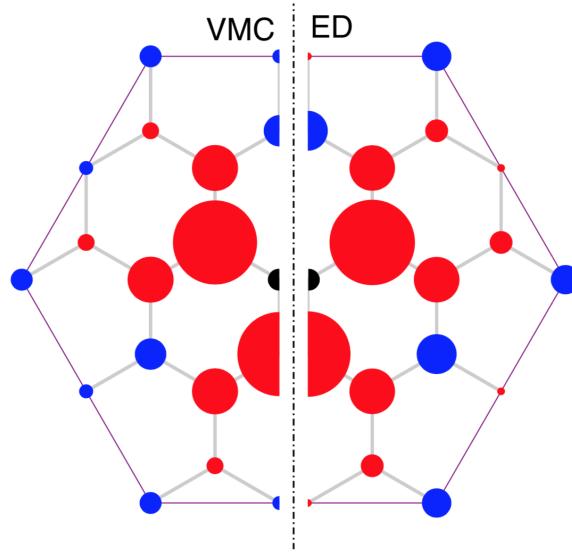


Figure 11: Comparison of color-color correlations. VMC stands for the method we shall use. ED stands for exact diagonalization carried out by Läuchli[3]. The calculation was done for a 24-site honeycomb lattice with a π -flux hopping configuration (see following sections). Blue corresponds to positive correlation, and red to negative. VMC gives a very accurate estimate of this quantity.

Part III Method

The Monte Carlo method is ubiquitous. Its central idea is to use randomness to produce accurate estimates of deterministic integrals. The term was coined by Nicolas Metropolis in 1949, first appearing in a seminal paper, in which it was described as a "statistical approach to the study of differential equations, or more generally, of integro-differential equations that occur in various branches of sciences" [12]. Although it was used as early as 1777 in an experiment known as Buffon's needle - where one obtains an estimate of the constant π by repeatedly

throwing a needle randomly onto a sheet of paper with evenly spaced lines - it was crucially developed in the Los Alamos National Laboratory during World War II where the development of the first atomic bomb was completed, the primary objective of the Manhattan Project. The method is particularly useful when one wants to sample from a probability distribution in an exponentially large state space. In fact, it can in principle be used to solve any problem allowing a probabilistic formulation.

The law of large numbers affords an approximation to integrals which can be written as an expectation of a random variable. Upon drawing enough independent samples, the sample mean gets arbitrarily close to the integral at stake. The idea is to first make an educated choice of a Markov Chain with a prescribed stationary distribution from which we ultimately desire to sample from. After a sufficiently high number of steps, a Markov Chain Monte Carlo (MCMC) algorithm generates samples from the target distribution. Imposing some conditions on this Markov Chain, namely that it should be irreducible, aperiodic and positive recurrent, the ergodic theorem guarantees that the empirical measures of the aforementioned sampler approach the target stationary distribution.

The usefulness of the Monte Carlo method in condensed matter systems is related to the fact that one frequently deals with a very large number of particles or degrees of freedom. The idea of MC is to sample a representative part of this large state space [10]. The method's power lies in the fact that it requires much less computing power than explicit methods, while still providing accurate and error-controlled results.

In the following section we provide detail on how to use a Monte Carlo method to compute physical quantities of interest using Gutzwiller projected ground state wave functions.

9 Color-color correlations

Let n_i^β be the onsite color density for color β . We aim to compute $\langle n_i^\beta n_j^\beta \rangle$, the color-color correlation functions. We choose as a variational wave function the Fermi sea ground state of the free fermionic hamiltonian

$$\mathcal{H} = - \sum_{\beta, i < j} (t_{i,j} f_{j,\beta}^\dagger f_{i,\beta} + t_{i,j}^* f_{i,\beta}^\dagger f_{j,\beta}) \quad (23)$$

Each color's energy levels are filled to $1/N$. This is equivalent to imposing that the number of fermions in the system is equal to the number of sites. Let the number of sites be N_s . Then the Fermi sea ground state wave function is

$$|\psi\rangle = \sum_{\{j\}} \prod_{\alpha=1}^N w_{\{j^\alpha\}} |j_1^\alpha j_2^\alpha \dots j_{N_s/N}^\alpha\rangle, \quad (24)$$

where the sum is carried out over the configurations $\{j\}$ and the colors α . j_l^α is the site occupied by fermion l with color α . $w_{\{j^\alpha\}}$ is a Slater determinant of one fermion wave functions of the hamiltonian in equation (23).

$$w_{\{j^\alpha\}} = \begin{vmatrix} \chi_1(j_1^\alpha) & \chi_1(j_2^\alpha) & \dots & \chi_1(j_{N_s/N}^\alpha) \\ \chi_2(j_1^\alpha) & \chi_2(j_2^\alpha) & \dots & \chi_2(j_{N_s/N}^\alpha) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_{N_s/N}(j_1^\alpha) & \chi_{N_s/N}(j_2^\alpha) & \dots & \chi_{N_s/N}(j_{N_s/N}^\alpha) \end{vmatrix} \quad (25)$$

Note that there are N (the number of colors) Slater determinants. The Gutzwiller projected states exclude multiply occupied or empty sites, including only singly occupied sites. Even taking this substantial simplifying assumption, the number of configurations is still a very large number, in general (since the number of sites is to be taken as large as possible): $\frac{N_s!}{(N_s/N)!^N}$.

A given Gutzwiller projected state $|j_1^\alpha j_2^\alpha \dots j_{N_s/N}^\alpha\rangle$ is constructed by applying creation operators to the vacuum of fermions. Here one must pay attention to the ordering of the operators. By convention, we assume that we create all the fermions of each color at a time. This is due to the fermion anti-commutation relations introducing negative signs upon exchange of these operators.

$$\bigotimes_{\alpha=1}^N |j_1^\alpha j_2^\alpha \dots j_{N_s/N}^\alpha\rangle = \prod_{\alpha=1}^N \left(f_{j_1^\alpha, \alpha}^\dagger f_{j_2^\alpha, \alpha}^\dagger \dots f_{j_{N_s/N}^\alpha, \alpha}^\dagger \right) |0\rangle \quad (26)$$

9.1 Importance sampling

The average color-color correlation function in a Gutzwiller projected state is

$$\langle \sum_\alpha n_i^\alpha n_j^\alpha \rangle = \frac{\langle \psi_G | \sum_\alpha n_i^\alpha n_j^\alpha | \psi_G \rangle}{\langle \psi_G | \psi_G \rangle} \quad (27)$$

Taking the hopping amplitudes to be the same for each color in equation (23), the operator $n_i^\alpha n_j^\alpha$ is diagonal, that is, taking the summation over repeated color indices implied, we may rewrite equation (27):

$$\langle n_i^\alpha n_j^\alpha \rangle = \frac{1}{\sum_{\{j\}} |w_{\{j\}}|^2} \sum_{\{j\}} |w_{\{j\}}|^2 \langle \{j\} | n_i^\alpha n_j^\alpha | \{j\} \rangle, \quad (28)$$

where $|\{j\}\rangle$ are fermion configurations, and $|w_{\{j\}}|^2 = \prod_{\alpha=1}^N |w_{\{j^\alpha\}}|^2$ is the product of all Slater determinants. The fermion configurations taken into account here are only the ones with strictly one fermion per site. Moreover, note that the expectation $\langle n_i^\alpha n_j^\alpha \rangle_{\{j\}}$ is simply 1 if the fermions on sites i and j have the same color, and 0 otherwise.

The first idea one might have when trying to simulate the random quantum oscillations of the system would be to simply draw a high number M of samples $\mu = 1, \dots, M$ and approximate the expectation by the estimator $E1$, such that:

$$\langle n_i^\alpha n_j^\alpha \rangle_{E1} = \frac{1}{\sum_{\mu=1}^M |w_{\{j\}_\mu}|^2} \sum_\mu |w_{\{j\}_\mu}|^2 \langle \{j\}_\mu | n_i^\alpha n_j^\alpha | \{j\}_\mu \rangle \quad (29)$$

However, one may end up sampling configurations that do not contribute significantly to the sum. Thus, we might need to increase the number of samples very much before the stochastic expectation $\langle n_i^\alpha n_j^\alpha \rangle_{E1}$ approaches $\langle n_i^\alpha n_j^\alpha \rangle$ satisfactorily, yielding an accurate estimate. In general, this estimator converges poorly. By applying importance sampling to our Monte Carlo algorithm, convergence improves drastically. This is because very often we find a difference of several orders of magnitude between the weights of successive configurations. The poor convergence is due to the fact that a greedy sampling scheme frequently picks states with a very low contribution to the quantum mechanical expectation value.

The convergence is sped up if we instead always pick a configuration with a probability proportional to its weight $P(\{j\}_\mu) \propto |w_{\{j\}_\mu}|^2$:

$$P(\{j\}_\mu) = \frac{|w_{\{j\}_\mu}|^2}{\sum_{\{j\}} |w_{\{j\}}|^2} \quad (30)$$

The estimator on equation (29) then becomes simply

$$\langle n_i^\alpha n_j^\alpha \rangle_{E2} = \frac{1}{M} \sum_\mu \langle \{j\}_\mu | n_i^\alpha n_j^\alpha | \{j\}_\mu \rangle \quad (31)$$

Now the weights are embedded in the sampling process, instead of in the estimator. This importance sampling technique is called a variance reduction technique, since the highly dispersed distribution of randomly chosen configurations is concentrated on a smaller more relevant region of state space.

The estimator $E2$ converges much faster to $\langle n_i^\alpha n_j^\alpha \rangle$ because all sampled configurations have the same contribution to the estimator, and consequently the spread in the weights of the chosen configurations is small.

It turns out that to pick configurations with probability $P(\{j\})$ we don't have to explicitly compute the probability, which would require computing the weights of all the states, yielding no improvement.

This concept is connected to that of an ensemble average in statistical physics. The ergodic theorem ensures that an ensemble average is equivalent to a time average. Therefore, the evolution in time according to the probabilities in equation (30) is analogous to the evolution of a system, according to probabilities of microscopic configurations given by $P(\{j\}) \propto e^{-\beta E_{\{j\}}}$, the Boltzmann weight. Here, β is the inverse temperature, and $E_{\{j\}}$ is the energy of the system in configuration $\{j\}$.

We mimic this concept of time evolution measurement by imposing a dynamics on the Gutzwiller subspace phase space.

10 Markov Chain. Detailed Balance. Metropolis Algorithm

Fermionic configurations are sampled according to the dynamics given by the probability distribution $P(\{j\}, t)$, and measurements are to be carried out in this system once equilibrium is reached.

Let the probability per unit time step, or rate, of going from configuration $\{j\}$ to $\{j'\}$ be $W_{jj'}$, to make the notation lighter. This corresponds to a stochastic matrix, describing a Markov chain. We may write a master equation:

$$\partial_t P(\{j\}, t) = \sum_{\{j'\}} [W_{j'j} P(\{j'\}, t) - W_{jj'} P(\{j\}, t)], \quad (32)$$

which actually constitutes a *set* of equations, describing a "global" balance between the transitions to and from $\{j\}$, by summing on all other configurations. The stationary distribution of the time-independent Markov Chain is that for which $\partial_t P(\{j\}, t) = 0$.

In fact, physical systems in equilibrium obey a stronger condition, holding for all pairs of configurations $\{j\}, \{j'\}$. The sort of "global" balance above holds in every path corresponding to a given set of transitions connecting configurations. There is no current between configurations in phase space, and processes are taken to be time-reversal symmetric, that is reversible in the thermodynamic sense. This stronger condition is called *detailed balance*:

$$W_{j'j} P(\{j'\}, t) - W_{jj'} P(\{j\}, t) = 0 \quad (33)$$

Note that detailed balance implies stationarity, but not vice-versa.

In this particular case, our phase space is that of the configurations with strictly one fermion per site. The transition matrix should satisfy detailed balance with the probability $P(\{j\}, t)$ defined in equation (30).

Upon choosing a suitable dynamics, the initial configuration is arbitrary since the system can be proven to equilibrate eventually (and sufficiently fast). Typically it is chosen randomly with uniform probability. The system then evolves, and after a given amount of time, it reaches the equilibrium distribution, i.e. it has a probability $P(\{j\})$ of realizing configuration $\{j\}$.

Up until now, we don't seem to have gained much. It seems that we would have to collect all the transition rates $W_{jj'}$ at each step to choose the next configuration. The computational cost of storing $W_{jj'}$ would take us back to square one, being impractical due to the exponentially large state space.

Let us introduce an *acceptance ratio* $A_{jj'}$, and a selection probability $g_{jj'}$, where notation is again made lighter, by making $\{j\} \mapsto j$, so that

$$W_{jj'} = g_{jj'} A_{jj'} \quad (34)$$

A new configuration j is always selected: $\sum_{j'} g_{jj'} = 1$, but the system only makes the transition to that configuration at a given time step with probability $A_{jj'}$. Detailed balance (equation (33)) may be rewritten

$$\frac{P(\{j\})}{P(\{j'\})} = \frac{g_{j'j} A_{j'j}}{g_{jj'} A_{jj'}} \quad (35)$$

The advantage of such a formulation is that, since the selection probability is uniform for all target configurations, it suffices to properly choose the acceptance ratios to satisfy detailed balance. We made a huge improvement: instead of generating all the transition rates at each step $W_{jj'}$, we simply compute the acceptance ratio of the transition to the target configuration.

In our case the selection process connects configurations where two differently colored fermions are exchanged. Thus each configuration has $\binom{N}{2} (\frac{N_s}{N})^2$ possible target configurations, from which we choose one with uniform probability at each step.

The Metropolis algorithm gives the most common choice for the acceptance ratios:

$$A_{jj'} = \begin{cases} 1 & \text{if } \frac{P(\{j'\})}{P(\{j\})} > 1 \\ \frac{P(\{j'\})}{P(\{j\})} & \text{if } \frac{P(\{j'\})}{P(\{j\})} \leq 1 \end{cases} \quad (36)$$

For our system, the ratio above is simply the ratio of the products of squared Slater determinants

$$\frac{P(\{j'\})}{P(\{j\})} = \prod_{\alpha=1}^N \frac{|w_{\{j'\alpha}\}|^2}{|w_{\{j\alpha}\}|^2} \quad (37)$$

11 Efficient determinant update

In complexity theory, the cost of computing the determinant of a $n \times n$ matrix is $\mathcal{O}(n^3)$ (for example using LU decomposition). However, the case in which we are interested is not this most general one. When exchanging two differently colored fermions, we modify only one column in two of the four Slater matrices, corresponding to the chosen colors. These columns correspond to the position of the exchanged fermions. The ratios of determinants where only a single column is modified allow a less expensive computation of order $\mathcal{O}(n^2)$, using widely known theorems, such as the matrix determinant lemma and the Sherman-Morrison formula. This optimization is important since the number of steps of a typical run can be really high, say at least of order 10^6 , and the accuracy improves with the number of measurements we make.

The determinant may be computed recursively, resorting to the Laplace expansion

$$\det \mathbf{A} = \sum_j (-1)^{k+j} a_{kj} \det \mathbf{A}_{kj}, \quad (38)$$

where \mathbf{A}_{kj} is the minor $(n - 1) \times (n - 1)$ matrix, obtained by deleting the i^{th} row and the j^{th} column of \mathbf{A} .

On the other hand, using a repeated index summation convention, the definition of the inverse matrix is

$$a_{ij}a_{jk}^{-1} = \delta_{ik} \quad (39)$$

and the elements of the inverse matrix are given by

$$a_{jk}^{-1} = \frac{1}{\det \mathbf{A}} (-1)^{k+j} \det \mathbf{A}_{kj}, \quad (40)$$

where we recognize the summand in equation (38).

Upon changing the k^{th} column of matrix \mathbf{A} , the determinant is modified, so that for the new matrix $\tilde{\mathbf{A}}$

$$\det \tilde{\mathbf{A}} = \sum_j \tilde{a}_{jk} \det \tilde{\mathbf{A}}_{jk} = \sum_j \tilde{a}_{jk} \det \mathbf{A}_{jk} = \det \mathbf{A} \sum_j \tilde{a}_{jk} a_{kj}^{-1}, \quad (41)$$

where we write the summations explicitly for clarity.

In the first step, we used the fact that the change of column k does not affect the minor matrices, and in the second step we used equation (40).

The improvement in computational cost is made at the expense of having to store the inverse Slater matrices. The ratio of determinants is given by

$$q \equiv \frac{\det \tilde{\mathbf{A}}}{\det \mathbf{A}} = \sum_j \tilde{a}_{jk} a_{kj}^{-1} \quad (42)$$

Since two Slater matrices are modified at each step, two determinant ratios must be computed at each step. The ratio between the new and old configuration probabilities is then $q_1^2 q_2^2$. The new configuration is either accepted or rejected as per equation (36), according to whether $q_1^2 q_2^2 > 1$ or $q_1^2 q_2^2 \leq 1$, respectively.

If we reject a step, the original configuration is left unchanged and a new target is selected. If we accept it, however, the Slater matrices and corresponding inverse matrices are updated and the simulation continues. The only result we need to derive now is the update rule for the inverse matrices. There are two different cases. The case of the exchanged column k and all others $i \neq k$.

$$\begin{aligned} \tilde{a}_{kj}^{-1} &= \frac{a_{kj}^{-1}}{q} \\ \tilde{a}_{ij}^{-1} &= a_{ij}^{-1} - \frac{a_{kj}^{-1}}{q} \sum_{l=1}^n a_{il}^{-1} \tilde{a}_{lk} \end{aligned} \quad (43)$$

The first result stems from the fact that we only change column k of matrix \mathbf{A} , so the \mathbf{A}_{jk} minor matrices are left unchanged. Simply plug it into the equation above to see that it works. The second is less trivial, so we present a proof, that is done by replacing the update and verifying that it gives the right inverse matrix. First note that the sum is the same for all elements of row i , so the update of a row of an inverse matrix is done in $\mathcal{O}(n)$.

We start by using the first update rule, splitting the new inverse matrix into the $i \neq k$ and the $i = k$ cases. Repeated summation convention is again used throughout.

$$\begin{aligned} \tilde{a}_{ij}^{-1} \tilde{a}_{jm} &= a_{ij}^{-1} \tilde{a}_{jm} - \frac{a_{kj}^{-1}}{q} \tilde{a}_{jm} a_{il}^{-1} \tilde{a}_{lk} \\ &= \tilde{a}_{ij}^{-1} \tilde{a}_{jm} - \delta_{km} a_{il}^{-1} \tilde{a}_{lk} \\ &= \tilde{a}_{ij}^{-1} a_{jm} + \tilde{a}_{ij}^{-1} (\tilde{a}_{jm} - a_{jm}) - \delta_{km} a_{il}^{-1} \tilde{a}_{lk} \\ &= \tilde{a}_{ij}^{-1} a_{jm} + \tilde{a}_{ij}^{-1} \delta_{km} (\tilde{a}_{jk} - a_{jk}) - \delta_{km} a_{il}^{-1} \tilde{a}_{lk} \\ &= \delta_{im} - \delta_{m(k)} \delta_{i(k)} \\ &= \delta_{im} \end{aligned} \quad (44)$$

We used $a_{il}^{-1} \tilde{a}_{lk} = \delta_{ik}$ repeatedly. This equality is due to the fact that \tilde{a}_{jm} and a_{jm} only differ if $m = k$. In the last step we just need to note that we have defined $i \neq k$ from the start. For clarity, the k index is inside parenthesis to remind the reader that it is not summed.

This part of the implementation is particularly challenging because one needs to take into account the different cases that might happen due to the difference in the updates for $i = k$ and $i \neq k$. This depends on which particular columns of the inverse matrices we are updating. We show a code snippet below to illustrate the difficulty.

```

// Update inverses

prefactor1 = invs ( fermion1 ,
span( color1*N_STATES, color1*N_STATES + N_STATES - 1) )/q1;

prefactor2 = invs ( fermion2 ,
span( color2*N_STATES, color2*N_STATES + N_STATES - 1) )/q2;

for (j = 0; j < N_STATES; j++){
    if (j == fermion1 and j == fermion2){
        invs ( j , span( color1*N_STATES,
color1*N_STATES + N_STATES - 1) )
= invs ( j , span( color1*N_STATES, color1*N_STATES + N_STATES - 1) )/q1;
        invs ( j , span( color2*N_STATES,
color2*N_STATES + N_STATES - 1) )
= invs ( j , span( color2*N_STATES, color2*N_STATES + N_STATES - 1) )/q2;
        continue;
    }
    else if (j == fermion1){
        invs ( j , span( color1*N_STATES,
color1*N_STATES + N_STATES - 1) )
= invs ( j , span( color1*N_STATES, color1*N_STATES + N_STATES - 1) )/q1;
        invs ( j , span( color2*N_STATES,
color2*N_STATES + N_STATES - 1) )
= invs ( j , span( color2*N_STATES, color2*N_STATES + N_STATES - 1) )
- as_scalar( invs( j , span( color2*N_STATES,
color2*N_STATES + N_STATES - 1) ) * column1 )*prefactor2;
        continue;
    }
    else if (j == fermion2){
        invs ( j , span( color2*N_STATES,
color2*N_STATES + N_STATES - 1) )
= invs ( j , span( color2*N_STATES,
color2*N_STATES + N_STATES - 1) )/q2;
        invs ( j , span( color1*N_STATES,
color1*N_STATES + N_STATES - 1) )
= invs ( j , span( color1*N_STATES,
color1*N_STATES + N_STATES - 1) )
- as_scalar( invs( j , span( color1*N_STATES,
color1*N_STATES + N_STATES - 1) ) * column2 )*prefactor1;
        continue;
    }
    invs ( j , span( color2*N_STATES,
color2*N_STATES + N_STATES - 1) )
= invs ( j , span( color2*N_STATES,
color2*N_STATES + N_STATES - 1) )
- as_scalar( invs( j , span( color2*N_STATES,
color2*N_STATES + N_STATES - 1) ) * column1 )*prefactor2;
    invs ( j , span( color1*N_STATES,
color1*N_STATES + N_STATES - 1) )
= invs ( j , span( color1*N_STATES,
color1*N_STATES + N_STATES - 1) )
- as_scalar( invs( j , span( color1*N_STATES,
color1*N_STATES + N_STATES - 1) ) * column2 )*prefactor1;
}

```

12 Measurements

We have designed a process generating configurations with our desired probability distribution $P(\{j\})$ once equilibrium is reached. After the path towards equilibrium is completed, the so called warm-up steps, we may start measuring. Starting measurements too early will yield distorted estimates because it would mean that we were not sampling from the right equilibrium distribution, but some other undefined one instead.

The equilibrium time can be estimated by looking at the weights of configurations that are picked as a function of time, as the Monte Carlo algorithm runs. We expect the weights of the chosen configurations to converge (starting from the very small weights of the first few configurations), and then stochastically oscillate around a given order of magnitude, dependent upon the total number of sites. This is afforded by our importance sampling scheme.

The first benchmark of the project was then to ensure that the implemented Monte Carlo was yielding this type of convergence. This was done by aiming to reproduce the work in the thesis by Dr Miklós Lajkó [8].

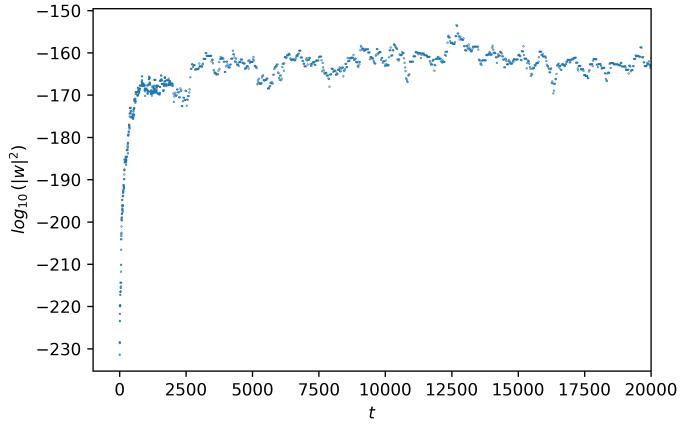


Figure 12: A 20000-step run of our Monte Carlo algorithm for a 200-site system, showing convergence at around $t = 3000$.

After making sure that the Monte Carlo algorithm is picking the configurations according to the right probability distribution by using importance sampling, we may turn to measurements.

The sampling distance, that is the number of steps that we wait between measurements, should be such that the samples are independent. We hope that as time goes by between measurements, they become stochastically uncorrelated. To quantify the sampling distance we use the auto-correlation function of some measured quantity Q :

$$\chi_Q(t) = \int dt' (Q(t+t') - \langle Q \rangle)(Q(t') - \langle Q \rangle), \quad (45)$$

where $\langle Q \rangle$ is the average of the quantity, and the integral is of course discretized in practice. In our case, we choose this quantity to be the color average of the correlations between site 0 and a nearest neighbor δ : $\langle n_0 n_\delta \rangle$.

The auto-correlation time is estimated as 5τ , where τ is the characteristic time of the exponential decay of the auto-correlation function. Since our measurements are performed in finite time, we introduce a cut-off time, such that there are enough points to fit an exponential reliably. Measuring the decay of χ_Q accurately is important: underestimating the auto-correlation time can lead to an underestimate of the error in the measurements.

13 Bond energies

The average bond energy $\langle P_{kl} \rangle$ is not as easy to measure since P_{kl} is an off-diagonal quantity in the $\{j\}$ basis, although it can still be calculated using the Gutzwiller projected variational state. However, now we must pay attention to the sign changes resulting from the exchange of fermionic operators. Recalling that we simplified the notation for a configuration $\{j\} \mapsto j$, the expectation of the exchange operator may be written, recalling equation (24):

$$\langle P_{kl} \rangle = \frac{1}{\sum_j |w_j|^2} \sum_{jj'} \bar{w}_{j'} w_j \langle j' | P_{kl} | j \rangle \quad (46)$$

, where $\bar{w_j}$ denotes the complex conjugate. The exchange operator in the fermionic representation is written, according to equation (20):

$$P_{kl} = \sum_{\alpha\beta} f_\alpha^\dagger(k) f_\beta(k) f_\beta^\dagger(l) f_\alpha(l) = - \sum_{\alpha\beta} f_\alpha^\dagger(k) f_\beta^\dagger(l) f_\beta(k) f_\alpha(l) \quad (47)$$

, where in the last equality the minus sign comes from switching the order of the two middle operators.

Now note that $\langle j' | P_{kl} | j \rangle$ is only non zero for the j' configuration where the colors of the fermions on sites k and l are exchanged. Let us denote this particular configuration by j^* . Omitting the implicit repeated index sum,

$$\left(\underbrace{\langle j^* | f_\alpha^\dagger(k) f_\beta^\dagger(l) }_{\langle j | f_\alpha^\dagger(l) f_\beta^\dagger(k)} \right) \left(f_\beta(k) f_\alpha(l) | j \rangle \right) = 1 \quad (48)$$

since the colors (assumed different) of the fermions on k and l are exchanged, and thus we have the same state inside both parenthesis. Thus, if the colors are different, the rearrangement of the fermion creation operators above gives $\langle j^* | P_{kl} | j \rangle = -1$. If the color of the fermion on sites k and l is the same, say γ , the configurations are equal $j^* = j$, so that equation (47) reduces to (no summation implied):

$$P_{kl} = f_\gamma^\dagger(k) f_\gamma(k) f_\gamma^\dagger(l) f_\gamma(l) \quad (49)$$

, and, again with no summation implied,

$$\left(\langle j | f_\gamma^\dagger(k) f_\gamma(k) \right) \left(f_\gamma^\dagger(l) f_\gamma(l) | j \rangle \right) = 1 \quad (50)$$

Collecting both results, equation (46) now becomes

$$\begin{aligned} \langle P_{kl} \rangle &= \mp \frac{1}{\sum_j |w_j|^2} \sum_j \bar{w_j} w_j \\ &= \mp \sum_j P(j) \frac{\bar{w_j}}{w_j} \end{aligned} \quad (51)$$

, where we simply multiplied and divided by $\bar{w_j}$ and the minus sign is for differently colored l and k sites, and the plus sign for l and k with the same color.

Since the Markov Chain that we designed leads us to distribution $P(j)$ after the warm-up steps, it suffices to measure $q(j) \equiv \frac{w_{j^*}}{w_j}$ at each step and take the average over the number of measurements m . Here we dropped the complex conjugation because, since equation (46) is symmetric under the exchange of the j , j' indices, we could have done exactly the same reasoning by switching the indices in the beginning, and at the previous step of the calculation above multiplying and dividing by w_j instead of $\bar{w_j}$

$$\langle P_{kl} \rangle_{MC} = \frac{1}{m} \sum_{\mu=1}^m q(j_\mu) s_{k,l}(j_\mu) \quad (52)$$

, where $s_{k,l}(j)$ is 1 if sites k and l have fermions with the same color, and -1 otherwise.

The measurement of the bond energies is straightforward. At a given measurement step, if the two sites - that we read from the provided neighbors list - have fermions with the same color, then, by equation (52), we simply measure a unit energy. However, if they have different colors we measure minus the ratio $q(j) = \frac{w_{j'}}{w_j}$. This is in essence the same process that goes into computing the acceptance ratio at each step. It is as if we propose a step in which the fermions on the two sites are exchanged. The measured bond energy for differently colored fermions is simply the negative ratio between the determinants of the proposed configuration and the current one.

We will take first, second and third nearest neighbor bonds into account. To do this, we read all the elements in our list of nearest and second and third next nearest neighbors, and "propose an exchange" for all bonds every time we measure, to compute the ratio between weights. Here we provide a snippet of our C++ code to clear up the explanation.

```
for (k = 0 ; k < 6*N_SITES ; k++){
```

```
// The measurement is as if you propose a change in the columns
```

```

site1_measure = bonds(k, 0);

site2_measure = bonds(k, 1);

fermion1_measure = fermion_color( site1_measure, 0 );

color1_measure = fermion_color( site1_measure, 1 );

fermion2_measure = fermion_color( site2_measure, 0 );

color2_measure = fermion_color( site2_measure, 1 );

if (color1_measure == color2_measure){

bond_energies(k) = bond_energies(k) + one_zero;

}

else {

    column1_measure = slaters_new( span( 0, N_STATES -1 ),
span( N_STATES*color1_measure + fermion1_measure ) );
    column2_measure = slaters_new( span( 0, N_STATES -1 ),
span( N_STATES*color2_measure + fermion2_measure ) );

    q1_measure = dot( invs( fermion1_measure, span( color1_measure*N_STATES,
(color1_measure + 1)*N_STATES -1 ) ), column2_measure ) ;

    q2_measure = dot( invs( fermion2_measure, span( color2_measure*N_STATES,
(color2_measure + 1)*N_STATES -1 ) ), column1_measure ) ;

    bond_energies(k) = bond_energies(k) - q1_measure*q2_measure;
}
}

```

All the variable names are self-explanatory, except for `one_zero`, which corresponds to the complex constant $1 + 0i$, which is used solely for data type convenience. `fermion_color` lists the color and the index on the configuration array of the fermion with a given color on each site. The `fermion_color` list is updated as fermions are exchanged at each Monte Carlo step:

```

// Make the switch; update configuration

config_t[N_STATES*color1 + fermion1] = site2;

config_t[N_STATES*color2 + fermion2] = site1;

fermion_color( site1 , 0 ) = fermion2;

fermion_color( site1 , 1 ) = color2;

fermion_color( site2 , 0 ) = fermion1;

fermion_color( site2 , 1 ) = color1;

```

14 MC algorithm outline and code flowchart

14.1 Outline

The main parts of the Monte Carlo algorithm are

1. Initialization;
2. MC time evolution;
3. Measurements.

Phase 1 corresponds to solving the eigenvalue problem of determining the one-fermion wave functions of the free fermionic hamiltonian; preparing an initial fermion configuration with one fermion per site, and calculating the initial Slater matrices and corresponding inverses.

Phase 2 comprises the iteration of the elementary MC step:

- Select 2 differently colored fermions in the current configuration, which we proposed to exchange;
- Calculate the ratio between the squared determinants of the target and current configurations by equation (42);
- This is where we use the Metropolis algorithm. We decide whether a target configuration is accepted or rejected using equation (36);
- If we accept the target configuration, the fermion configuration, and the Slater and corresponding inverse matrices are updated as per equation (43). The update on the fermion configuration and the Slater matrices corresponds to simply switching the colors of the exchanged fermions and the columns of the Slater matrices corresponding to the switch.
- If we reject a step, we keep the current configuration and the corresponding Slater and inverse matrices.

Phase 3 follows the graphically estimated warm-up time W . An initial run is used to estimate the auto-correlation time $A = 5\tau$. After warm-up, a measurement is made every A steps, yielding a total of $\frac{S-W}{A}$ measurements, where S is the total number of steps. For $n_i^\alpha n_j^\alpha$ it suffices to increment a counter by one whenever the fermions on sites i and j have the same color to calculate the average at the end. The measurement of the bond energies requires the use of the scheme described in section 13. These are also stored in a vector in which each entry corresponds to a given bond, so as to allow one to later map these average bond energies on the lattice. In the end, an average is performed over all nearest neighbor and second and third next-nearest neighbor bonds.

14.2 How to attack the problem and technicalities

The first step taken into solving the problem was to think of the simple particular case of an eight-site honeycomb lattice. One can draw the periodic boundary conditions on the lattice (see next figure), and explicitly write down the hopping matrix \mathbf{t} , for both a uniform hopping amplitude and a π -flux configuration.

$$\mathbf{t}_{uniform} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad \mathbf{t}_{\pi-flux} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & 0 & 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (53)$$

Writing an algorithm that formulates the problem for this small simple system turned out to be a good approach to tackle the implementation for the general case.

The strategy that was followed was to first code in Python, using the `numpy`, `scipy`, and `matplotlib` libraries, and then translate the program into C++, where we used a linear algebra library integrated with LAPACK called `armadillo`. This is a very powerful and fast library that provides a good compromise between ease of use and speed. The file `main.cpp` in the project folder contains the final program that was written from scratch for this project. Additionally, we include a `main_8.cpp` file, which consists of an optimization of the main code that ran on the MacBook Pro with an 8-core processor kindly provided by the cosupervisor Dr Miklós Lajkó.

The Monte Carlo algorithm for this small system was then first coded in Python. In fact this strategy allowed a faster solution of the problem to gain some intuition. Both a brute force determinant calculation at each step

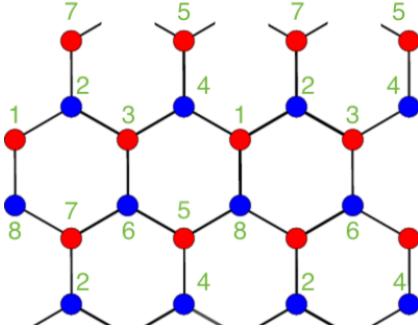


Figure 13: Periodic boundary conditions on the honeycomb lattice.

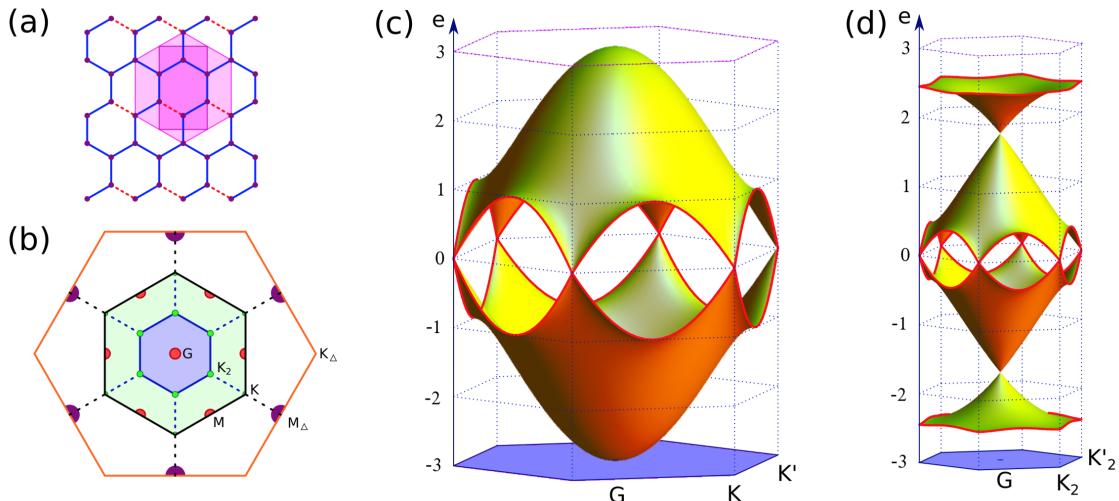


Figure 14: (a) Minimal rectangular 4-site and 8-site hexagonal unit cell of the π -flux hopping configuration. The dashed red lines represent the negative hoppings. (b) First Brillouin zone of the honeycomb lattice (middle black hexagon); Reduced Brillouin zone of the π -flux hopping configuration (inner blue hexagon); Brillouin zone of the triangular lattice (obtained by filling the "centers" of the honeycomb lattice) - outer orange hexagon. (c-d) Comparison of the free fermion band structures of the 0-flux and π -flux cases. In the π -flux case, a Dirac cone appears at quarter filling, corresponding to the $SU(4)$ Fermi level (taken from [8]).

and the efficient determinant update of section 11 were used at first. This was done to make sure that the update rules were well implemented and the correct acceptance ratio was being computed at each step. Then some basic correlation function measurements were implemented and checked against baseline results from [8]. This exact program was then translated into C++, and adapted for use with the `armadillo` library, providing exactly the same results as the initial Python code. The problem was then generalized for an arbitrary number of sites, and we finished the implementation of measurements.

After having a working code where we implemented the Gutzwiller Projection VMC, we moved on to optimize it, finish the implementation and perform several measurements to conclude the final goal of the project, that of exploring a new variational state, and how it might probe the nature of the ground state, and of the topological insulator transition.

14.3 Planning the code

As we have stated in the outline, the code consists of three blocks, whose functionality we summarize below.

In addition to what is sketched in the flowchart, the measurement of the auto-correlation function is only done once for each number of sites in the measurements block of the algorithm. Note that this is done in a similar manner to the storing of the weights that are used to make the convergence plots. We make an initial shorter run of the program, where, after warm-up, we measure nearest neighbor correlations $\langle n_0 n_\delta \rangle$ at every step. The resulting data is used to make a plot of the exponentially decaying auto-correlation function $\chi_{\langle n_0 n_\delta \rangle}(t)$. We fit an exponential to this data to obtain the auto-correlation time between measurements.

We used a pre-existing `Mathematica` notebook that generates the hamiltonian matrices and exports them into `.txt` format. The C++ code reads these matrices to start the algorithm. After obtaining the exported

Flowchart

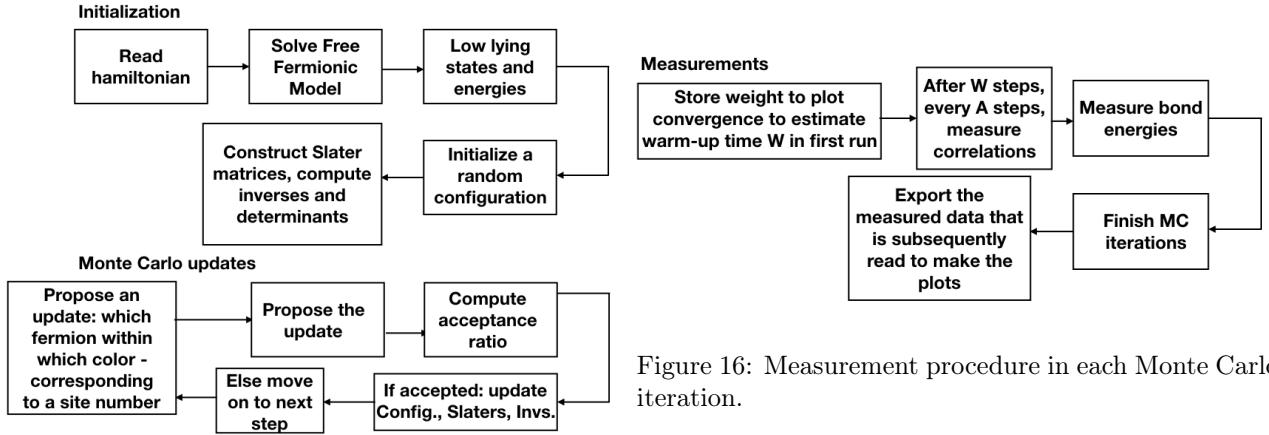


Figure 15: The layout of the code, sketching the process carried out in tackling the problem.

outputs of the main program `main.cpp`, we read the files into Python and post-process them to make the plots using the `matplotlib` and `scipy` libraries.

From our point of view, the `Mathematica` notebook essentially serves as a black box creating the input files to start the program. Two families of finite-size clusters were considered, with the full D_6^2 symmetry of the honeycomb lattice[3]:

- Clusters with $N = 2(2n)^2$ sites defined by the lattice vectors $\mathbf{g}_1 = (\sqrt{3}, 0)n$ and $\mathbf{g}_2 = (\sqrt{3}/2, 3/2)n$, where n is an integer.
- Clusters with $N = 6(2n)^2$ sites defined by $\mathbf{g}_1 = (3\sqrt{3}/2, 3/2)n$ and $\mathbf{g}_2 = (0, 3)n$, where n is an integer.

The bond length was always set to unity. We used the following numbers of sites: 24, 72, 96, 200, corresponding to elements of these two families.

Finally, for the 72-site case, we considered additional second nearest neighbor hopping t_2 , and introduced a parameter ϕ_t , specifying a phase $e^{i\phi_t}$ added to the matrix elements within each honeycomb cell, as shall be specified in part V. The idea is that by varying these parameters, we may study the phase diagram of the model.

The `read.md` file included in the folder of the project explains how to use the code to reproduce the results that will be presented in the following sections.

Part IV

Benchmarks

15 Naïve calculation of determinant ratios vs. Efficient update

The validation of the code was first done by comparing the acceptance ratios computed by the $\mathcal{O}(n^3)$ determinant calculation method, and by the update rule that has the lower cost $\mathcal{O}(n^2)$. We present an example of the output of the first program that was implemented in Python for the 8-site system shown in the previous section. A small anisotropy was included in the π -flux configuration to prevent ill-conditioning of the matrices in the beginning of the algorithm. In practice this condition is always ensured if our algorithm converges because it wouldn't converge if it computed the wrong inverse matrices in the beginning.

²This is the group of symmetry of a six-sided polygon.

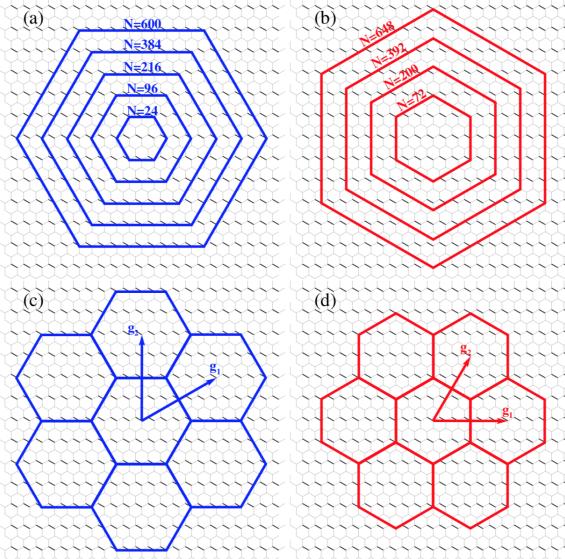


Figure 17: The two families of finite-size clusters used in the simulations (taken from [8]).

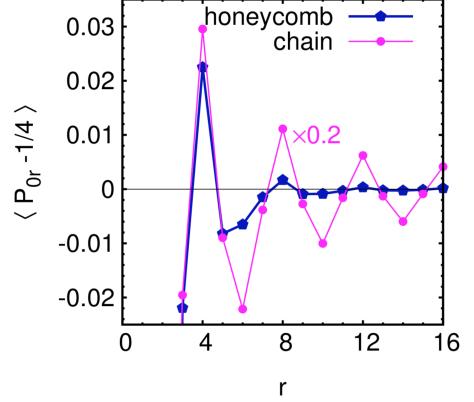


Figure 18: The quantity on the y -axis is proportional to $\langle n_0 n_r \rangle - \frac{1}{4}$, which we will measure. This plot compares the decay of correlations for the 1D Heisenberg spin chain and the honeycomb configurations to be used in this project. We were able to reproduce the result in this plot as we shall show in the following sections (taken from [8]).

$$t_{\pi-flux} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & -0.98 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & -1 & 0 & 1 & 0 & 1 & 0 \\ 0 & -0.98 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (54)$$

First we show an example of the output of the initialization part of the algorithm

SU(4) Honeycomb Hubbard Simple 8-site example

Solve the free one fermion problem: $H = - \sum_{\{ij\}} t_{ij} |i\rangle \langle j|$
Eigenvalues/Energies

$[-2.4157 \ -2.4057 \ -0.4157 \ -0.4057 \ 0.4057 \ 0.4157 \ 2.4057 \ 2.4157]$

Eigenstates/States – Columns

```
[[ 0.4619 -0.1914  0.4605  0.1947 -0.1947  0.4605  0.1914 -0.4619]
 [-0.1924  0.4596 -0.1938 -0.4628 -0.4628  0.1938  0.4596 -0.1924]
 [ 0.1914 -0.4638 -0.1901 -0.4606  0.4606 -0.1901  0.4638 -0.1914]
 [-0.4615  0.1924  0.4629  0.1891  0.1891 -0.4629  0.1924 -0.4615]
 [ 0.4615  0.1924 -0.4629  0.1891 -0.1891 -0.4629 -0.1924 -0.4615]
 [-0.1914 -0.4638  0.1901 -0.4606 -0.4606 -0.1901 -0.4638 -0.1914]
 [ 0.1924  0.4596  0.1938 -0.4628  0.4628  0.1938 -0.4596 -0.1924]
 [-0.4619 -0.1914 -0.4605  0.1947  0.1947  0.4605 -0.1914 -0.4619]]
```

Find low-lying states and energies $\phi_1, \dots, \phi_{N_S/N}; E_1, \dots, E_{N_S/N}$

Low-lying energies: $[-2.4157 \ -2.4057]$

Low-lying states

```
[[ 0.4619 -0.1924  0.1914 -0.4615  0.4615 -0.1914  0.1924 -0.4619]
 [-0.1914  0.4596 -0.4638  0.1924  0.1924 -0.4638  0.4596 -0.1914]]
```

Initialize in a random configuration (e.g. R in 5,0; G in 3,7, B in 4, 6, Y in 1,2)

Initial configuration [array([6, 2]), array([1, 7]), array([3, 0]), array([5, 4])]

Then we show the output corresponding to an elementary Monte Carlo step. Note that the acceptance ratios calculated from both methods are the same, thus validating the code.

Compute initial Slater matrices, determinants, and inverses

Slater's

```
[array([[ 0.19240258,  0.19144099],
       [ 0.459587 , -0.46380582]], array([[ -0.19240258, -0.46189781],
       [ 0.459587 , -0.1914372 ]]), array([[ -0.46149974,  0.46189781],
       [ 0.19239478, -0.1914372 ]]), array([[ -0.19144099,  0.46149974],
       [-0.46380582,  0.19239478]]])]
```

Determinants

```
[-0.17722123  0.24911524 -0.00051851  0.17721402]
```

Inverses

```
[[ 2.61710084e+00  1.08023734e+00 -7.68468454e-01  1.85415317e+00
   3.69207915e+02  8.90821244e+02  1.08566342e+00 -2.60419430e+00]
 [ 2.59329542e+00 -1.08566332e+00 -1.84487710e+00 -7.72343703e-01
   3.71054712e+02  8.90053532e+02  2.61720727e+00 -1.08028127e+00]]
```

Product of squared determinants 1.6456553412e-11

Monte Carlo

Step: 1

Colors: 1 2

Fermions: 0 1

Configuration change: 1 0

Switch the following columns [-0.19240258 0.459587]
[0.46189781 -0.1914372]

Acceptance ratio by smart update rule: 57461.4182463

Current slater matrices

```
[[ 0.19240258  0.19144099  0.46189781 -0.46189781 -0.46149974 -0.19240258
```

```

-0.19144099  0.46149974]
[ 0.459587   -0.46380582 -0.1914372  -0.1914372   0.19239478  0.459587
-0.46380582  0.19239478]]

```

Previous slater matrices

```

[[ 0.19240258  0.19144099 -0.19240258 -0.46189781 -0.46149974  0.46189781
-0.19144099  0.46149974]
[ 0.459587   -0.46380582  0.459587   -0.1914372   0.19239478 -0.1914372
-0.46380582  0.19239478]]

```

Determinants: [-0.17722123 -0.17684885 -0.17508203 0.17721402]

New product of squared determinants 9.45616898498e-07

Acceptance Ratio by ratio of squared determinants 57461.4182463

Current slater matrices

```

[[ 0.19240258  0.19144099  0.46189781 -0.46189781 -0.46149974 -0.19240258
-0.19144099  0.46149974]
[ 0.459587   -0.46380582 -0.1914372  -0.1914372   0.19239478  0.459587
-0.46380582  0.19239478]]

```

Current configuration

```
[6 2 0 7 3 1 5 4]
```

New inverses

```

[[ 2.61710084  1.08023734  1.08249053 -2.61182256 -2.62498101 -1.09892825
 1.08566342 -2.6041943 ]
[ 2.59329542 -1.08566332 -1.08249053 -2.61182256  1.09888367  2.63590587
 2.61720727 -1.08028127]]

```

16 Warm-up time

After the code was translated into C++ and generalized for an arbitrary number of sites, we tried to reproduce the first figure of section 12, and compared the time evolution of the weights computed through determinant computation and update rule.

17 Auto-correlation time

The auto-correlation time was measured by taking the fermion at site 0 and one of its nearest neighbors. The list of nearest (and second and third next nearest) neighbors is also a file provided by the **Mathematica** notebook. The idea is to save a list where at each step we store a 0 if the sites 0 and one of its nearest neighbors (say site 37) have differently colored fermions, and 1 if the sites have fermions with the same color.

The auto-correlation function was then computed from equation (45) in post-processing. We plotted this function for a time frame with a sufficient amount of points to obtain an accurate fit.

18 Correlation function measurements

The next step was to try to reproduce the previous measurements of the decay of correlations in the honeycomb lattice as a function of distance (and not site number!). This requires the use of a table that contains the correspondences between sites and distances on the honeycomb lattice, taking boundary conditions into account.

The idea is similar to that of the last section. We fix the fermion at site 0 and its color, and we check the colors of all other fermions. We initialize an array of zeros with the size of the lattice and whenever a site has a fermion with the same color as that on site 0, we increment 1 on the corresponding position of the counter

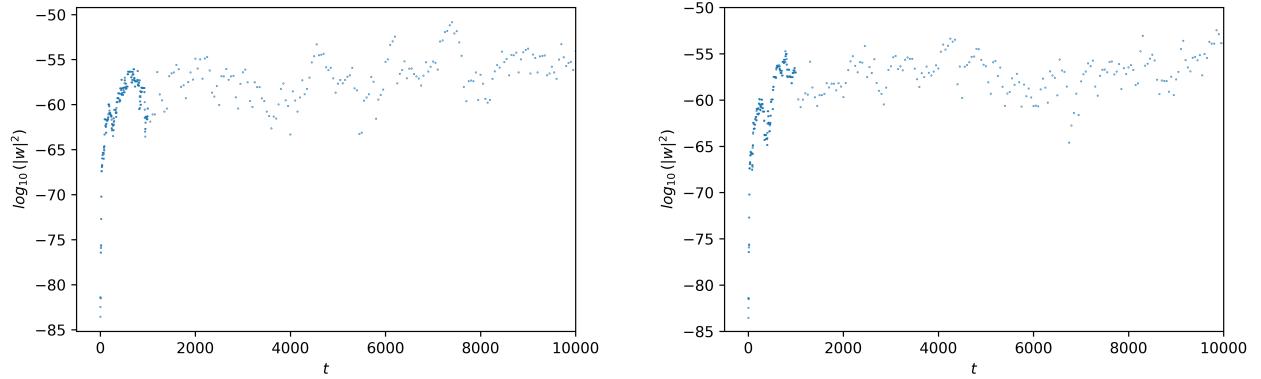


Figure 19: On the left, the convergence of the weights computed by explicit determinant calculation. On the right, a different run of the code showing a similar convergence behavior for the weights computed by the efficient update rule. For both runs we took more points in the beginning, when the function varies more, and then less points after the time estimated for convergence.

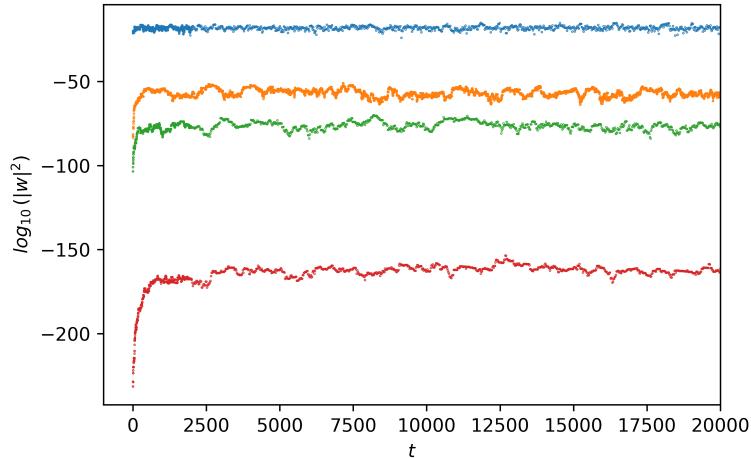


Figure 20: The warm-up process for varying system size for the cases $N_S \in \{24, 72, 96, 200\}$. The number of sites increases from top to bottom. As the system size increases so does the warm-up time. In fact, the equilibration is almost immediate for the 24-site system. Note also that the larger the system, the smaller the initial weight, which shows how crucial importance sampling is. Here we took 20000 steps of the Monte Carlo.

array. The process of making the correspondence between site distance on the lattice and physical distance is done on post-processing. To obtain the average, we simply start by finding the sites corresponding to the same distances on the lattice. Say sites 3, 64, and 28 could all correspond to the same distance from 0 on the honeycomb lattice. We then average over these. In the end, we also average the total correlation array by dividing the whole array by the total number of measurements.

The obtained result reproduces the one obtained by Lajkó [8].

Part V

Discussion

19 Introducing a novel variational state

The variational state that we now consider is a solution of a different free fermionic hamiltonian (with respect to that of equation (22)). A next nearest neighbor hopping term is added to the original mean field hamiltonian. Note that some of the matrix elements corresponding to these next nearest neighbor hoppings are complex, since we introduce a phase $e^{i\phi_t}$ when cycling through next nearest neighbors, as specified in figure 25.

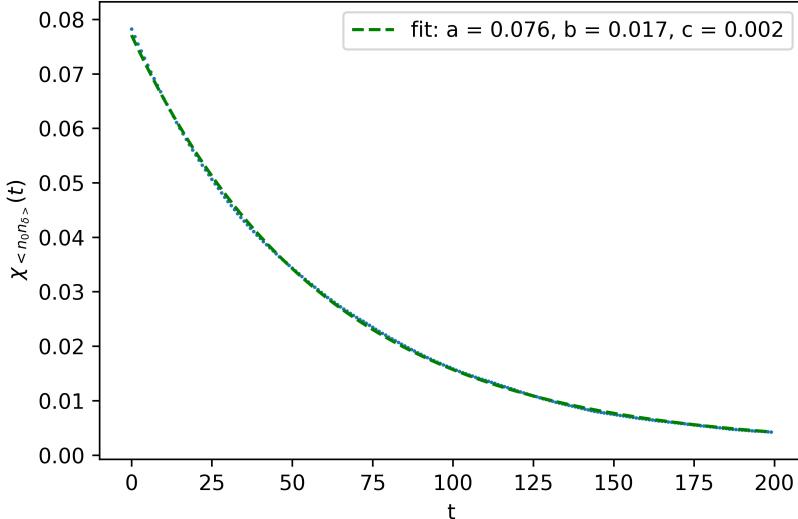


Figure 21: We fit an exponentially decaying function to the auto-correlation post-processed data: $f(t) = ae^{-bt} + c$. We obtain $b = 0.017$, corresponding to an auto-correlation time $A = 5\tau \approx 300$.

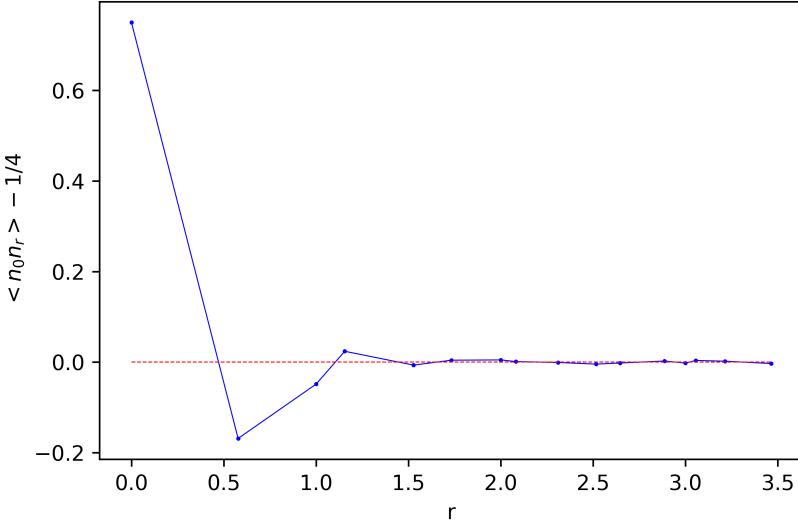


Figure 22: For a number of measurements of 10000, the correlations are seen to show the expected behavior, related to the quantum spin liquid character of the ground state. This was measured for the nearest neighbor hopping π -flux configuration, for the 72-site system.

Thus, we now consider the variational hamiltonian

$$\mathcal{H}_{var} = - \sum_{\langle i,j \rangle} \left[(t_{1,i,j} f_{j,\beta}^\dagger f_{i,\beta} + t_{1,i,j}^* f_{i,\beta}^\dagger f_{j,\beta}) \right] - \sum_{\langle\langle i,j \rangle\rangle} \left[(t_{2,i,j} f_{j,\beta}^\dagger f_{i,\beta} + t_{2,i,j}^* f_{i,\beta}^\dagger f_{j,\beta}) \right], \quad (55)$$

where $\sum_{\langle\langle i,j \rangle\rangle}$ represents a sum over next nearest neighbors.

We are interested in studying what variational parameters (t_2, ϕ_t) minimize the average bond energy of a generalized Heisenberg hamiltonian that couples up to third nearest neighbors. Denoting the pairs of second and third nearest neighbors $\langle\langle i,j \rangle\rangle \equiv \langle i,j \rangle_2$ and $\langle\langle\langle i,j \rangle\rangle \equiv \langle i,j \rangle_3$.

$$\mathcal{H} = J_1 \sum_{\langle i,j \rangle} P_{ij} + J_2 \sum_{\langle i,j \rangle_2} P_{ij} + J_3 \sum_{\langle i,j \rangle_3} P_{ij}, \quad (56)$$

which takes up to third nearest neighbor exchange into account.

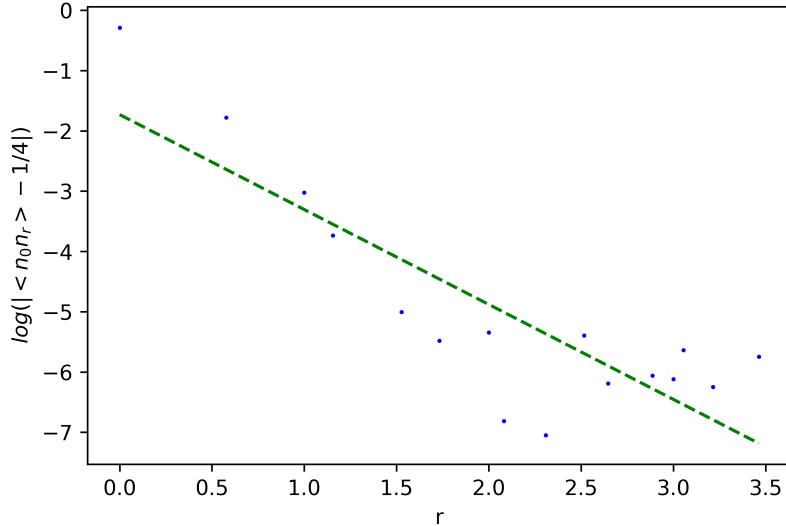


Figure 23: The same plot in log-scale compares the decay to what would be an exponential decay. The decay is known to be polynomial, but to see it we would need to run the simulations for larger systems.

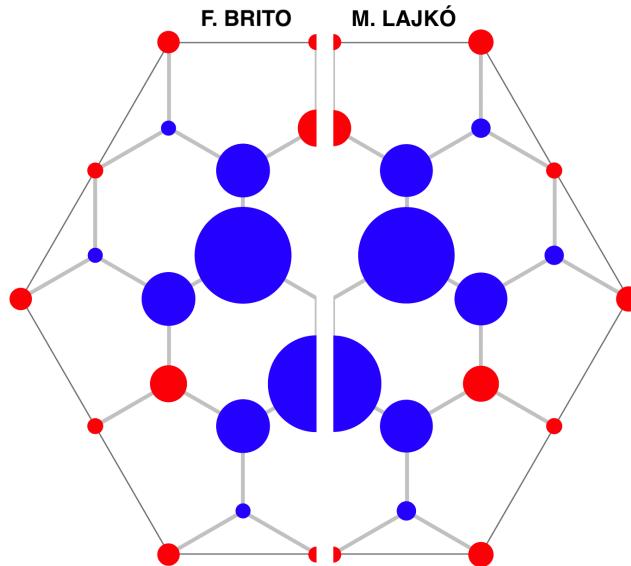


Figure 24: Comparison of our correlation measurements for the 24-site system with those of Lajkó. The two results are indistinguishable, thus validating our code. The comparison of the two results with exact diagonalization is shown in figure 11 of section 8.

In particular, we are interested in how the minimizing variational parameters (t_2, ϕ_t) change when the coupling constants J_2, J_3 are varied, that is in studying the phase diagram of the model using the novel variational state.

On the honeycomb lattice, the total number of first and third nearest neighbor pairs is $3N_S/2$ and the total number of next nearest neighbor pairs is $3N_S$, where N_S is the number of lattice sites. We measure the bond energies for these $6N_S$ pairs. Although in general the measured energy for a given bond at one step may be complex, in the end these contributions should cancel out. If one would consider the exact expectation $\langle P_{kl} \rangle$, the imaginary part would be exactly zero, since it corresponds to an energy. However, due to finite sampling, the real check to do is that the measured imaginary part is much smaller than the real part, vanishing in the many measurement limit. We stored the imaginary parts of the measured bond energies and in the end we checked that they were only of the order of 1% of the real parts, thus negligibly smaller.

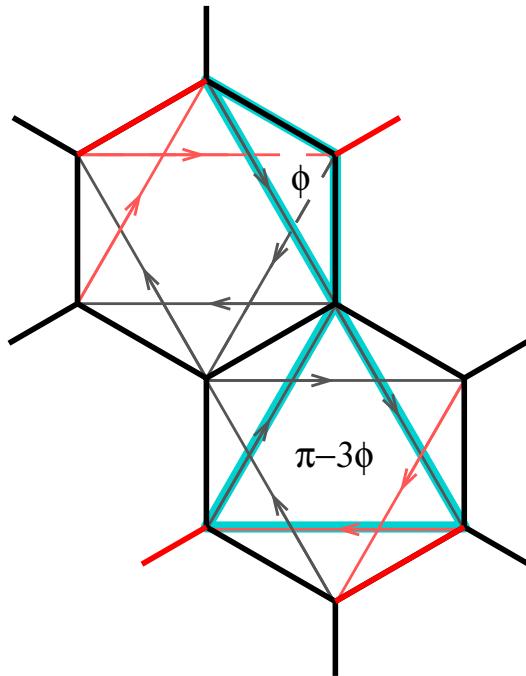


Figure 25: A phase $e^{i\phi_t}$ is introduced on the hopping matrix elements corresponding to next nearest neighbors so that a phase $\pi - 3\phi$ is gained in circling through the next nearest neighbors. The hoppings in red have a sign opposite to that of the ones in black. This condition ensures that hexagons have a total flux of π .

20 Correlation measurements

We repeated the measurements of color-color correlations for the 96-site system considering the variational state with parameters $t_2 = 0.3$ and $\phi_t = \pi/6$. We found a similar behavior to that of the nearest neighbor variational state, characteristic of the spin-liquid ground state.

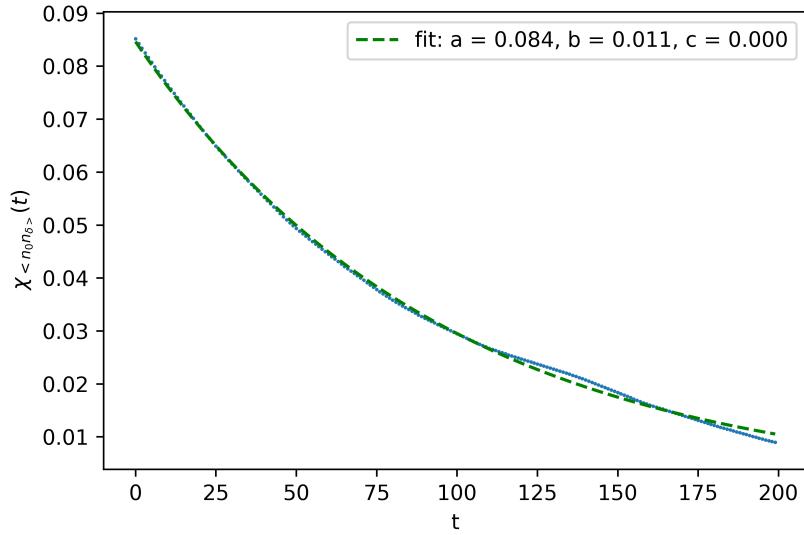


Figure 26: In a similar manner to what was done in the previous section, we obtain an estimate of 462 steps for the auto-correlation time of the 96-site system. In practice we always used $A = 700$, well in excess, to ensure no compromise in accuracy, since the increase in computational cost is marginal.

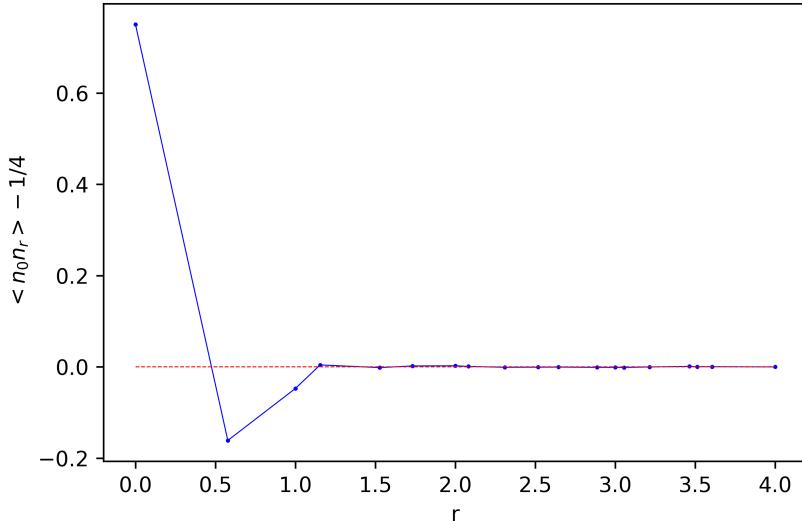


Figure 27: For a number of measurements of 100000 for the 96-site system, the correlation function shows again the typical spin liquid behavior. The parameters that were considered were $t_2 = 0.3$ and $\phi_t = \pi/6$. The hopping configuration is still a π -flux, with the inner triangle of the honeycomb having a flux of $\pi - 3\phi_t = \pi/2$.

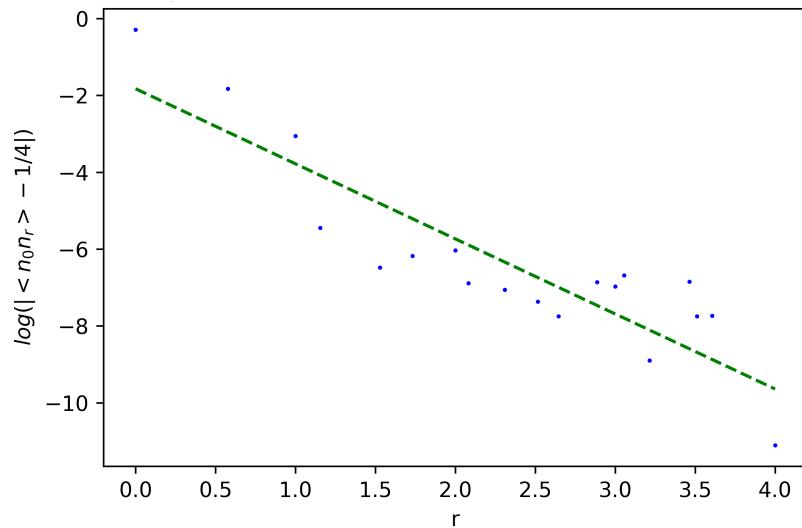


Figure 28: The same plot in logarithmic scale, compared to what would be an exponential decay, again showing that we would need larger systems (computationally too expensive for this project) to properly capture the (polynomial) decay of the correlations.

21 Phase Diagram

We made a final run of the program `main_8.cpp` optimized for the MacBook Pro with an 8-core processor. This run spanned 25 phase angles between 0 and π , and 8 hopping amplitudes t_2 between 0 and 0.7, both evenly spaced. We took 100 000 measurements (the run took $\mathcal{O}(\text{day})$), for each hopping configuration (in total $8 \times 25 = 200$ configurations) so as to obtain the maximum possible accuracy within reach of our computational power.

The first check is that for $t_2 = 0$ the result cannot depend on ϕ_t , so the oscillation in the measurements in figure 29 are solely due to the random nature of the Monte Carlo. For N measurements, the standard deviation goes like $N^{-1/2}$, so that for $N = 100000$ measurements it is ~ 0.003 . We find it to be around 0.001, for example for the $J_2/J_1 = J_3/J_1 = 1$ case (see section 19), giving an estimate of the uncertainty of our total bond energy results.

To illustrate how we constructed the phase diagrams we fix for now $J_2/J_1 = J_3/J_1 = 1$. For example, in this case we found that the variational parameters minimizing the total bond energy were $t_2 = 0.3$ and $\phi_t = 17\pi/24$.

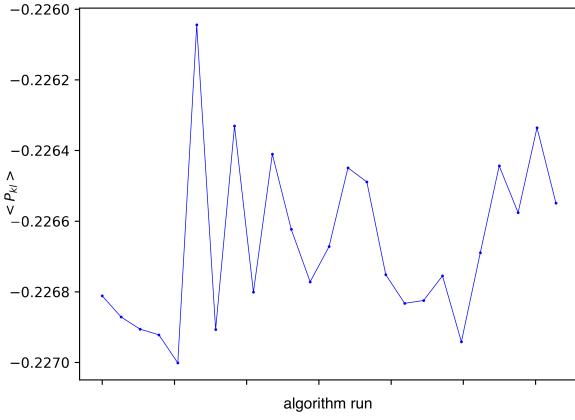


Figure 29: Fluctuations of the measured total bond energy for different runs of the algorithm for $t_2 = 0$. The uncertainty is estimated as 0.001, consistent with the central limit theorem estimate.

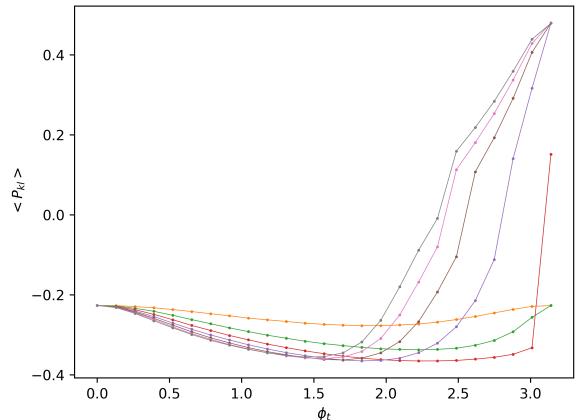


Figure 30: Average bond energy as a function of ϕ_t for different values of t_2 , for $J_{1,2,3} = 1$. Our aim is to find the parameters (t_2, ϕ_t) corresponding to the minimum average bond energy out of all these cases.

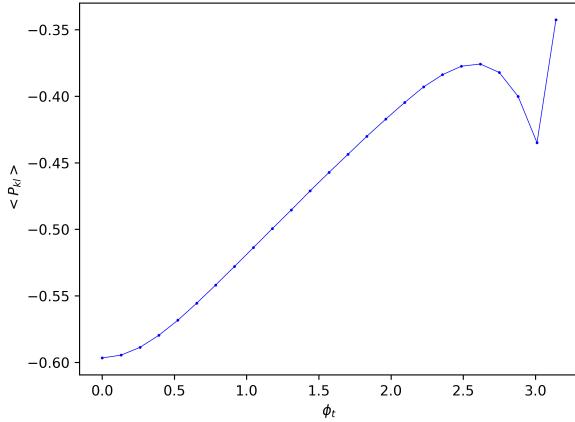


Figure 31: NN contribution to the average bond energy for $t = 0.3$.

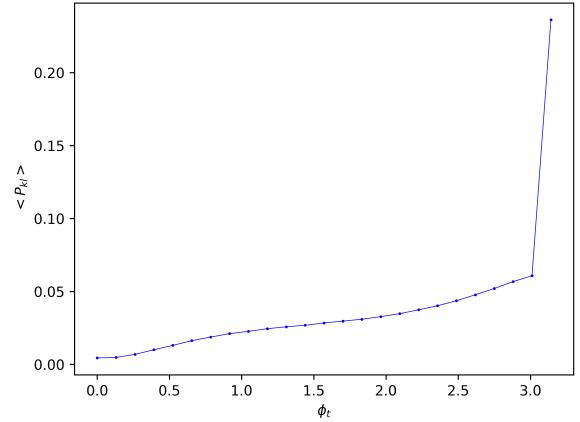


Figure 32: NNN contribution to the average bond energy for $t_2 = 0.3$.

The phase diagram of the model studied by our newly introduced variational state is constructed by finding this minimizing variational parameters for varying coupling constants $J_{2,3}$ (fixing $J_1 = 1$) of the generalized (up to third) nearest neighbor Heisenberg hamiltonian. Now we present close-ups of different regions of this phase diagram.

Let us remark that only the $J_{2,3} < 1$ part of the phase diagram has a physical meaning in terms of interpreting the Heisenberg model as representing the low energy limit of the Hubbard model. However, the remainder of the phase diagram might be relevant for other physical instances. For the regions of the diagram where $t_2 = 0$, the value of ϕ_t has no meaning, and the differently colored areas in the plots stem only from numerical finite sampling error. This is because the definition of ϕ_t only makes sense for $t_2 \neq 0$.

Part VI Outlook

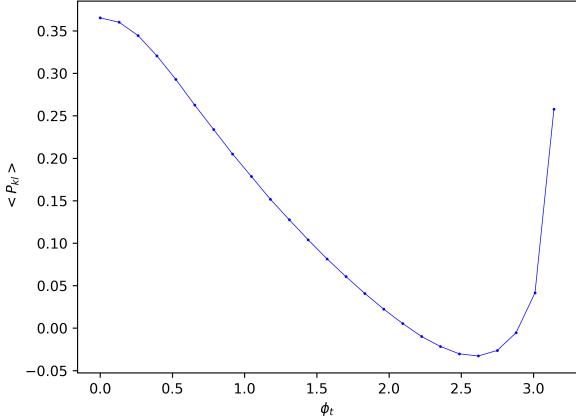


Figure 33: NNNN contribution to the average bond energy for $t_2 = 0.3$.

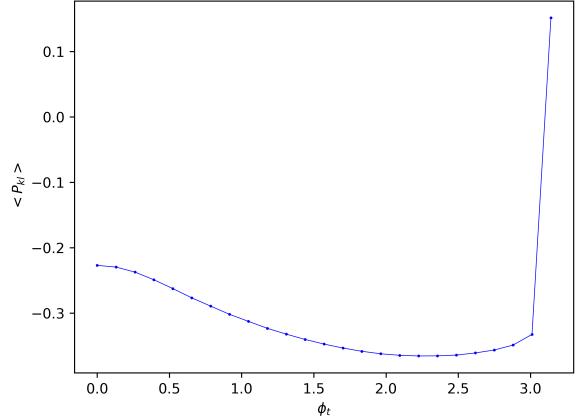


Figure 34: Sum of up to third NN contributions for $t_2 = 0.3$.

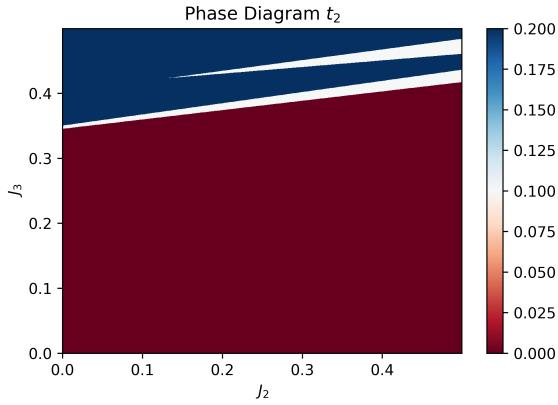


Figure 35: Phase diagram on the $J_2 - J_3 - t_2$ hyperplane

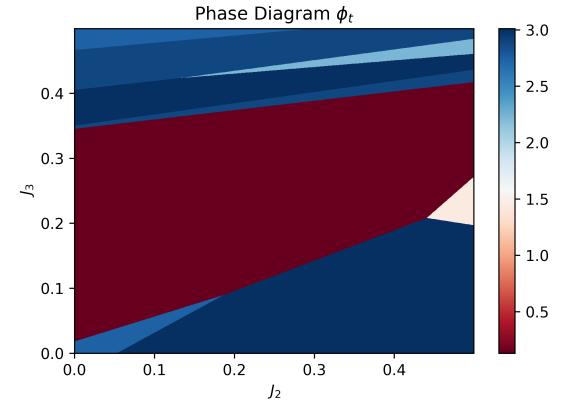


Figure 36: Phase diagram on the $J_2 - J_3 - \phi_t$ hyperplane

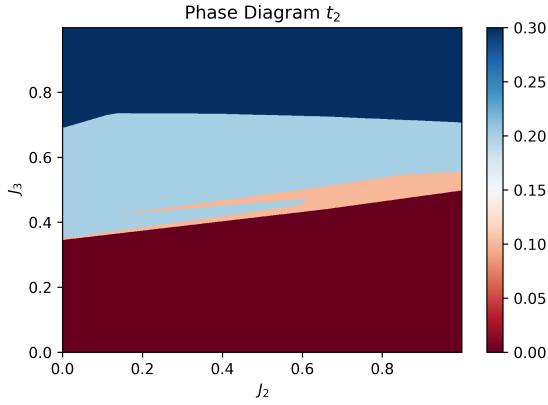


Figure 37: Phase diagram on the $J_2 - J_3 - t_2$ hyperplane

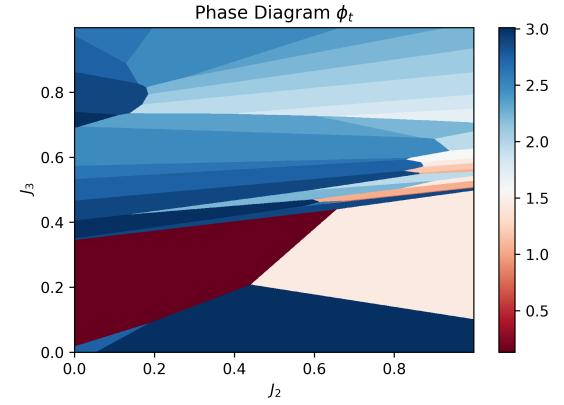


Figure 38: Phase diagram on the $J_2 - J_3 - \phi_t$ hyperplane

It can prove crucial to optimize the variational wave function in VMC, since it can allow more accurate calculations of quantities of interest. This can make the difference between capturing a given phenomenon occurring within the material or not. In this project we probe the spin-orbital quantum liquid on the $SU(4)$ Hubbard model on the honeycomb lattice in the $t/U \ll 1$ limit. We are motivated by the possibility of the opening of a gap in the dispersion relation, leading to a topological insulator phase hosting what is called anomalous quantum Hall effect, that is quantized Hall effect with finite transverse conductivity in the absence

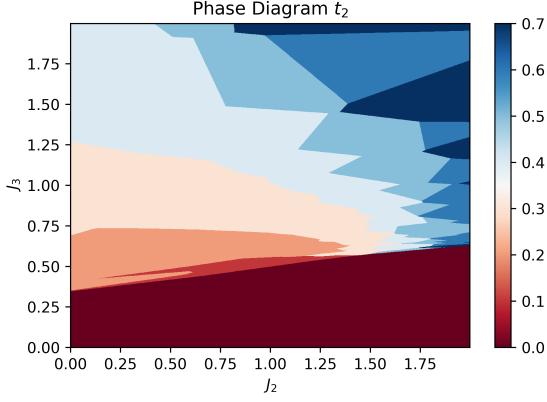


Figure 39: Phase diagram on the $J_2 - J_3 - t_2$ hyperplane

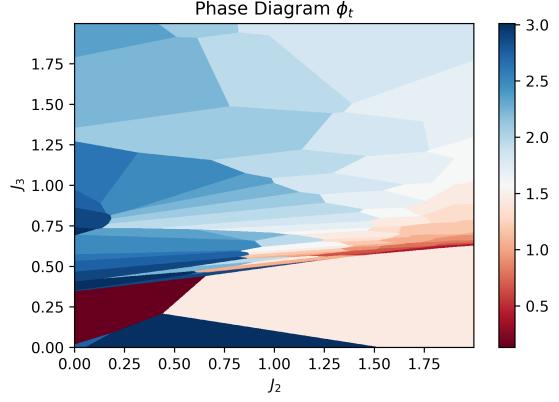


Figure 40: Phase diagram on the $J_2 - J_3 - \phi_t$ hyperplane

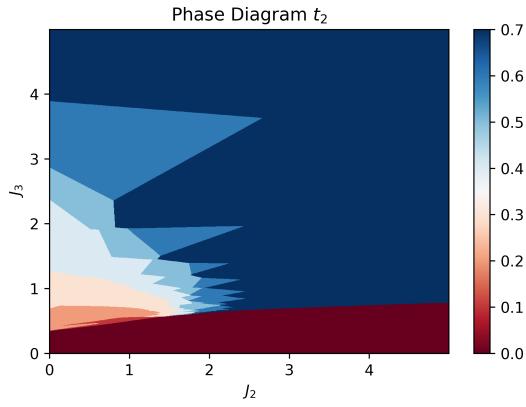


Figure 41: Phase diagram on the $J_2 - J_3 - t_2$ hyperplane

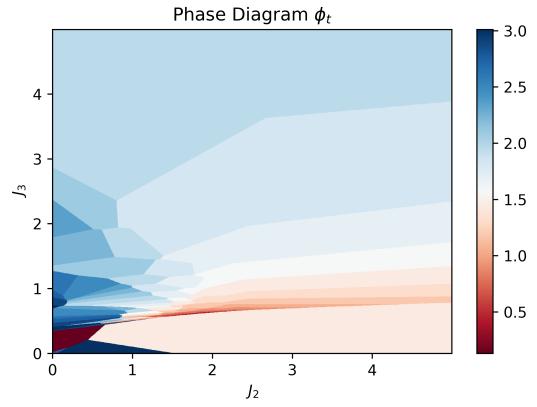


Figure 42: Phase diagram on the $J_2 - J_3 - \phi_t$ hyperplane

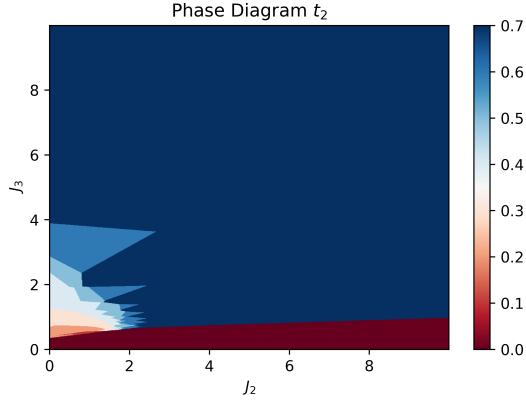


Figure 43: Phase diagram on the $J_2 - J_3 - t_2$ hyperplane

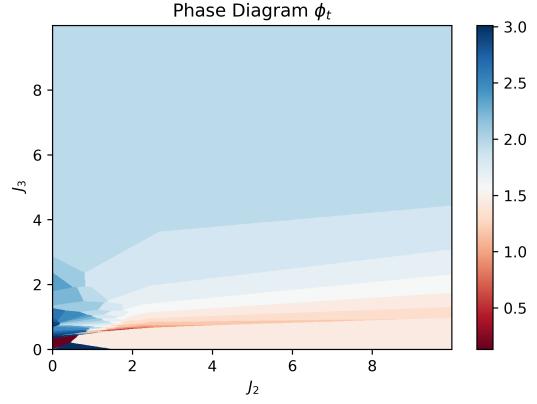


Figure 44: Phase diagram on the $J_2 - J_3 - \phi_t$ hyperplane

of an external magnetic field. Anomalous QHE is a manifestation of topological structure in many-electron systems.

We optimized the variational state by finding hopping configurations leading to lower average exchange energy $\langle P_{kl} \rangle$ of first, second, and third nearest neighbors. By varying the coupling constants corresponding to these exchange interactions, we found the phase diagram of a generalized Heisenberg model, coupling up to third nearest neighbors. The physical motivation to study this model comes from the $t/U \ll 1$ limit expansion of the Hubbard model that maps to it, when we consider two-site exchange terms up to sixth order in degenerate perturbation theory. The expansion gives $J_2 \sim \mathcal{O}(t^4/U^3)$, and $J_3 \sim \mathcal{O}(t^6/U^5)$. More concretely, this means

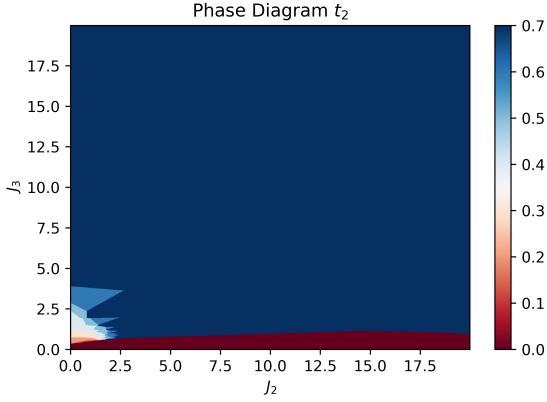


Figure 45: Phase diagram on the $J_2 - J_3 - t_2$ hyperplane

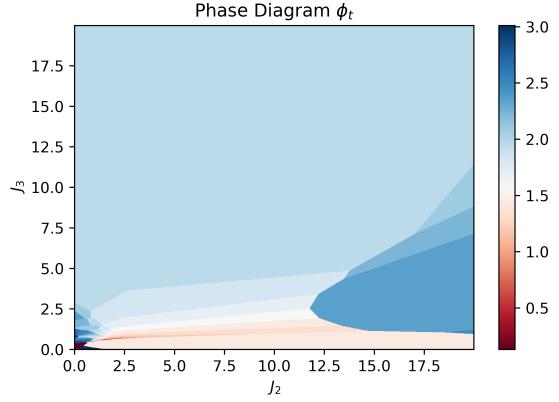


Figure 46: Phase diagram on the $J_2 - J_3 - \phi_t$ hyperplane

that only realizations of the model with $J_3 < J_2 < J_1$ correspond to an interpretation in terms of a limiting case of the Hubbard model.

Hubbard model beyond the Mott-insulating limit

For this project we always take the $t/U \ll 1$ limit of the Hubbard model, meaning that the interaction energy is so high that a singly occupied configuration is decisively favored. In fact, the Gutzwiller projection method starts by assuming only configurations with strictly one fermion per site. While we always remained in the $t/U \ll 1$ limit, the method can be extended to the more general case.

To generalize the procedure that was carried out in this project to the Hubbard model we would have to consider the possibility of sites with no fermions, or doubly occupied sites. We would have to devise a way to keep track of the fermions with each color, and now, because fermions can simply move between sites without necessarily exchanging sites with another fermion, we need to have a separate list for each color, keeping track of the sites containing fermions with that color. Another aspect to take into account is that each site can only have up to 4 - the number of colors - fermions.

In the Hubbard model case, the update rule which allows us to compute acceptance ratios also changes. Equation (24) is changed to account for the penalization of doubly occupied sites. Note that we should recover the method we used in the $t/U \ll 1$ limit. A natural choice is to introduce a factor proportional to $e^{-U_{\{j\}}}$, so that now

$$|\psi\rangle = \sum_{\{j\}} \prod_{\alpha=1}^N w_{\{j^\alpha\}} e^{-U_{\{j\}}} |j_1^\alpha j_2^\alpha \dots j_{N_s/N}^\alpha\rangle \quad (57)$$

If U is large, the configurations with doubly occupied sites have negligible contributions, in the sense that we rarely sample them, giving back the method used throughout the project.

We would also have to derive new update rules. This is because one must now consider the action of the fermion creation and annihilation operators on empty or doubly occupied sites. For example, the $n_i^\alpha n_j^\alpha$ operator is not diagonal anymore.

These generalizations constitute a natural continuation of this project. They would yield a more complete description of the system, for example describing other eventual non trivial phases arising within the model, with possible realization in the examples of $\alpha - ZrCl_3$ and alkaline-earth atoms on an optical lattice.

References

1. Anderson, P. W. Resonating valence bonds: A new kind of insulator? *Mater. Res. Bull.* **8**, 153–160 (1973).
2. Cazalilla, M. A. & Rey, A. M. Ultracold Fermi Gases with Emergent SU(N) Symmetry. *Reports on Progress in Physics* **77** (2014).
3. Corboz, P. *et al.* Spin-Orbital Quantum Liquid on the Honeycomb Lattice. *Physical Review X* **2** (2012).
4. Fazekas, P. *Lecture Notes on Electron Correlation and Magnetism* (World Scientific, 1999).

5. Haldane, F. D. M. Model for a Quantum Hall Effect without Landau levels: Condensed Matter Realization of the "Parity Anomaly". *Physical Review Letters* **61**, 2015–2018 (1988).
6. Jérôme Dufour, P. N. & Mila, F. Variational Monte Carlo investigation of $SU(N)$ Heisenberg chains. *Phys. Rev. B* **91**, 174427 (2015).
7. Kugel', K. I. & Khomski, D. I. The Jahn-Teller effect and magnetism: transition metal compounds. *Sov. Phys. Usp.* **25** (1982).
8. Lajkó, M. *Fluctuating moments in one and two dimensional Mott insulators* (PhD Thesis, 2013).
9. Lajkó, M. & Penc, K. Tetramerization in a $SU(4)$ Heisenberg model on the honeycomb lattice. *Physical Review B* **87** (2013).
10. M. E. J. Newman, G. T. B. *Monte Carlo Methods in Statistical Physics* (Clarendon Press, 1999).
11. Masahiko G. Yamada, M. O. & Jackeli, G. Emergent $SU(4)$ Symmetry in $\alpha - ZrCl_3$ and Crystalline Spin-Orbital Liquids. (*submitted on arxiv.org on 15 Sep 2017 https://arxiv.org/abs/1709.05252*) (2017).
12. Nicolas Metropolis, S. The Monte Carlo Method. *Journal of the American Statistical Association* **44**, 335–341 (1949).
13. Oudovenko, G. K. V. Spin and orbital excitation spectrum in the Kugel-Khomskii model. *Phys. Rev. B* **56**, 243–246 (1997).
14. Smerald, A. & Mila, F. Exploring the spin-orbital ground state of $Ba_3CuSb_2O_9$. *Phys. Rev. B* **90**, 094422 (9 Sept. 2014).