Alma Mater Studiorum · University of Bologna

School of Science
Department of Physics and Astronomy
Master Degree in Physics

Machine Learning approach to the Extended Hubbard Model

Supervisor:
Prof. Elisa Ercolessi

Co-supervisor:
Simone Tibaldi

Submitted by:
Filippo Caleca

Academic Year 2021/2022

To my friend Giulio, still the best theoretical physicist I know.

Abstract

Il modello di Hubbard $1d$ esteso con potenziale soft-shoulder si è dimostrato estremamente difficile da studiare a causa della sua non risolvibilità analitica e della competizione tra i diversi termini dell'Hamiltoniana. Per questo motivo abbiamo deciso di utilizzare tecniche di Machine Learning per studiarne il diagramma di fase nel caso di riempimento $\rho = 2/5$ e range del potenziale soft-shoulder $r_c = 2$, ottenendo un diagramma di fase estremamente ricco. Detti U, V i parametri dell'Hamiltoniana associati rispettivamente al potenziale di Hubbard e al potenziale soft-shoulder, abbiamo accertato che per $V \lesssim 5$ e $U \gtrsim 3$ il sistema è sempre un liquido di Tomonaga Luttinger, diventando poi un liquido di Luttinger composto da cluster per $5 \lesssim V \lesssim 7$ (con struttura a blocchi dipendente dal valore di relativo di U e V), e infine subisce un processo di cristallizzazione per $V \gtrsim 7$, diventando un cristallo nella regione $U \lesssim 3V/2$ e $U \gtrsim 5$. Infine abbiamo constatato che per $U \lesssim 5$ e $V \gtrsim 2-3$ il sistema mantiene la struttura di liquido di cluster con un residuo grado di mobilità intra-blocco delle particelle.

Abstract

The $1d$ extended Hubbard model with soft-shoulder potential has proved itself to be very difficult to study due its non solvability and to competition between terms of the Hamiltonian. Given this, we tried to investigate its phase diagram for filling $\rho = 2/5$ and range of soft-shoulder potential $r_c = 2$ by using Machine Learning techniques. That led to a rich phase diagram; calling U, V the parameters associated to the Hubbard potential and the soft-shoulder potential respectively, we found that for $V \lesssim 5$ and $U \gtrsim 3$ the system is always in Tomonaga Luttinger Liquid phase, then becomes a Cluster Luttinger Liquid for $5 \lesssim V \lesssim 7$ (with different block structure depending on the relative values of U and V), and finally undergoes a general crystallization for $V \gtrsim 7$, with a quasi-perfect crystal in the $U \lesssim 3V/2$ and $U \gtrsim 5$ region. Finally we found that for $U \lesssim 5$ and $V \gtrsim 2-3$ the system shall maintain the Cluster Luttinger Liquid structure, with a residual in-block single particle mobility.

Contents

Introduction **7**

1 Many-fermion systems **11**
- 1.1 The Fermi gas 11
- 1.2 Fermi Liquids 12
  - 1.2.1 Failure at $d = 1$ 15

2 Bosonization **19**
- 2.1 Symmetries of the model 20
- 2.2 Luttinger Hamiltonian 22
  - 2.2.1 Interactions 26
  - 2.2.2 Systems with spin 27
- 2.3 Phenomenological bosonization 29
- 2.4 The Hubbard term 31
  - 2.4.1 Sine-Gordon Hamiltonian 32
- 2.5 The soft-shoulder potential 38

3 The extended Hubbard model with soft-shoulder potential: a review **43**
- 3.1 The classical limit 43
- 3.2 Liquid of clusters 45
- 3.3 Phase diagram proposal 46
- 3.4 Perturbation theory $V = 0$, $U \gg t$ 47
- 3.5 Perturbation theory $U \gg V \gg t$ 51

4 Machine Learning techniques and DMRG **55**
- 4.1 Machine Learning 55
  - 4.1.1 Principal Component Analysis 56
  - 4.1.2 k-means clustering 57
  - 4.1.3 Learning by confusion 57

4.2 DMRG 59
4.3 Applications to physics models 61
  4.3.1 The Ising model 61
  4.3.2 The Kitaev model 61
  4.3.3 Applications of Machine Learning 62

5 ML techniques for the Extended Hubbard Model 69
5.1 Observables 69
  5.1.1 Charge structure factor 69
  5.1.2 Spin structure factor 71
  5.1.3 Local density 72
  5.1.4 Double occupancy and nearest and next-to-nearest neighbors numbers 72
5.2 k-means 73
5.3 PCA 80
5.4 Learning by Confusion 83
5.5 Order parameters 86
5.6 Phases of the system 88

6 Conclusions 95

Introduction

The study of strongly correlated electron systems has proved itself fundamental in order to understand peculiar aspects of nature such as magnetism and superconductivity. Some examples are the Kitaev model [9, 26], Bardeen–Cooper–Schrieffer (BCS) theory [2, 3, 8] and the Hubbard model [18, 37, 43]. In particular the latter Hamiltonian describes the behaviour of spin 1/2 particles in a lattice and is composed by two elements: the quantistic hopping term, which makes particles jump from a site to its neighbours, and the on-site potential, which can disfavour or favour double occupancies depending on its repulsive or attractive nature.

This model has been widely studied in $d = 1$ due to its exact solvability thanks to Lieb and Wu's 1968 Bethe ansatz solution [22], which made it a theoretical laboratory for non-perturbative effects. An almost complete review of the model can be found in [20]. Here we briefly review its phase diagram, which depends onto two parameters, those being the Hamiltonian parameter associated to the on-site potential and the filling of the chain. For systems with repulsive potential (which is the natural choice if one wants to reproduce the behaviour of electrons), for a generic filling spin fluctuations dominate over superconducting ones. In fact, as one would expect, particles are forced not to be on the same site so the system is a spin density wave. That consists in anti-ferromagnetic and low-conducting state, which is more and more dominant as the on-site potential becomes more repulsive. That changes only at half-filling (i.e. one particle per site), when the system becomes a Mott insulator (which is an insulating state which cannot be described by conventional band theories [29]) for all repulsive interactions. On the other hand, when the potential is attractive, fermions tend to pair in singlets for every filling and the system becomes superconducting. It is important to underline the richness of this phase diagram; in fact, despite the simpleness of the Hamiltonian, the model exhibits both superconducting and Mott insulator phases, which are non-trivial states of matter. One of the main factors which causes such a richness is low dimensionality. It shall be stressed in fact (and it will be done in the following chapters) that the $d = 1$ case is dramatically different to the others. As a matter of fact, differently to all other frameworks, in such a scenario single-particle excitations do not exist, as low dimensionality force all excitations to be collective. Because of this, Fermi liquids [40]

do not exist in $d = 1$, and are substituted by Tomonaga Luttinger Liquids (commonly referred to as TLL), which will be carefully examined in the following.
Moreover, technological progress has made possible the experimental realization of the Hubbard model, promoting it from a simple theoretical toy model to an experimental prototype for strongly correlated electrons. Some examples of quantum simulation of the Hubbard model are metal-insulator transitions in organic solids [11], quantum simulation with cold atoms [41], and dopant atoms in Silicon [36].
As the Hubbard model has been widely studied due to the previously mentioned nice properties which characterize it, many extension have been proposed and examined. Some of them consists in modifying the hopping term by introducing second and third nearest neighbour hopping [23], while others consists in modification of the potential [16]. In this thesis we are going to focus on one of the latter, which is obtained by including a new term into the Hamiltonian, i.e. the soft-shoulder potential (this is sometimes called the U, V model, e.g. in [15]). This new term introduces an off-site and spin-independent interaction to the previous framework and drastically changes the physics of the model. One example of this dramatic change is the fact that the Mott insulator phase [28] can be reached at fillings which are not necessarily half-filling (i.e. one particle per site); for example taking just a nearest-neighbour repulsive interaction, that can be achieved for $\rho = 1/4$ (i.e. one particle every two sites). That is due to effect of frustration of particles [30, 5], which is caused by the competition between the three terms of the Hamiltonian. Another example is the emergence of phases with underlying block-structures (in particular we will see Cluster Luttinger Liquids CLL), as pointed out in [6, 27].
In order to study the phase diagram of this model, which can be a difficult task due to the previously mentioned competition, we decided to use Machine Learning (often referred to as ML in the following) techniques. Machine Learning has become more and more important in physics, as its use has had an exponential growth in the last years, see Fig. 1. In particular we will focus on unsupervised Machine Learning techniques for phase recognition; unsupervised methods were prefered to supervised ones as they do not require any a priori knowledge of the system.
The work is structured as follows:

§ In Chapter 1 we briefly review some key aspects of quantum many-body theory, with a particular focus on Fermi gas. We also give an empirical and almost-formal prove of the failure of Fermi gas theory in the case of $d = 1$.

§ In Chapter 2 we move on the $d = 1$ picture. We introduce the $1d$ extended Hubbard Hamiltonian and use it as a model to study bosonization, which is the main topic of this chapter. Bosonization, which is a mathematical technique which makes fermions and bosons the same (this is only possible in $d = 1$), is tackled from both

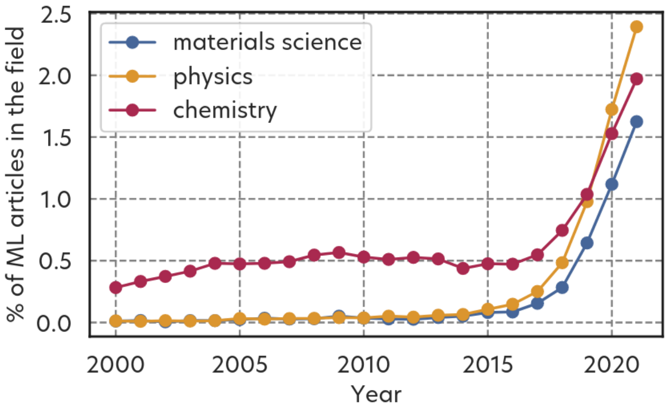


Figure 1: Percentage of ML-based publications in material science, physics and chemistry from 2000 to 2020. Image from [10], adapted from [4].

formal and phenomenological arguments. That naturally leads to the discussion of Luttinger liquids, which are the analogue of Fermi gases in $d = 1$. Finally this chapter also includes a brief discussion of the Renormalization Group flow of the $1d$ extended Hubbard Hamiltonian.

§ Chapter 3 is dedicated to the discussion of previously proposed phases of the system. That includes the introduction of the liquid of clusters phase, perturbation theory for $U \gg t$ with $V = 0$ and perturbation theory with $U \gg V \gg t$. A previous proposal for the phase diagram of the $1d$ extended Hubbard model is also included.

§ Chapter 4 focuses on numerical and ML techniques which were used in this work. That therefore includes the discussion of Density Matrix Renormalization Group (DMRG), Principal Component Analysis (PCA), k-means clustering and Learning by Confusion.

§ Chapter 5 includes the result of the application of the ML techniques to the $1d$ extended Hubbard model. After the exam of these, our proposal for the phase diagram of the system, along with a detailed explanation of all possible phases, is discussed.

§ Finally in the conclusions we briefly resume and discuss the content of the main outcomes of this work.

Chapter 1

Many-fermion systems

> *"The more, the merrier!"*
>
> Common quote

By studying Quantum Mechanics one immediately realizes that the exam of quantum many-body systems presents great technical difficulties. On the other hand, studying these systems is necessary to understand matter's strangest aspects, such as superconductivity or superfluidity.
In the following, we will mainly focus on systems of fermions (we'll see that in one dimension it does not make sense to make such a strict distinction), first considering non-interacting systems (the Fermi gas), then we'll show a good approximation of interacting ones (Landau's Fermi liquids) and finally, we'll show that this approximation fails in one dimension.

1.1 The Fermi gas

Taking $\hbar = 1$ (we will always do so in the following), a non-interacting, translationally invariant system has plane waves as single particle eigenstates with energy $\varepsilon_k = k^2/2m$. That being said, at $T = 0$, for a system of N fermions, the ground state corresponds to the notorious Fermi sea, which means that every momentum eigenstate is filled up to the *Fermi momentum* k_F, whose correspondent energy is the *Fermi level* $\varepsilon_F = k_F^2/2m$. The Hamiltonian therefore reads

$$H = \sum_{\vec{k}} \xi_{\vec{k}} n_{\vec{k}} \tag{1.1}$$

where $n_{\vec{k}}$ is the number of particles with momentum $\vec{k}$, the sum over spin is implicit and $\xi_{\vec{k}} = \varepsilon_{\vec{k}} - \mu$, where μ is the chemical potential. Starting from this expression one can derive [1] the specific heat

$$C_V(T) = \gamma T = \frac{2\pi^2}{3} N(\varepsilon_F) k_B^2 \, T \tag{1.2}$$

where k_B is the Boltzmann constant and $N(\varepsilon)$ is the number of particles with energy ε (commonly called *density of states*). Moreover it is possible to compute an expression for spin susceptibility

$$\chi = 2N(\varepsilon_F)\mu_B^2 \tag{1.3}$$

where μ_B is the Bohr magneton $\mu_B = e/2m_e$ (with e and m_e being the electron charge and mass respectevely). Finally one also find an expression for the compressibility

$$\kappa = 2N(\varepsilon_F)/\rho^2. \tag{1.4}$$

where ρ is the mass density of the system. Given these one can easily find that the so-called Wilson ratio, defined as

$$R_W = \frac{1}{3}\left(\frac{\pi k_B}{\mu_B}\right)^2 \frac{\chi}{C_V/T} \tag{1.5}$$

is equal to the unity for non-interacting fermionic systems, independently of the density $N(\varepsilon)$.

1.2 Fermi Liquids

Non-interacting systems are as nice as, in a sense, useless. They do have nice features which theoretical physicists enjoy (one above all their nice analytical representation), but they do not present real physical interest. The question of how interaction modifies this picture naturally arises and Landau's theory represents, with some limits due to its nature of approximation, a good answer.

The fundamental hypothesis of Fermi liquids theory is that interactions do not change the picture *dramatically.* More formally, we can imagine switching on the interaction by using a continuous parameter V and our assumptions consist in taking the ground state energy and the low-lying eigenstate as a continuous function of V. Consider, for example, a low-lying eigenstate, the state obtained by adding one particle with momentum $\vec{p}$ (which must satisfy $|\vec{p}| > k_F$ because of Pauli exclusion principle)

---

[1] See [33] for more details

$$a^\dagger_{\vec{p}}|0,N\rangle = |\vec{p},N+1\rangle$$

where $a^\dagger_{\vec{p}}$ is the fermionic one particle creation operator for states with momentum $\vec{p}$ and $|0,N\rangle$ is the non-interacting ground state with N particles. Following our continuity assumption, in the interacting case, the momentum $\vec{p}$ still must be greater than the Fermi momentum, but interaction will modify the energy of the state and the $\vec{k}$-distribution of the particles. The complex of the added particles and the disturbed distribution will be the so-called *Landau's quasiparticle*. One can consider analogously the state with total momentum $\vec{p}$ with an added hole, which is nothing but the state obtained by acting on $|0,N\rangle$ with the destruction operator $a_{-\vec{p}}$.

We now want to investigate the physical properties of low-lying excitations in the interacting system. Starting from the ground state quasiparticle $\vec{k}$ distribution

$$n(\vec{k}) = \begin{cases} 1 & \text{if } |\vec{k}| < k_F \\ 0 & \text{if } |\vec{k}| > k_F \end{cases} \tag{1.6}$$

the shift in the energy due to the disturbance in $n(\vec{k})$ caused by the interactions (which we will indicate with $\delta n(\vec{k})$) is

$$\delta E = \sum_{\vec{k}} \varepsilon^0_{\vec{k}}\ \delta n(\vec{k}) + \frac{1}{2V}\sum_{\vec{k}\vec{k'}} f(\vec{k},\vec{k'})\ \delta n(\vec{k})\delta n(\vec{k'}). \tag{1.7}$$

The last expression needs further examination. The energy shift is divided into two different terms, the first one representing the energy of a single "bare" quasiparticle while the second takes into consideration the interaction between quasiparticles. One could now argue what $\varepsilon^0_{\vec{k}}$ and $f(\vec{k},\vec{k'})$ are. The first one depends on the microscopic aspects of the theory, but can be expanded near the Fermi surface as

$$\varepsilon^0_{\vec{k}} = \frac{k_F}{m^*}(k-k_F) + \mathcal{O}(k^2) \tag{1.8}$$

where m^* defines the *effective mass* and shall be thought of as an expansion parameter. The analysis of $f(\vec{k},\vec{k'})$ is less trivial. First, we cannot assume that it does not depend on the spin, so we shall divide it into symmetric and anti-symmetric parts w.r.t. spin as

$$\begin{aligned} f(\vec{k}\uparrow,\vec{k'}\uparrow) &= f^s(\vec{k},\vec{k'}) + f^a(\vec{k},\vec{k'}) = f(\vec{k}\downarrow,\vec{k'}\downarrow) \\ f(\vec{k}\uparrow,\vec{k'}\downarrow) &= f^s(\vec{k},\vec{k'}) - f^a(\vec{k},\vec{k'}) = f(\vec{k}\downarrow,\vec{k'}\uparrow) \end{aligned} \tag{1.9}$$

then assuming rotational invariance and that all processes under exam happen near the Fermi surface one finds that $f^{s,a}(\vec{k},\vec{k'})$ shall only depend on the angle between $\vec{k}$ and $\vec{k'}$

which will be called θ and whose cosine can be expressed as

$$\cos\theta \simeq \frac{\vec{k}\cdot\vec{k'}}{k_F^2}.$$

One can now expand $f^{s,a}(\vec{k},\vec{k'})$ in terms of Legendre polynomials

$$f^{s,a}(\vec{k},\vec{k'}) = \sum_{L=0}^{\infty} f_L^{s,a} P_L(\cos\theta) \equiv \sum_{L=0}^{\infty} \frac{\pi^2}{k_F m^*} F_L^{s,a} P_L(\cos\theta) \tag{1.10}$$

where the last equality defines the quantities $F_L^{s,a}$ and it is just a rescale to obtain dimensionless coefficients. We can now express the full energy of a quasiparticle as the difference in δE between a state with $\delta n(\vec{k}) = 1$ and a state with $\delta n(\vec{k}) = 0$ as

$$\varepsilon_{\vec{k}} = \varepsilon_{\vec{k}}^0 + \frac{1}{V}\sum_{\vec{k'}} f(\vec{k},\vec{k'})\ \delta n(\vec{k'}). \tag{1.11}$$

The continuity assumption we made at the beginning states that the interaction does not change *dramatically* the picture. It is therefore reasonable to assume that quasiparticles obey the same statistics as non-interacting particles, that is, in this case, the Fermi-Dirac statistics. For Fermi-Dirac statistics the occupation number is

$$n(k) = \frac{1}{e^{\beta\varepsilon_k} + 1} \tag{1.12}$$

and it is fundamental in order to compute observables. We start by looking at the specific heat C_V. Recalling its definition

$$C_V = \frac{1}{V}\frac{\partial U}{\partial T}$$

where U is the internal energy, one sees [40] that "new" contributions due to thermally excited quasiparticles come from T-dependence of the second term in (1.11), but these only contribute at T^2 order and can therefore be neglected in the low-temperature limit. That implies that in this limit one can simply substitute $\varepsilon_{\vec{k}}$ with $\varepsilon_{\vec{k}}^0$ and get

$$C_V = \frac{m^* k_F}{3} k_B^2 T \tag{1.13}$$

which is identical to the non-interacting expression (apart from the substitution of the bare mass by m^*). For spin susceptibility, the situation is significantly different. Recalling

$$\chi = \left[V\frac{\partial^2 E_0}{\partial M}\right]^{-1}$$

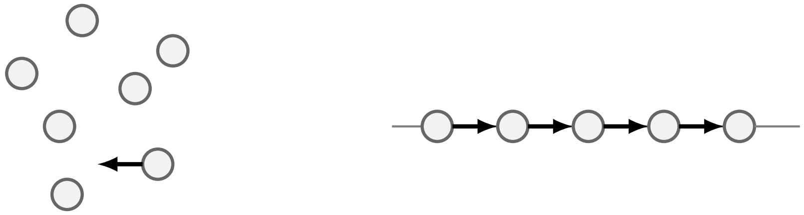


Figure 1.1: Difference between $d = 1$ and higher dimension.

where M is the spin magnetization, which is $M = \mu_B = (N_\uparrow - N_\downarrow)$, one sees that varying M corresponds to a positive (negative) shift in the Fermi momentum for $\uparrow$ ($\downarrow$) particles. Using now (1.7) one has

$$\chi = \frac{1}{1+F_0^a}\frac{\mu_B^2 k_F m^*}{\pi^2} \tag{1.14}$$

which means that now interaction enters also via the adimensional coefficient F_0^a, which also modifies the Wilson ratio

$$R_W = \frac{1}{1+F_0^a}. \tag{1.15}$$

One can also find the compressibility

$$\kappa = -\frac{1}{V}\frac{\partial V}{\partial P} = \left[V\frac{\partial^2 E_0}{\partial V^2}\right]^{-1} = \frac{m^* k_F}{\pi^2 \rho^2 (1+F_0^s)} \tag{1.16}$$

and obtain the m^* dependence on interaction coefficients (that is done by using Galilean invariance)

$$\frac{m^*}{m} = 1 + \frac{F_1^s}{3}. \tag{1.17}$$

1.2.1 Failure at $d = 1$

The idea of Fermi liquids is extremely powerful at high dimensions but fails at $d = 1$. This can be seen in a quite intuitively way as the whole theory is based on the idea of single particle-hole excitations. Now imagining many body systems e.g. in $d = 3$, one particle getting excited does not imply a collective phenomenon, i.e. it does not necessarily hurt other particles exciting them as depicted in Fig. 1.1.

A more formal argumentation to prove the failure is given in [15] and I recall here the basics concepts. It was proven [50] that the susceptibility that measures the response of $\langle \rho(x) \rangle$ to an external potential in the form

$$H_{int} = \int d^d x \; V(x,t)\rho(x)$$

is given by

$$\chi(q,\omega) = \frac{1}{V}\sum_{k}\frac{f_F(\xi_k) - f_F(\xi_{k+q})}{\omega + \xi(k) - \xi(k+q) + i\delta} \tag{1.18}$$

where $\xi(k) = \varepsilon(k) - \mu$, f_F is the Fermi factor, V is the volume of the system and $\delta \to 0^+$. Now in the case of $\omega = 0$ (static susceptibility), if one finds some values of q such that both $\xi(k)$ and $\xi(k+q)$ are zero, this would lead to singularities. Luckily in high dimensions, this happens for a very limited set of points and it is smoothed out by integration. On the other hand, there may be some values Q such that

$$\xi(k+Q) = -\xi(k) \tag{1.19}$$

which brings

$$\text{Re }\chi(Q,\omega=0) = -\frac{1}{V}\sum_{k}\frac{\tanh(\beta\xi(k)/2)}{2\xi(k)} \tag{1.20}$$

which in the thermodynamic limit becomes

$$\text{Re }\chi(Q,\omega=0) = -\int d\xi \; N(\xi)\frac{\tanh(\beta\xi/2)}{2\xi} \tag{1.21}$$

which is logarithmic divergent near the Fermi level (where $N(\xi)$ is roughly constant). Luckily enough this *nesting* property in high dimension is rather the exception than the rule, and that implies that the susceptibility itself does not diverge while its derivatives are generally singular. On the other hand in $d = 1$ the nesting property is *always* satisfied. Close to the Fermi surface (which are Fermi points in $d = 1$), one has

$$\begin{aligned} \xi(k) &\simeq v_F(k - k_F) \quad k \sim k_F \\ \xi(k) &\simeq v_F(-k - k_F) \quad k \sim -k_F \end{aligned}$$

from which is clear that $Q = 2k_F$ shall be a nesting vector for all values of k near the Fermi points (we have assumed that the Fermi velocity is the same on the two Fermi points; this is in general the case because of the inversion properties of practically all systems).

It is possible to examine other reasons why Landau's theory fails in the one-dimensional case, e.g. the singularity which arises in the particle pairing, which is the response to a pair creation/destruction potential as

$$H_{pair} = \int dx \; V(x,t)[\psi(x)^\dagger\psi(x')^\dagger + \text{h.c.}]$$

but we will not examine them in this thesis as we are more interested in examining the properties of the extended Hubbard model.
What I want to underline now is the physical reason which makes Landau's theory fail. That is that Fermi liquids rely on the fundamental hypothesis that the excitation is indeed a small perturbation of the non-interacting theory, the reason for it to be small being the fact that the excitation is indeed *single-particle*. In one dimension that is never the case. In fact, due to the reduction of momentum space, every excitation is a collective one and every the physical phenomenon is itself due to the collective behavior of particles.

Chapter 2

Bosonization

> *"Everything in this world is magic,*
> *except to the magician."*
>
> Robert Ford - Westoworld

To make our considerations useful and clear we shall directly examine the extended Hubbard Hamiltonian in $1d$:

$$H = -t \sum_{i,\sigma=\uparrow,\downarrow} \left(c^{\dagger}_{i+1,\sigma} c_{i,\sigma} + c^{\dagger}_{i,\sigma} c_{i+1,\sigma} \right) + U \sum_{i} n_{i\uparrow} n_{i\downarrow} + V \sum_{i} \sum_{l=1}^{r_c} n_i n_{i+l} \tag{2.1}$$

where $c_{i,\sigma}$ and $c^{\dagger}_{i,\sigma}$ are the fermionic annihilation and creation operators of fermions with spin σ at the site i of a chain respectively, $n_{i\sigma} = c^{\dagger}_{i,\sigma} c_{i,\sigma}$ is the number operator for particles of species σ at site i and $n_i = n_{i\uparrow} + n_{i\downarrow}$ is the total number operator of particles at site i. The first term is the kinetical term and it is often referred to as *hopping term* as its effect is to make a paticle jump from a site to its neighbour. The second term is called the *Hubbard term* and it is an on-site potential which can be either attractive or repulsive depending on the sign of U. Finally the last term represents the extension to the standard Hubbard model due to its non-zero range. In fact, differently to the Hubbard term, this makes particles on different sites interact independently on their spin (in fact only total number operators n_i are involved). The range of the potential depends on the parameter r_c, which represents the range of the potential. In the following we shall consider $r_c = 2$, which implies that particles interact with their nearest neighbours and with next-to-nearest neighbours. We will often refer to this last term as *soft-shoulder* potential. We start by looking at the symmetries of the model, then we'll divide the Hamiltonian into three different terms and study each of them through bosonization.

2.1 Symmetries of the model

As pointed out in [12] the Hubbard model (i.e. $V = 0$) enjoys many symmetries. We start by looking at conserved quantities. One can trivially check that the total number of particles

$$N = N_\uparrow + N_\downarrow \qquad \text{with} \qquad N_\sigma = \sum_i c^\dagger_{i\sigma} c_{i\sigma} \tag{2.2}$$

and the total magnetization

$$S^z = \frac{1}{2}\big(N_\uparrow - N_\downarrow\big) \tag{2.3}$$

are conserved (in particular $N_\uparrow$ and $N_\downarrow$ are separately conserved i.e. $[H, N_\sigma] = 0$). As the number of particles is conserved, it is possible to add a term proportional to N to the Hamiltonian without modifying the physics of the system. That implies a modification to the Hubbard term which will be useful when using bosonization dictionary

$$U \sum_i n_{i\uparrow} n_{i\downarrow} \;\longrightarrow\; U \sum_i \left(n_{i\uparrow} - \frac{1}{2}\right)\left(n_{i\downarrow} - \frac{1}{2}\right).$$

It is possible to define the other two components of the magnetization $S^{x,y}$ (which also commute with the Hamiltonian) as

$$S^\alpha = \sum_i S^\alpha_i = \frac{1}{2} \sum_{\sigma,\sigma'=\uparrow,\downarrow} \sum_i \; c^\dagger_{i\sigma} (\sigma^\alpha)_{\sigma\sigma'} c_{i\sigma'} \tag{2.4}$$

where σ^α, with $\alpha = x, y, z$ are the Pauli matrices

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Consequently

$$\begin{aligned} S^x_i &= \frac{1}{2}\left(c^\dagger_{i\uparrow} c_{i\downarrow} + c^\dagger_{i\downarrow} c_{i\uparrow}\right) \\ S^y_i &= \frac{1}{2i}\left(c^\dagger_{i\uparrow} c_{i\downarrow} - c^\dagger_{i\downarrow} c_{i\uparrow}\right) \\ S^z_i &= \frac{1}{2}\left(n_{i\uparrow} - n_{i\downarrow}\right). \end{aligned} \tag{2.5}$$

The conservation of the three total components of the spin (it shall be underlined that only total components are conserved, while local are not) implies the symmetry under the $SU(2)$ rotation group in spin space. This symmetry will be useful when dealing with the soft-shoulder potential term. Another $SU(2)$ symmetry is related to the components of charge operators, which can be obtained by spin operators through a particle-hole transformation only on ↓ fermions combined with a change of sign on the odd-sites sublattice

$$c_{i\uparrow} \to c_{i\uparrow} \qquad , \qquad c_{i\downarrow} \to (-1)^i c^\dagger_{i\downarrow}. \tag{2.6}$$

That brings

$$\begin{aligned} C^x_i &= \frac{(-1)^i}{2}\Big(c^\dagger_{i\uparrow}c^\dagger_{i\downarrow} + c_{i\uparrow}c_{i\downarrow}\Big) \\ C^y_i &= \frac{(-1)^i}{2i}\Big(c^\dagger_{i\uparrow}c^\dagger_{i\downarrow} - c_{i\uparrow}c_{i\downarrow}\Big) \\ C^z_i &= \frac{1}{2}\Big(n_{i\uparrow} + n_{i\downarrow} - 1\Big). \end{aligned} \tag{2.7}$$

It shall be noted that the symmetry relation

$$[H, C^\alpha] = 0$$

only holds for an even number of sites (which will always be the case in the examined numerical simulations). This imposes some constraints on the joint irreducible representation and implies that the total symmetry group is

$$SU(2) \times SU(2)/\mathbb{Z}_2 = SO(4). \tag{2.8}$$

It is worth mentioning that the transformation (2.6) maps the charge sector into the spin sector and the repulsive Hubbard term into the attractive one, i.e.

$$H(t, U) \to H(t, -U).$$

We also observe that a magnetic field term like $B\sum_i S^i_z$ in the Hamiltonian would break spin-rotational symmetry, while a chemical potential term like $\mu\sum_i n_i$ term would break charge symmetry.

We finally list the discrete symmetries of the model

- Under periodic boundary conditions, translation of length L, $c_{i\sigma} \to c_{i+L,\sigma}$
- Reflection $c_{i\sigma} \to c_{L-i+1,\sigma}$
- Particle hole symmetry $c_{i\sigma} \to (-1)^i c^\dagger_{i\sigma}$ (for bipartite lattices only)
- Spin flips $c_{i\sigma} \to c_{i\bar\sigma}$ (where $\bar\sigma =\downarrow$ if $\sigma =\uparrow$ and viceversa)

* Time reversal $c_{i\sigma} \to \sigma c_{i\sigma}$ with $\uparrow = +1$, $\downarrow = -1$.

The soft-shoulder potential $V \sum_i \sum_{l=1}^{r_c} n_i n_{i+l}$ reduces the $SU(2)$ charge symmetry into $U(1)$. The total symmetry is therefore

$$SU(2) \times U(1) = U(2). \tag{2.9}$$

2.2 Luttinger Hamiltonian

We divide the Hamiltonian into three terms

$$H \equiv H_K + H_U + H_V$$

and start by considering the kinetical part, i.e. H_K. Fourier transforming it as

$$c_{i,\sigma} = \sum_k e^{ikx_i} c_{k,\sigma} \quad , \quad c_{i,\sigma}^\dagger = \sum_k e^{-ikx_i} c_{k,\sigma}^\dagger$$

one has

$$H_K = \sum_{k,\sigma} \varepsilon_k c_{k,\sigma}^\dagger c_{k,\sigma} \tag{2.10}$$

with $\varepsilon_k = 2\cos(ka)$ where a is the lattice spacing. If we imagine to ignore the exact expression for ε_k and being interested only in low-level excitations, we can approximate

$$\varepsilon_k \simeq v_F \left(\pm k - k_F \right)$$

where ± is there because of the different signs that the expansions at the two Fermi points carry. Introducing a multiplicative factor $r = \pm 1$ in order to differentiate the two different expansions, the resulting form of the Hamiltonian is (from now on we are omitting spin index to simplify the notation)

$$H_K = \sum_{k,r=\pm 1} v_F (rk - k_F) \; c_{r,k}^\dagger c_{r,k} \tag{2.11}$$

which is known as the Luttinger Hamiltonian. In the following, we are going to call particles created near $k = k_F$ *right-going* ($r = +1$ or $r = R$) and particles created near $k = -k_F$ *left-going* ($r = -1$ or $r = L$).
Within this frame, particle-hole excitations' energy is (for right-going)

$$E_{R,k} = v_F (k + q) \; - \; v_F k = v_F q. \tag{2.12}$$

We can now define the density fluctuation operator as

$$\rho_r(q) = \sum_k c^\dagger_{k+q,r} c_{k,r} \tag{2.13}$$

which is indeed a superposition of particle-hole excitations. Since it is the product of two fermionic operators, it has bosonic character. We shall now prove, to ensure a correct mapping from fermions to bosons, that the Hamiltonian can be written in terms of these and that the correspondence is, so to say, complete (one-to-one correspondence between operators). We shall now pause and focus on the nature of the ground state. The Hamiltonian (2.11) implies an infinite number of occupied states, so one needs to be careful in defining density operators to avoid infinities. To do so in the following we will use normal ordering, which is defined as

$$:AB: \equiv AB - \langle 0| AB |0\rangle \tag{2.14}$$

where $|0\rangle$ is the ground state of (2.11), that is the $1D$ Dirac sea. The normal ordered density operator is defined as

$$:\rho_r(x): \equiv :\psi^\dagger_r(x)\psi_r(x): \tag{2.15}$$

and has Fourier expansion

$$:\rho_r(x): = \sum_k :\rho_r(k): e^{ikx}$$

so that

$$\begin{aligned} :\rho^\dagger_r(p): &= \sum_k c^\dagger_{r,k+p} c_{r,k} \quad (p \neq 0) \\ &= \sum_k [c^\dagger_{r,k} c_{r,k} - \langle 0| c^\dagger_{r,k} c_{r,k} |0\rangle] \equiv N_r \quad (p = 0). \end{aligned} \tag{2.16}$$

It is trivial to see that the fact that $\rho(x)$ is real implies $\rho^\dagger(q) = \rho(-q)$. Let us now look at commutations relations between bosonic operators. It is immediate to verify

$$[\rho_R(q), \rho_L(q)] = 0.$$

Considering now fermions od the same species and the relation

$$[AB, C] = ABC - CAB + ACB - ACB = A\{B, C\} - \{C, A\}B$$

one obtains

$$[\rho^\dagger_r(p), \rho_r(p')] = \sum_k (c^\dagger_{r,k+p-p'} c_{r,k} - c^\dagger_{r,k+p} c_{r,k+p'})$$

One would now be led to a change of variable to obtain zero. That would be true if we were not operating with infinities, which complicate (but indeed save) the situation. To remove these infinities we shall operate with normal-ordered operators. By doing so one finds

$$[\rho_r^\dagger(p), \rho_r(p)] = \delta_{p,p'} \sum_k (\langle 0| c^\dagger_{r,k} c_{r,k} |0\rangle - \langle 0| c^\dagger_{r,k+p} c_{r,k+p} |0\rangle)$$

As the momentum is quantized, by taking periodic boundary conditions $k_n = 2\pi n/L$ one finds

$$[\rho_r^\dagger(p), \rho_r'(p')] = -\delta_{r,r'}\delta_{p,p'}\frac{rpL}{2\pi}. \tag{2.17}$$

This is crucial as it shows that density operators can be mapped to bosonic creation/annihilation operators. It is easy to realize that

$$\begin{aligned} \rho_R^\dagger(p<0)\,|0\rangle &= 0 \\ \rho_L^\dagger(p>0)\,|0\rangle &= 0 \end{aligned}$$

thus we define the new bosonic operators as (here $\Theta(x)$ is the step function)

$$\begin{aligned} b_p^\dagger &= \left(\frac{2\pi}{L|p|}\right)^{1/2} \sum_r \Theta(rp)\rho^\dagger(p) \\ b_p &= \left(\frac{2\pi}{L|p|}\right)^{1/2} \sum_r \Theta(rp)\rho(p) \end{aligned} \tag{2.18}$$

which are only defined for $p \neq 0$. We now want to find out if the Hamiltonian can be expressed in terms of these new operators. We compute

$$[b_{p_0}, H] = v_F p_0 b_{p_0}$$

assuming the basis is complete one arrives at the conclusion

$$H \simeq \sum_{p\neq 0} v_F |p| b_p^\dagger b_p. \tag{2.19}$$

This result is remarkable as it shows that the kinetic Hamiltonian is indeed quadratic in bosonic operators, and not quartic as one would naively guess. The same can be done for real space one particle creation and annihilation operator $\psi_r(x)$, $\psi_r^\dagger(x)$, obtaining

$$[\rho_r^\dagger(p), \psi_r(x)] = -e^{ipx}\psi_r(x) \qquad \psi_r(x) \simeq e^{\sum_p e^{ipx}\rho(p)(\frac{2\pi r}{pL})} \tag{2.20}$$

To complete the mapping one needs to add the $p = 0$ term to the Hamiltonian and consider the loss of charge due to particle destruction (charge is on the other hand conserved by density operators). Taking care of this

$$\psi_r(x) = U_r e^{\sum_p e^{ipx}\rho(p)(\frac{2\pi r}{pL})} \tag{2.21}$$

where U_r suppresses a charge, and another term is added to the Hamiltonian

$$H_{p=0} = \sum_{p\neq 0} v_F|p|b_p^\dagger b_p + \frac{\pi v_F}{L}\sum_r N_r^2. \tag{2.22}$$

It is now convenient to introduce field operators

$$\varphi(x), \theta(x) = \mp(N_R \pm N_L)\frac{\pi x}{L} \mp \frac{i\pi}{L}\sum_{p\neq 0}\frac{1}{p}e^{-\alpha|p|/2-ipx}(\rho_R^\dagger(p) \pm \rho_L^\dagger(p)) \tag{2.23}$$

or in terms of bosonic creation/annihilation operators

$$\begin{aligned}
\varphi(x) &= -(N_R + N_L)\frac{\pi x}{L} - \frac{i\pi}{L}\sum_{p\neq 0}\left(\frac{L|p|}{2\pi}\right)^{1/2}\frac{1}{p}e^{-\alpha|p|/2-ipx}(b_p^\dagger + b_{-p}) \\
\theta(x) &= (N_R - N_L)\frac{\pi x}{L} + \frac{i\pi}{L}\sum_{p\neq 0}\left(\frac{L|p|}{2\pi}\right)^{1/2}\frac{1}{p}e^{-\alpha|p|/2-ipx}(b_p^\dagger - b_{-p})
\end{aligned} \tag{2.24}$$

where α is a cutoff scale or a mimic of a finite bandwidth (roughly speaking the bandwidth is $\Lambda \sim 1/\alpha$). By working out commutation rules (in the thermodynamic limit and limit $\alpha \to 0$) one gets

$$[\varphi(x), \theta(x')] = i\pi \; Sign(x - x').$$

Similarly one can prove

$$[\varphi(x), \nabla\theta(x')] = i\pi\delta(x - x') \tag{2.25}$$

so the conjugate momentum to the field $\varphi(x)$ is

$$\Pi(x) = \frac{1}{\pi}\nabla\theta(x). \tag{2.26}$$

The Hamiltonian can be rewritten in terms of these field operators as

$$H = \frac{1}{2\pi}\int dx \; v_F[\pi\Pi(x))^2 + (\nabla\varphi(x))^2]. \tag{2.27}$$

In principle, one could now add higher order terms of the Hamiltonian, compute again commutation relations, and modify the density operator definition to maintain a diagonal Hamiltonian. That being said, the picture soon becomes quite cumbersome. For this reason, to include higher order terms we prefer a different approach that we describe in Sect. 2.3. We now look at the effect of point-like interactions.

2.2.1 Interactions

We shall now see how this picture changes if we introduce short-range (indeed point-like) interactions. Possible interactions with coupling constant g are

- with $g \equiv g_4$ scattering between particles lying near the same Fermi point, i.e. with $V(q \sim 0)$;
- with $g \equiv g_2$ scattering particles lying near opposite Fermi points (that is equivalent to saying they are different species fermions) which "stay" on their side after interacting, i.e. $V(q \sim 0)$;
- with $g \equiv g_1$ scattering of particles lying near opposite Fermi points which "switch" sides after interacting, i.e. $V(q \sim 2k_F)$.

It is clear that for spinless fermions the last two are the same as one cannot distinguish them. We shall therefore start by looking only at g_4 and g_2 and then turn back to g_1 when dealing with systems with spin. Dividing

$$\psi(x) = \psi_R(x) + \psi_L(x)$$

the g_4 interaction is (for right-going)

$$\frac{g_4}{2}\psi_R^\dagger(x)\psi_R(x)\psi_R^\dagger(x)\psi_R(x) = \frac{g_4}{2}\rho_R(x)\rho_R(x) = \frac{g_4}{2}\frac{1}{(2\pi)^2}(\nabla\varphi - \nabla\theta)^2 \tag{2.28}$$

and similarly for left-going (with $R \to L$, $\varphi - \theta \to \varphi + \theta$). Adding the two contributions is therefore equivalent to a velocity rescale

$$u = v_F\left(1 + \frac{g_4}{\pi v_F}\right). \tag{2.29}$$

A similar treatment can be applied to g_2 interaction

$$g_2\rho_R(x)\rho_L(x) = \frac{g_2}{(2\pi)^2}[(\nabla\varphi)^2 - (\nabla\theta)^2]. \tag{2.30}$$

The entire Hamiltonian can now be rewritten as

$$H = \frac{1}{2\pi}\int dx[uK(\pi\Pi(x))^2 + \frac{u}{K}(\nabla\varphi(x))^2] \tag{2.31}$$

so the entire effect has been reabsorbed into two parameters u, K defined by

$$\begin{aligned} uK &= v_F\left(1 + \frac{g_4}{2\pi v_F} - \frac{g_2}{2\pi v_F}\right) \\ \frac{u}{K} &= \left(1 + \frac{g_4}{2\pi v_F} + \frac{g_2}{2\pi v_F}\right) \end{aligned} \tag{2.32}$$

which can be solved as

$$\begin{aligned} u &= v_F[(1+y_4/2)^2 - y_2^2/2]^{1/2} \\ K &= \left(\frac{1+y_4/2-y_2/2}{1+y_4/2+y_2/2}\right)^{1/2} \end{aligned} \tag{2.33}$$

with $y_i = g_i/\pi v_F$. Note that quite generally $K > 1$ for attractive interactions ($y_2 < 0$) and $K < 1$ for repulsive ones.

2.2.2 Systems with spin

One can use the bosonization technique for the two degrees of freedom separately, introducing two sets $(\varphi_\uparrow, \theta_\uparrow)$ and $(\varphi_\downarrow, \theta_\downarrow)$, then the kinetic Hamiltonian reads

$$H_{kin} = H_\uparrow + H_\downarrow.$$

The situation changes by introducing interactions. In fact in this case we have two different possibilities for g_4 and g_2 interactions as they can couple particles with aligned or antialigned spin. That is equivalent to write

$$\begin{aligned} H_4 &= \int dx \sum_{r=R,L} \sum_{\sigma=\uparrow,\downarrow} \left[\frac{g_{4\parallel}}{2}\rho_{r,\sigma}(x)\rho_{r,\sigma}(x) + \frac{g_{4\perp}}{2}\rho_{r,\sigma}(x)\rho_{r,-\sigma}(x)\right] \\ H_2 &= \int dx \sum_{\sigma=\uparrow,\downarrow} \Big[g_{2\parallel}\rho_{R,\sigma}(x)\rho_{L,\sigma}(x) + g_{2\perp}\rho_{R,\sigma}(x)\rho_{L,-\sigma}(x)\Big]. \end{aligned} \tag{2.34}$$

By using this formulation the Hamiltonian remains quadratic in φ, θ, but is not diagonal w.r.t. spin. To diagonalize it one can introduce the new densities

$$\begin{aligned} \rho(x) &= \frac{1}{\sqrt{2}}[\rho_\uparrow(x) + \rho_\downarrow(x)] \\ \sigma(x) &= \frac{1}{\sqrt{2}}[\rho_\uparrow(x) - \rho_\downarrow(x)] \end{aligned} \tag{2.35}$$

which of course define new fields, called the charge and spin field respectevely

$$\begin{aligned} \varphi_\rho(x) &= \frac{1}{\sqrt{2}}[\varphi_\uparrow(x) + \varphi_\downarrow(x)] \\ \varphi_\sigma(x) &= \frac{1}{\sqrt{2}}[\varphi_\uparrow(x) - \varphi_\downarrow(x)] \end{aligned} \tag{2.36}$$

and similar for θ_ρ, θ_σ. It is easy to prove that ρ and σ commute and that $(\varphi_\rho, \theta_\rho)$, $(\varphi_\sigma, \theta_\sigma)$ obey standard bosonic commutation relations. The full Hamiltonian now reads

$$\begin{aligned} H = H^0_\rho + H^0_\sigma + \frac{1}{4\pi^2}\int dx[g_{4\|} + g_{4\perp}][(\nabla\varphi_\rho)^2 + (\nabla\theta_\rho)^2] + \\ \frac{1}{4\pi^2}\int dx[g_{4\|} - g_{4\perp}][(\nabla\varphi_\sigma)^2 + (\nabla\theta_\sigma)^2] + \\ \frac{1}{4\pi^2}\int dx[g_{2\|} + g_{2\perp}][(\nabla\varphi_\rho)^2 - (\nabla\theta_\rho)^2] + \\ \frac{1}{4\pi^2}\int dx[g_{2\|} - g_{2\perp}][(\nabla\varphi_\sigma)^2 + (\nabla\theta_\sigma)^2]. \end{aligned} \tag{2.37}$$

Now we need to consider the g_1 process, which is present for systems with spin and needs a bit more care. In fermionic language the interaction term reads as

$$\int dx\ g_{1\|}\sum_{\sigma=\uparrow,\downarrow}\psi^\dagger_{L,\sigma}\psi^\dagger_{R,\sigma}\psi_{L,\sigma}\psi_{R,\sigma} + g_{1\perp}\sum_{\sigma=\pm1}\psi^\dagger_{R,\sigma}\psi^\dagger_{L,-\sigma}\psi_{R,-\sigma}\psi_{L,\sigma}. \tag{2.38}$$

By using anticommutation relations it is easy to see that $g_{1\|}$ has the same form of $g_{2\|}$ (but with a minus sign), so it will be sufficient to substitute $g_{2\|} \to g_{2\|} - g_{1\|}$. On the other hand, for $g_{1\perp}$ one finds (in field representation)

$$\begin{aligned} H_{1\perp} &= \int dx \frac{2g_{1\perp}}{(2\pi\alpha)^2}\sum_{s=\uparrow,\downarrow} e^{-2i\varphi_s(x)}e^{2i\varphi_{-s}(x)} \\ &= \int dx \frac{2g_{1\perp}}{(2\pi\alpha)^2}\cos(2\sqrt{2}\varphi_\sigma(x)). \end{aligned} \tag{2.39}$$

Putting everything together one arrives at the remarkable result of spin-charge complete separation

$$H = H_\rho + H_\sigma \tag{2.40}$$

with (upper sign for ρ, lower for σ)

$$\begin{aligned} u_\nu &= v_F[(1 + y_{4\nu})^2 - (y_{2\nu}/2)^2]^{1/2} \\ K_\nu &= \left[\frac{1 + y_{4\nu} + y_{2\nu}}{1 + y_{4\nu} - y_{2\nu}}\right]^{1/2} \\ g_\nu &= g_{1\|} - g_{2\|} \mp g_{2\perp} \\ g_{4\nu} &= g_{4\|} \pm g_{4\perp} \end{aligned} \tag{2.41}$$

and

$$H_\sigma = H^0_\sigma + \int dx \frac{2g_{1\perp}}{(2\pi\alpha)^2}\cos(2\sqrt{2}\varphi_\sigma(x)). \tag{2.42}$$

2.3 Phenomenological bosonization

While in the previous section we have examined a rigorous approach to bosonization, we will now see a more "*physical*" one. Considering a 1d chain of particles we can define

$$\rho(x) = \sum_i \delta(x - x_i) \tag{2.43}$$

where x_i is the position of the i-th particle. Following Haldane's intuition, we introduce a labeling field $\varphi_l(x)$ which shall be continuous and take definite values at particles positions, e.g. $\varphi_l(x_i) = 2\pi i$ (in one dimension this is always well defined as one can always enumerate the particles from left to right). Now using the rule for the Dirac delta of a function we can rewrite

$$\rho(x) = \sum_n |\nabla\varphi_l(x)| \delta(\varphi_l(x) - 2\pi n) \tag{2.44}$$

and now, using the Poisson resummation formula,

$$\rho(x) = \frac{\nabla\varphi_l(x)}{2\pi} \sum_p e^{ip\varphi_l(x)}. \tag{2.45}$$

It is now convenient to define a new labeling field that takes the difference with the perfect crystalline distribution ρ_0 i.e.

$$\varphi_l(x) = 2\pi\rho_0 x - 2\varphi(x) \tag{2.46}$$

so that now

$$\rho(x) = \left(\rho_0 - \frac{\nabla\varphi(x)}{\pi}\right) \sum_p e^{2ip(\pi\rho_0 x - \varphi(x))} \tag{2.47}$$

which implies that averaging over large distances the only term which contributes is the $p \sim 0$ one

$$\rho_{p\sim 0}(x) = \rho_0 - \frac{\nabla\varphi(x)}{\pi}. \tag{2.48}$$

We now pass to single particle operators by making a reasonable ansatz of the form

$$\psi^\dagger(x) = [\rho(x)]^{1/2} e^{-i\theta(x)} \tag{2.49}$$

where $\theta(x)$ is an operator which still has to be defined. Imposing bosonic commutation relation

$$[\psi(x), \psi^\dagger(x')] = \delta(x - x')$$

and combining it with (2.49) one gets

$$e^{i\theta(x)}[\rho(x)]^{1/2}[\rho(x')]^{1/2}e^{-i\theta(x')} - [\rho(x')]^{1/2}e^{-i\theta(x')}e^{i\theta(x)}[\rho(x)]^{1/2}.$$

Assuming $[\theta(x), \theta(x')] = 0$, a sufficient condition to satisfy the bosonic commutation relation is

$$[\rho(x), e^{-i\theta(x')}] = \delta(x - x')e^{-i\theta(x')}. \tag{2.50}$$

We start by looking at the $p \sim 0$ term (which shall be the most physically relevant) that is satisfied if

$$[\frac{1}{\pi}\nabla\varphi(x), \theta(x')] = -i\delta(x - x') \tag{2.51}$$

which resembles the previous chapter φ, Π relation (2.25). Higher harmonics give contributions as

$$[e^{-2ip\varphi(x)}, e^{-i\theta(x')}] = e^{-2ip\varphi(x)}e^{-i\theta(x')}(1 - e^{ip\pi Sign(x-x')})$$

which always gives zero for $x = x'$ while gives contribution for $x \neq x'$ (odd terms). That being said we shall consider that these have a $e^{-2ip\varphi(x)}$ factor in front, which due to its rapidly oscillating nature makes them negligible in the continuum limit. It is therefore sufficient that $\varphi(x)$ and $\theta(x)$ satisfy (2.51). We can now write the full expression for the bosonic single particle operator

$$\psi^\dagger(x) = \left(\rho_0 - \frac{\nabla\varphi(x)}{\pi}\right)^{1/2} \sum_p e^{2ip(\pi\rho_0 x - \varphi(x))}e^{-i\theta(x)} \tag{2.52}$$

To go from bosonic representation to fermionic one it is sufficient to go from commutation to anticommutation relations. In order to do so one can just multiply the bosonic field for $e^{i\frac{1}{2}\varphi_l(x)}$, which jumps between ±1 for consecutive particles. This is commonly known as the Jordan-Wigner transformation.

This representation and this formulation are richer than what we have obtained in the previous section. It can be proved [15] indeed that the former represents just the $p \sim 0$ term as the higher harmonics are de facto cut by the complete linearization of the spectrum.

Going back to us we now want to find the Hamiltonian of this system. The presence of the term $\int dx\rho(x)^2$ implies a contribution proportional to $(\nabla\varphi(x))^2$. Similar reasons imply the presence of $(\nabla\theta(x))^2$ terms, that being for example a kinetical single particle term as

$$H_K = \int dx \frac{1}{2m}(\nabla\psi^\dagger(x))(\nabla\psi(x))$$

which, if one only keeps the most relevant terms, that is using $\psi^\dagger(x) = \rho_0^{1/2} e^{-i\theta(x)}$, results in

$$H_K = \int dx \frac{\rho_0}{2m} (\nabla\theta)^2. \tag{2.53}$$

It can be proved that $\nabla\varphi\nabla\theta$ are forbidden by inversion symmetry [14].
We can now parameterize the Hamiltonian as

$$H = \frac{1}{2\pi} \int dx [\ uK(\nabla\theta(x))^2 + \frac{u}{K}(\nabla\varphi(x))^2\ ] \tag{2.54}$$

where u, K are parameterize the unknown coefficients.

2.4 The Hubbard term

We now want to include the Hubbard term

$$H_U = U \sum_i n_{i\uparrow} n_{i\downarrow}.$$

Using the field definitions (2.24) (sometimes called *bosonization dictionary*) we introduce two different fields $\varphi_\uparrow(x)$, $\varphi_\downarrow(x)$ for different spins particles and get

$$\begin{aligned} \sum_i n_{i\uparrow} n_{i\downarrow} \to \int dx \Big[-\frac{\nabla\varphi_\uparrow(x)}{\pi} + \frac{1}{2\pi\alpha}(e^{2i(k_F x - \varphi_\uparrow(x))} + \text{h.c.}) \Big] \times \\ \Big[-\frac{\nabla\varphi_\downarrow(x)}{\pi} + \frac{1}{2\pi\alpha}(e^{2i(k_F x - \varphi_\downarrow(x))} + \text{h.c.}) \Big] = \\ \int dx \Big[\frac{1}{\pi^2} \nabla\varphi_\uparrow(x) \nabla\varphi_\downarrow(x) + \frac{2}{(2\pi\alpha)^2} \cos(2\varphi_\uparrow(x) - 2\varphi_\downarrow(x)) \Big] = \\ \Big(\varphi_\rho(x) \equiv \frac{1}{\sqrt{2}}[\varphi_\uparrow(x) + \varphi_\downarrow(x)]\ ,\ \varphi_\sigma(x) \equiv \frac{1}{\sqrt{2}}[\varphi_\uparrow(x), \varphi_\downarrow(x)] \Big) \\ \int dx \Big[\frac{1}{2\pi^2}((\nabla\phi_\rho)^2 - (\nabla\phi_\sigma)^2) + \frac{2}{(2\pi\alpha)^2} \cos(\sqrt{8}\phi_\sigma(x)) \Big] \end{aligned} \tag{2.55}$$

where we have neglected the highly oscillating ($\sim e^{\pm 4ik_F x}$) terms and introduced the charge/spin fields to decouple the kinetical terms. The full Hamiltonian therefore reads

$$\begin{aligned} H = \frac{1}{2\pi}\int dx \Big[& u_\rho K_\rho (\pi \Pi_\rho(x))^2 + \frac{u_\rho}{K_\rho}(\nabla\varphi_\rho(x))^2 + \\ & u_\sigma K_\sigma (\pi \Pi_\sigma(x))^2 + \frac{u_\sigma}{K_\sigma}(\nabla\varphi_\sigma(x))^2 + \\ & \frac{2U}{2\pi\alpha^2}\cos(\sqrt{8}\phi_\sigma(x)) \Big] \end{aligned} \tag{2.56}$$

where

$$\begin{aligned} & u_\rho K_\rho = u_\sigma K_\sigma = v_F \\ & \frac{u_\rho}{K_\rho} = v_F \\ & \frac{u_\sigma}{K_\sigma} = v_F \left(1 - \frac{U}{\pi v_F}\right). \end{aligned} \tag{2.57}$$

This is the well-known and studied Sine-Gordon Hamiltonian [25]. It is useful to review some general aspects of it, in particular some features of its Renormalization Group.

2.4.1 Sine-Gordon Hamiltonian

Following [25] let's consider the Hamiltonian

$$H[\Pi,\phi] = \int dx \left[\frac{v}{2}\Big(\Pi^2 + (\partial_x\phi)^2\Big) - \tilde{g}\cos(\beta\phi)\right] \tag{2.58}$$

so that after Wick rotating $t \to -it$ the action reads

$$\begin{aligned} & S[\phi] \equiv S_0[\phi] + S_I[\phi] \\ & S_0[\phi] = \int dx\ \frac{1}{2}\phi\nabla_x^2\phi \quad , \ \nabla_x^2 = \partial_x^2 + \frac{1}{v^2}\partial_t^2 \\ & S_I[\phi] = \int dx\ g\cos(\beta\phi) \quad , \ g = \frac{\tilde{g}}{v} \end{aligned} \tag{2.59}$$

where $x \to (x, vt)$. The general procedure for Wilsonian RG is to decompose $\phi(x)$

$$\begin{aligned} \phi(x) = & \int_{|q|<\Lambda} \frac{dq}{(2\pi)^2}\ \tilde{\phi}(q)\ e^{iqx} = \\ & \int_{|q|<\Lambda/s} \frac{dq}{(2\pi)^2}\ \tilde{\phi}(q)e^{iqx} + \int_{\Lambda/s<|q|<\Lambda} \frac{dq}{(2\pi)^2}\ \tilde{\phi}(q)\ e^{iqx} = \\ \equiv\ & \phi^S(x) + \delta\phi(x), \end{aligned} \tag{2.60}$$

where Λ is some cutoff, $s \simeq 1$ and the field is divided into the *bulk* and *shell* components or *slow* and *fast* modes. One can prove that the free part of the action S_0 decouples

$$S_0[\phi] = S_0[\phi^S] + S_0[\delta\phi]$$

while this is false for the interacting term. We, therefore, decide to treat the interaction term in a perturbative way by expanding in $\delta\phi \ll \phi^S$. One finds:

$$S_I[\phi] \simeq S_I[\phi^S] + \int dx\ a^S(x)\ \delta\phi(x) + \int dx dx'\ \delta\phi(x)\ b^S(x,x')\ \delta\phi(x')$$

where

$$\begin{aligned} a^S(x) &= \frac{\delta l_I[\phi]}{\delta\phi(x)}\Big|_{\phi^S(x)} \qquad l_I[\phi] = g\cos(\beta\phi) \\ b^S(x,x') &= \frac{1}{2}\frac{\delta^2 l_I[\phi]}{\delta\phi(x)\delta\phi(x')}\Big|_{\phi^S(x)} \end{aligned} \tag{2.61}$$

so that by using the expanded $S_I[\phi$ one has

$$S[\phi] = S[\phi^S] + \delta S[\phi^S, \delta\phi]$$

with

$$\begin{aligned} \delta S[\phi^S, \delta\phi] = &\ S_0[\delta\phi] + \int dx\ a^S(x)\delta\phi(x)\ + \int dx dx'\ \delta\phi(x) b^S(x,x')\delta\phi(x') = \\ &\int dx dx'\ \delta\phi(x)\Big[\ \delta(x-x')\ \frac{1}{2}\ \nabla^2_{x'} + b^S(x,x')\Big]\ \delta\phi(x') + \int dx\ a^S(x)\delta\phi(x). \end{aligned} \tag{2.62}$$

We will now derive an effective action for the slow modes by averaging $\delta S[\phi^S, \delta\phi]$ w.r.t. the unperturbed ground state of the fast modes $\delta\phi(x)$. The RG procedure consists indeed in re-obtaining $S[\phi]$ from this average $S_{eff}[\phi^S]$ through a scale renormalization $\Lambda \to \Lambda/s$. First, we derive the unperturbed Green function $G_0(x,x')$

$$\begin{cases} G_0^{-1}(x,x') = \delta(x-x')\ \frac{1}{2}\ \nabla^2_{x'} \\ \int dx'' G_0^{-1}(x''-x')\ G_0(x''-x) = \delta(x-x') \end{cases} \quad \to\ G_0(q,\omega) = -\frac{2}{q^2 + \omega^2/v^2} \tag{2.63}$$

where

$$G_0(x-x') = \int \frac{dq}{(2\pi)^2}\ G_0(q)\ e^{iq(x-x')} \qquad \text{with} \quad x = (x, vt) \quad q = (q, \omega).$$

Now the full Green function is defined through [25]

$$G^{-1}(x-x') = G_0^{-1}(x-x') - \Sigma(x,x') \tag{2.64}$$

where Σ is the self-energy which in our case is $\Sigma(x,x') = b^S(x,x')$. Inserting (2.64) into the relation which defines its inverse, which is

$$\int dx''\; G^{-1}(x''-x')G(x''-x) = \frac{1}{(2\pi)^2}\delta(x-x'),$$

one gets

$$\begin{aligned} &\int \frac{dq}{(2\pi)^2}\frac{dq'}{(2\pi)^2}\; G(q,q')\;\frac{1}{2}\Big(-q^2-\frac{\omega^2}{v^2}\Big)\; e^{i(qx+q'x')} - \\ \int dx'' \int \frac{dq}{(2\pi)^2}\frac{dq'}{(2\pi)^2}\frac{dk}{(2\pi)^2}\frac{dk'}{(2\pi)^2}\;&\Sigma(q,k)G(q',k')\; e^{iqx+i(k+k')x''+iq'x'} = \int \frac{dq}{(2\pi)^2}e^{iq(x-x')} \end{aligned}$$

$$\begin{aligned} \int \frac{dq}{(2\pi)^2}\frac{dq'}{(2\pi)^2}\Big[G(q,q')G_0^{-1}(q) - \int \frac{dq''}{(2\pi)^2}\Sigma(q,q'')G(-q'',q')\Big]e^{iqx+iq'x'} &= \\ \int \frac{dq}{(2\pi)^2}\frac{dq'}{(2\pi)^2}e^{iqx+iq'x'}dq. \end{aligned}$$

This leads to

$$G(q,q') = G_0(q)\delta(q+q') + G_0(q)\int \frac{dq''}{(2\pi)^2}\;\Sigma(q,q'')G(-q'',q')$$

which is the well-known *Dyson equation*. This is solved iteratively as

$$G^{(0)}(q,q') = G_0(q)\delta(q+q')$$

$$\begin{aligned} G^{(1)}(q,q') &= G_0(q)\delta(q+q') - G_0(q)\;\int \frac{dq''}{(2\pi)^2}\;b^S(q,q'')G^{(0)}(q'',q) \\ &= G_0(q)\delta(q+q') - \frac{1}{(2\pi)^2}\;G_0(q)\;b^S(q,q')\;G_0(-q') \end{aligned} \tag{2.65}$$

$$\begin{aligned} G^{(2)}(q,q') &= G_0(q)\delta(q+q') - G_0(q)\;\int \frac{dq''}{(2\pi)^2}\;b^S(q,q'')G^{(1)}(q'',q) \\ &= \ldots \end{aligned}$$

Turning back to the action

$$\delta S[\phi^S, \delta\phi] = \int dx dx' \; \delta\phi(x) \; G^{-1}(x,x') \; \delta\phi(x') + \int dx \; a^S(x)\delta\phi(x)$$

we see that to average we need it to be quadratic. We rewrite

$$\delta\phi(x) = \bar{\phi}(x) + r(x) \tag{2.66}$$

so that

$$\begin{aligned}\delta S[\phi^S, \delta\phi] = & \int dx dx' \; (\bar{\phi}(x) + r(x)) \; G^{-1}(x,x') \; (\bar{\phi}(x') + r(x')) + \\ & \int dx \; a^S(x)(\bar{\phi}(x) + r(x)) = \\ = & \int dx dx' \; \bar{\phi}(x) G^{-1}(x,x') \bar{\phi}(x') \; + \\ & \int dx dx' \; \bar{\phi}(x) \Big[2G^{-1}(x,x') \; r(x') + a^S(x')\delta(x'-x) \Big] \; + \\ & \int dx dx' \; r(x) \Big[G^{-1}(x,x') \; r(x') + a^S(x')\delta(x-x') \Big].\end{aligned} \tag{2.67}$$

Assuming now

$$\int dx' G^{-1}(x,x') \; r(x') = \; -\frac{1}{2} a^S(x) \longrightarrow r(x) = -\frac{1}{2} \int dx' \; G(x,x') a^S(x) \tag{2.68}$$

one gets

$$\delta S[\phi^S, \bar{\phi}] = \int dx dx' \Big(\bar{\phi}(x) G^{-1}(x,x') \bar{\phi}(x') - \frac{1}{4} a^S(x) G(x,x') a^S(x') \Big) \tag{2.69}$$

which is ready to be averaged. As previously done, one can again find

$$H_0[\Pi, \phi] = H_0[\Pi^S, \phi^S] + H_0[\bar{\Pi}, \bar{\phi}] \tag{2.70}$$

which implies

$$\left|0\right\rangle_\phi = \left|0\right\rangle_{\phi^S} \; \left|0\right\rangle_{\bar{\phi}} . \tag{2.71}$$

We now want to re-obtain $G_0(x,x') = {}_\phi\left\langle 0\right| \phi(x)\phi(x') \left|0\right\rangle_\phi$ in the fast mode subspace. One finds

$$\begin{aligned}G_0(x,x') = & {}_\phi\left\langle 0\right| \phi^S(x)\phi^S(x') \left|0\right\rangle_\phi + {}_\phi\left\langle 0\right| \phi^S(x)\bar{\phi}(x') \left|0\right\rangle_\phi \\ & {}_\phi\left\langle 0\right| \phi^S(x)\bar{\phi}(x') \left|0\right\rangle_\phi + {}_\phi\left\langle 0\right| \bar{\phi}(x)\bar{\phi}(x') \left|0\right\rangle_\phi .\end{aligned}$$

The second and fourth terms vanish because they act on different subspaces, while scale invariance implies that the first and third terms are equal. That, combined with the fact that ${}_{\phi^S}\langle 0|\,|0\rangle_{\phi^S}$ brings

$$G_0(x,x') = 2_{\bar{\phi}}\langle 0|\,\bar{\phi}(x)\bar{\phi}(x')\,|0\rangle_{\bar{\phi}}$$

so that we can write

$$G_0(x,x') = G_0(x-x') = \int_{shell} \frac{dq}{(2\pi)^2}\; G_0(q)e^{iq(x-x')}.$$

The effective action is found by averaging over $\bar{\phi}$

$$\delta S[\phi^S] = {}_{\bar{\phi}}\langle 0|\,\delta S[\phi^S,\bar{\phi}]\,|0\rangle_{\bar{\phi}}$$

$$\delta S[\phi^S] = \int dx dx' \left[\frac{1}{2}G_0(x,x')G^{-1}(x,x') - \frac{1}{4}a^S(x)G(x,x')a^S(x')\right].$$

Keeping terms up to the 2nd order in g, in the Sine-Gordon case, the computation reduces to

$$\begin{aligned} F_1[\phi^S] &= \frac{1}{2}\int dx dx'\; G_0(x,x')b^S(x,x') \\ F_2[\phi^S] &= -\frac{1}{4}\int dx dx'\; G_0(x,x')a^S(x)a^S(x'). \end{aligned} \tag{2.72}$$

By working everything out (see [25] for more details) we finally come to

$$S_{eff}[\phi^S] = \int dx \left[\left(1 + \frac{3\;dl\;\beta^4 g^2}{4\pi\Lambda^3}\right)\frac{1}{2}\phi^S\nabla_x^2\phi^S + g\left(1 - \frac{dl\;\beta^2}{4\pi}\right)\cos(\beta\phi^S)\right] \tag{2.73}$$

where $dl = -\ln s$. The RG flow therefore is

$$\beta_R^{-2} = \beta^{-2}\left(1 + \frac{3\;dl\;\beta^4 g^2}{4\pi\Lambda^3}\right) \qquad g_R = gs^{-2}\left(1 - \frac{dl\;\beta^2}{4\pi}\right). \tag{2.74}$$

One can now define the differentials $d\beta^2 \equiv \beta_R^2 - \beta^2$ and $dg \equiv g - g_R$ and get

$$\frac{d\beta^2}{dl} = -\frac{3\beta^6 g^2}{4\pi\Lambda^3} \qquad \frac{dg}{dl} = 2g\left(1 - \frac{\beta^2}{8\pi}\right). \tag{2.75}$$

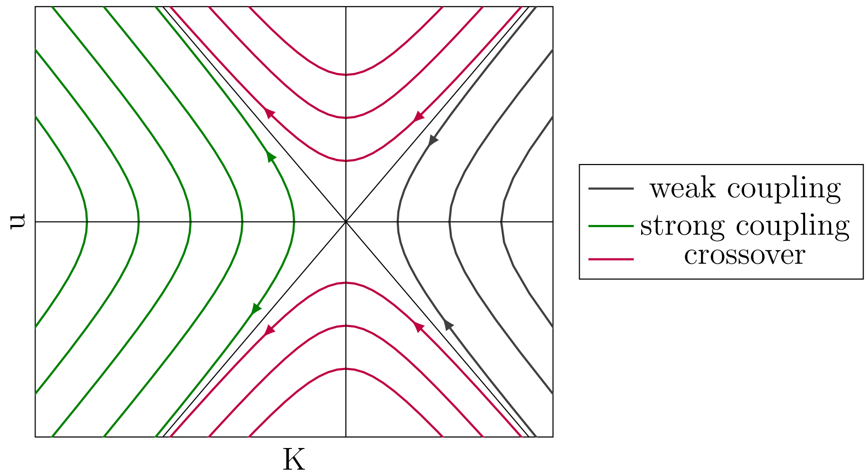
Finally defining $K = \frac{\beta^2}{8\pi}$ and $u = 4g\sqrt{\frac{3\pi}{\Lambda^3}}$

$$\begin{cases} \frac{dK}{dl} = 2u(1-K) \\ \frac{du}{dl} = -u^2K^3. \end{cases} \tag{2.76}$$

From the last equations, it is clear that $K = 1$ is a critical point where physics changes dramatically. We can prove that the trajectories in the $u-K$ space are indeed hyperbolas as

$$c = u^2 - 2(K-1)^2$$

is invariant. In particular for $c < 0$ the RG flow has two fixed points $(K_f, u_f) = \left(1 \pm \sqrt{\frac{-c}{2}}, 0\right)$, for $c = 0$ there are two lines with one fixed point $(1, 0)$ while for $c > 0$ there are no fixed points. The latter is called *the crossover regime* while the former is divided into the *weak coupling* regime (i.e. $K > 1$) where the fixed points are stable and the *strong coupling* regime, where the fixed points are unstable.
We see that both in the crossover and strong coupling regime the system flows towards large values of u. It is thus reasonable to assume that at some scale l_c (a good estimate is considering when $|u|$ reaches unity) the system encounters a phase transition. Recalling (2.57), the charge sector has trivial solution $u = v_F$, $K = 1$ [1] while for the spin sector one has that for $U > 0$ (i.e. repulsive regime) the $\cos(\sqrt{8}\phi_\sigma)$ is always marginal, so the spin sector is always massless.



---

[1] To be precise we recall that at half filling one should also include the umklapp term which makes the charge sector less trivial. As in the following, we will consider $\rho = 2/5$ we do not include this feature here and refer to [15] for more details.

2.5 The soft-shoulder potential

We now pass to analyze the soft-shoulder potential

$$\begin{aligned} H_V =& V\sum_i\sum_{l=1}^{r_c} n_i n_{i+l} = \\ & V\sum_i\sum_{l=1}^{r_c}(n_{i\uparrow}+n_{i\downarrow})(n_{i+l\uparrow}+n_{i+l\downarrow}) = \\ & V\sum_i\sum_{l=1}^{r_c}\Big[(n_{i\uparrow}n_{i+l\uparrow}+n_{i\downarrow}n_{i+l\downarrow})+(n_{i\uparrow}n_{i+l\downarrow}+n_{i\uparrow}n_{i+l\downarrow})\Big]. \end{aligned} \tag{2.77}$$

By considering it term by term (we indicate the lattice spacing with a), first we have:

$$\begin{aligned} \sum_i\sum_{l=1}^{r_c} n_{i\uparrow}n_{i+l\uparrow} \to \int dx \sum_{l=1}^{r_c}\Big[-\frac{\nabla\varphi_\uparrow(x)}{\pi}+\frac{1}{2\pi\alpha}(e^{2i(k_F x-\varphi_\uparrow(x))}+\text{h.c.})\Big]\times \\ \Big[-\frac{\nabla\varphi_\uparrow(x+la)}{\pi}+\frac{1}{2\pi\alpha}(e^{2i(k_F(x+la)-\varphi_\uparrow(x+la))}+\text{h.c.})\Big] = \\ \int dx\sum_{l=1}^{r_c}\left[\frac{1}{\pi^2}\nabla\varphi_\uparrow(x)\nabla\varphi_\uparrow(x+la)+\frac{2}{(2\pi\alpha)^2}\cos(2k_F la-2(\varphi_\uparrow(x+la)-\varphi_\uparrow(x)))\right] \end{aligned}$$

where we have eliminated $e^{2ik_F(2x+la)}$ terms as they are highly oscillating. It is the same for the $\varphi_\downarrow(x)$ field. The other two terms instead lead to

$$\int dx\sum_{l=1}^{r_c}\left[\frac{1}{\pi^2}\nabla\varphi_\uparrow(x)\nabla\varphi_\downarrow(x+la)+\frac{2}{(2\pi\alpha)^2}\cos(2k_F la-2(\varphi_\uparrow(x+la)-\varphi_\downarrow(x)))\right].$$

Using now basic goniometric identities and the fact that $\sin(\varphi_\uparrow(x+la)-\varphi_\uparrow(x))\simeq 0$ we find ourselves with a term $\cos(k_F la)\cos(2\varphi_\uparrow(x+la)-2\varphi_\uparrow(x))$. It is now reasonable to expand the cosine, but this has to be done with some care. Because in a normal ordered operator destruction operators are on the right, by a simple expansion one can find

$$\langle :\cos\varphi:\rangle = 1.$$

Then from direct expansion, one can see that

$$\cos(\varphi) = \langle\cos(\varphi)\rangle : \cos(\varphi) :$$

which leads to

$$e^{2i(\varphi_\uparrow(x)-\varphi_\uparrow(x+la))} =: e^{2i(\varphi_\uparrow(x)-\varphi_\uparrow(x+la))} : e^{-\frac{1}{2}\langle[2(\varphi_\uparrow(x)-\varphi_\uparrow(x+la))]^2\rangle}.$$

The last expression thus becomes at the lowest order in V [2]

$$e^{2i(\varphi_\uparrow(x)-\varphi_\uparrow(x+la))} =: e^{2i(\varphi_\uparrow(x)-\varphi_\uparrow(x+la))} : \frac{\alpha^2}{(la)^2}. \tag{2.78}$$

Finally in the normal ordered operator, one can expand $\varphi_\uparrow(x+la)-\varphi_\uparrow(x) \simeq (la)\nabla\varphi(x)$ and that leads to a contribution to the quadratic part of the charge Luttinger Hamiltonian, that is

$$\frac{V}{\pi^2}\int dx\ (\nabla\phi_\uparrow(x))^2 \sum_{l=1}^{r_c}[1+2\cos(k_F la)] \tag{2.79}$$

combining the ↑, ↓ components

$$(\nabla\varphi_\uparrow(x))^2 + (\nabla\varphi_\uparrow(x))^2 = (\nabla\varphi_\rho(x))^2 + (\nabla\varphi_\sigma(x))^2$$

that leads to corrections

$$\begin{aligned} \frac{u_\rho}{K_\rho} &= v_F\left(1+\frac{V\sum_{l=1}^{r_c}(1-2\cos(k_F la)}{\pi v_F}\right) \\ \frac{u_\sigma}{K_\sigma} &= v_F\left(1-\frac{U-V\sum_{l=1}^{r_c}(1+2\cos(k_F la)}{\pi v_F}\right) \end{aligned} \tag{2.80}$$

where in the spin term we considered the previously listed corrections due to the on-site potential (see eq (2.57)). Coming to the other terms, in this case, it is easier to compute the $\cos(\sqrt{8}\phi_\sigma(x))$ coefficient and then use spin rotation symmetry (this is due to the fact that $\sin(2\varphi_\uparrow(x+la)-2\varphi_\downarrow(x)) \neq 0$). In fact by looking at the cosine, in this case, the argument the difference between $\varphi_\uparrow(x)$ and $\varphi_\uparrow(x+la)$ or vice-versa. By expanding again and keeping only the most relevant terms that clearly becomes $\cos(\sqrt{8}\varphi_\sigma(x))$. That implies a new contribution to the *Sine-Gordon term* coefficient, which is the backscattering $g_{1\perp}$ in the g-ology theory. The *Sine-Gordon term* is thus now

$$H_{SG} = \frac{2}{(2\pi\alpha)^2}\left(U+2V\sum_{l=1}^{r_c}\cos(k_F la)\right)\int dx\ \cos(\sqrt{8}\varphi_\sigma(x)) \tag{2.81}$$

which means

---

[2] The expression for $e^{-\frac{1}{2}\langle[2(\varphi_\uparrow(x)-\varphi_\uparrow(x+la))]^2\rangle}$ at the lowest order in V can be found in [15]

$$g_{1\perp} = U + 2V\sum_{l=1}^{r_c}\cos(2k_F la). \tag{2.82}$$

That implies

$$\begin{aligned}\frac{u_\rho}{K_\rho} &= v_F\left(1 + \frac{V(1+\sum_{l=1}^{r_c}(1-2\cos(k_F la))}{\pi v_F}\right)\\ \frac{u_\sigma}{K_\sigma} &= v_F\left(1 - \frac{U + 2V\sum_{l=1}^{r_c}\cos(2k_F la)}{\pi v_F}\right).\end{aligned} \tag{2.83}$$

Let us end this chapter by underlying the differences between the U and V potentials. While in the first case we did not make any assumption or expansion, in the latter we only keep the leading terms in V. We, therefore, make the hypothesis that the Luttinger Liquid theory holds for every value of U while breaks down for finite values of V. In fact, for some values of V, the RG flow moves to the strong coupling region. Solving as before for the spin and charge sectors we have

$$\begin{aligned}K_\rho &= \left(1 + \frac{V(1+\sum_{l=1}^{r_c}(1+2\cos(k_F la))}{\pi v_F}\right)^{-1/2}\\ K_\sigma &= \left(1 - \frac{U + 2V\sum_{l=1}^{r_c}\cos(2k_F la)}{\pi v_F}\right)^{-1/2}.\end{aligned} \tag{2.84}$$

As for $\rho = \frac{2}{5}$, $k_F = \frac{2}{5a}\pi$, and taking $r_c = 2$ one finds

$$\begin{aligned}K_\rho &= \left(1 + \frac{V\big(3 + 2\cos(2\pi/5) + 2\cos(4\pi/5)\big)}{\pi v_F}\right)^{-1/2}\\ &= \left(1 + \frac{V}{\pi v_F}\right)^{-1/2}.\end{aligned} \tag{2.85}$$

which is always less than one for positive values of V. By looking now at the spin sector the condition $K_\sigma < 1$ which is reached if

$$\begin{aligned}&U{+}2V\sum_{l=1}^{r_c}\cos(2k_F la) < 0\\ &U{+}2V\left[\cos\left(\frac{2\pi}{5}\right) + \cos\left(\frac{4\pi}{5}\right)\right] < 0\\ &\quad\implies V > U.\end{aligned} \tag{2.86}$$

Of course, this cannot be exact as for finite values V other terms shall emerge which can modify the picture, but shall work for small values of U, V.

Chapter 3

The extended Hubbard model with soft-shoulder potential: a review

"In a dark place we find ourselves and a little more knowledge lights our way."

Yoda - Star Wars Episode III:
Revenge Of The Sith

In this chapter, we will first examine the classical version of the extended Hubbard model. That is the Hamiltonian (2.1) with $t = 0$. In fact the previously described hopping term is the only quantistic element of the model as it introduces tunnel effect, while the on-site Hubbard potential and the soft-shoulder potential are indeed classical. The classical Hamiltonian therefore reads

$$H = U \sum_i n_{i\uparrow} n_{i\downarrow} + V \sum_i \sum_{l=1}^{r_c} n_i n_{i+l}. \tag{3.1}$$

Once we will have examined the classical limit we will move on to examine the liquid of clusters phase first proposed in [27], then we will see the proposal for the extended Hubbard model made in [6]. Finally we will study strong coupling perturbation theory, starting with the $V = 0$, $U \gg t$ case and then looking at $U \gg V \gg t$.

3.1 The classical limit

A key implicit parameter of the model is the density ρ. Taking the classical limit as a guide in fact, defining $r^* = 1/\rho$ (which is the average distance between particles) we

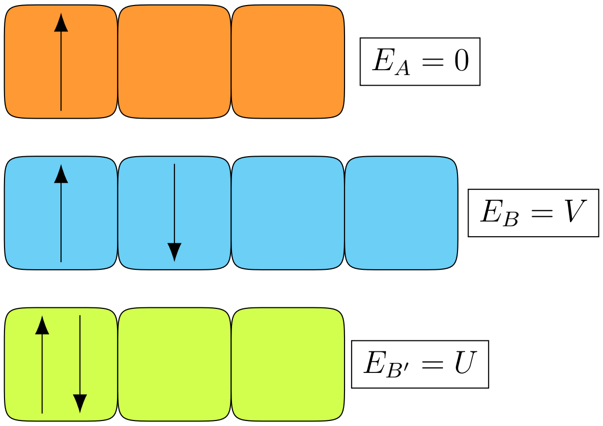


Figure 3.1: Different blocks for $r_c = 2$

can distinguish three different regimes created by the competition between it and the soft-shoulder potential range r_c:

(i) For $r^* > r_c + 1$, particle density is small, the system is in a "liquid-like" phase

(ii) For $r^* = r_c + 1$, particles are exactly distributed (one every $r_c + 1$ sites) in a crystalline order

(iii) For $r^* < r_c$, the competition between r^* and r_c leads to the formation of clusters.

The last regime is the most interesting and can tackle in an analytical way as in [27]. We define block configuration as follows (see Fig. 3.1):

- Block of type A has a length of $r_c + 1$ sites and contains only one particle (with spin $\uparrow$ or $\downarrow$); that corresponds to a classical energy $E_A = 0$.
- Block of type B has a length of $r_c + 2$ sites and contains two particles (which can have any spin configuration) located in nearest neighbours sites; that corresponds to a classical energy $E_B = V$.
- Block of type B' has a length of $r_c + 1$ sites and contains two particles located in the same site (so the two particles must have different spins); that corresponds to a classical energy $E_{B'} = U$.

If $U \gg V$ the system organizes in blocks A and B. Using the same notation of [6] we call this Cluster Luttinger Liquid of nearest neighbors (CLL_{nn}) phase. In this phase the number of blocks is

$$M_{nn} = \frac{L - N}{r_c}. \tag{3.2}$$

As one reduces U, there is a phase transition at $V = \frac{r_c}{r_c+1}U$ where $(r_c + 1)$ B blocks turn into r_c B' and 2 A blocks. This is called the Cluster Luttinger Liquid of doublons (CLL_d) phase and has a number of blocks equal to

$$M_d = \frac{L}{r_c + 1}. \tag{3.3}$$

3.2 Liquid of clusters

Starting from the classical ground state, in [27] the formalism we developed in Sect. 2.3 was used to capture the effect of the hopping term at intermediate value of the couplings U, V (actually in [27] there was no U term and fermions were spinless, but that concept can easily be generalized by introducing B' blocks). In particular the hypothesis made by the authors was that the system would still had liquid-lie features, the difference being in the elementary enetities of the Luttinger liquid, which shall now be clusters and not particles. This is based on two assumptions: (i) the number of clusters is fixed by the classical ground state and (ii) the internal cluster degree of freedom is frozen. We can therefore rewrite the density as

$$\rho(x) = \sum_{m=1}^{M} f(x_m)\delta(x - x_m) \tag{3.4}$$

where M is the total number of clusters, x_m is the position of the m-th cluster and $f(x_m) \in \{1, 2\}$ represents the m-th cluster size. We can now proceed in the same way as before by introducing the collective cluster counting field $\varphi_{cl}(x)$ and finally getting to

$$\rho(x) = \left[\rho_0 - \frac{N}{M\pi}\nabla\varphi_{cl}(x)\right] \sum_p a_p e^{2ip(\pi\rho_0 \frac{M}{N} x - \phi_{cl}(x))} \tag{3.5}$$

with a_p constants. From here one gets a Hamiltonian with an identical functional form but has evident differences in some observables, such as the Fourier transform of the denisty-density correlation function $G_2(l - j)$

$$S(k) = \frac{1}{L}\sum_{l,j} e^{ik(l-j)} G_2(l - j) \quad \text{where} \quad G_2(l - j) = \langle n_l n_j \rangle - \langle n_l \rangle \langle n_j \rangle.$$

In our case the two possibilities would be liquid of clusters phases with block structures identical to the different ground state ones. We shall call the phase with blocks B and A (i.e. for $U > 3V/2$ in the classical limit) the Cluster Luttinger Liquid of nearest neighbours (CLL_{nn}), while we will refer to the phase with blocks B', A' as Cluster Luttinger Liquid of doublons (CLL_d). The two phases are represented in Fig. 3.2.

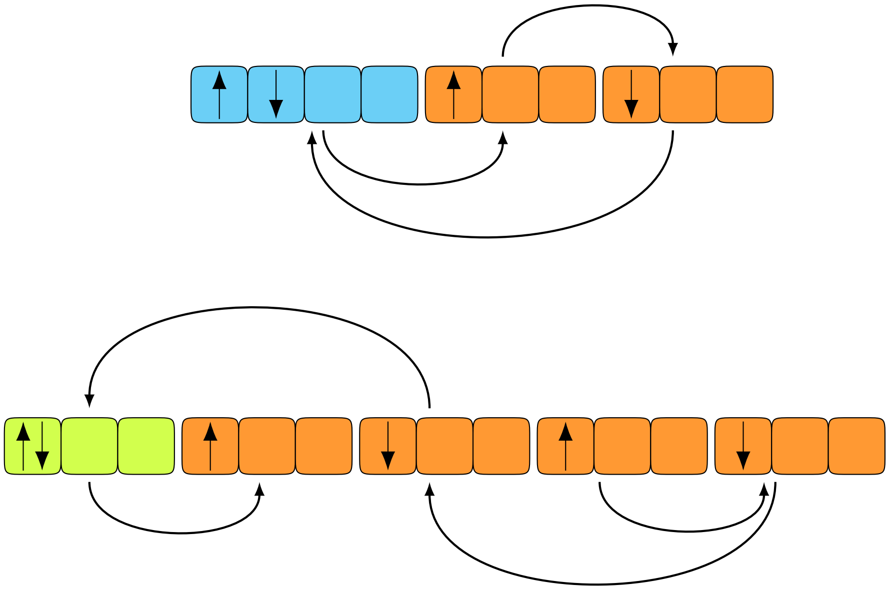


Figure 3.2: Upper panel: representation of Cluster Luttinger Liquid of nearest neighbours (CLL_{nn}). Lower panel: representation of Cluster Luttinger Liquid of doublons (CLL_d). The arrows and colors indicate that the elemetary units of the system are now the clusters.

3.3 Phase diagram proposal

The model under our examination was previously studied in [6], in which the authors used the knowledge of the classical ground state and the previously explained liquid of cluster formalism in order to build a proposal for the phase diagram of the extended Hubbard model. Using Density Matrix Renormalization Group (DMRG) they numerically calculated various observables (charge and spin structure factor, double occupancy number) for different values of U and V (rigorously talking they varied U/t and V/t, in the following we are going to assume $t \equiv 1$ when dealing with the phase diagram). We are not going to examine here the precise working of DMRG and the significance of the observables as we will explain these in the following chapters. The sketch of the phase diagram is depicted in Fig. 3.3. Starting from low values of U (and arbitrary values of V) the first phase predicted by the authors was Tomonaga Luttinger Liquid of doublons (TLL_d). That corresponds to liquid phase whose elementary elements are not particles, but doublons (i.e. doubly occupied sites). Increasing U, doubly occupied sites are less energetically favourable and, for low values of V, i.e. $V \lesssim 5$, the system undergoes a transition to simple Tomonaga Luttinger Liquid (TLL). As V increases the system un-

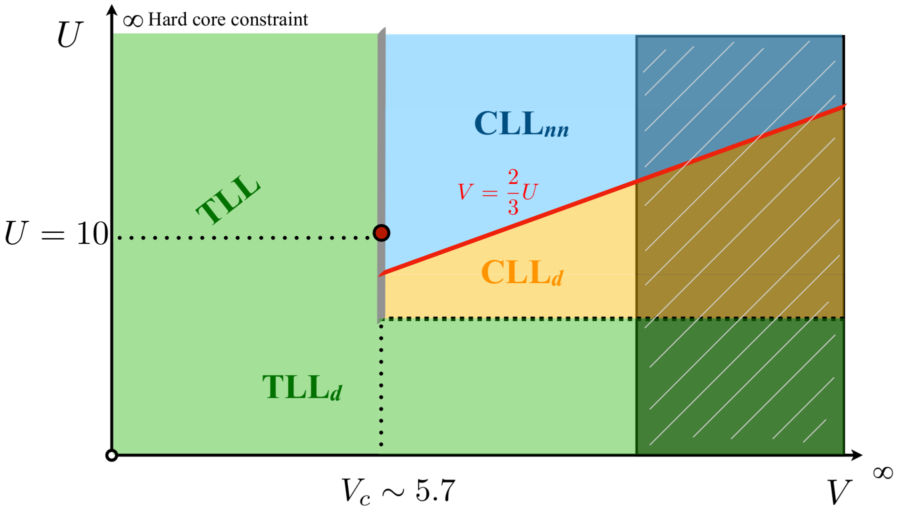


Figure 3.3: Sktech of phase diagram of the extended Hubbard model taken from [6]. The transition from Tomonaga Luttinger Liquid (TLL) to Cluster Luttinger Liquid of nearest neighbours (CLL_{nn}) was well studied for $U = 10$, finding the approximate value of $V_c \simeq 5.7$. The shaded area corresponding to $V \gtrsim 7$ corresponds to an area where numerical results are particularly hard to extract.

dergoes a new phase transition to Cluster Luttinger Liquid (CLL). That can be of two kinds, which are Cluster Luttinger Liquid of nearest neighbours (CLL_{nn}) and Cluster Luttinger Liquid of doublons (CLL_d). The difference between the two is the nature of the elementary entities which compose the system, which are A and B blocks in the former case and A and B' in the latter. The two liquid of clusters phases are divided by the classical phase transition line $U = 3V/2$. Phase transitions from TLL/TLL_d to CLL_d were not studied in detail, while the transition point from TLL to CLL_{nn} was numerically determined for $U = 10$ at $V \simeq 5.7$.

3.4 Perturbation theory $V = 0$, $U \gg t$

By following what was done in [1] we start by dividing the Hamiltonian into

$$H = -t \sum_i \left(c^\dagger_{i+1} c_i + \text{h.c.} \right) + U \sum_i n_{i\uparrow} n_{i\downarrow} \equiv H_K + H_U \tag{3.6}$$

and by defining

$$\Pi(x) \equiv \prod_{i=1}^{L}[1-(1-x)\nu_i] \qquad \nu_i \equiv n_{i\uparrow}n_{i\downarrow}; \quad 0 \leq x \leq 1 \tag{3.7}$$

whose expansion is w.r.t x is

$$\Pi(x) = \sum_{l=1}^{M} x^l P_l$$

where P_l is the projector onto the subspace $\mathcal{H}_l$ of the full Hilbert space $\mathcal{H}$ with l doubly occupied sites. By direct expansion, one finds

$$P_l = \sum_{(i_1 \ldots i_l)} \left[\nu_{i_1} \ldots \nu_{i_l} \prod_j{}' (1-\nu_j)\right] \tag{3.8}$$

where the sum is intended over all possible distinct sets of l sites and $\prod_j'$ stands, for any given set, for a product over all the sites j which are not in the considered set. In particular

$$P_0 = \prod_{i=1}^{L}(1-\nu_i)$$

projects onto the subspace containing no doubly occupied states and is known as the *"Gutzwiller projector"*. Defining

$$P_\eta \equiv \sum_{l>0} P_l$$

this will project onto the subspace containing at least one doubly occupied site. Of course, has

$$P_0 + P_\eta = 1. \tag{3.9}$$

We now want to make a distinction between diagonal and off-diagonal terms in the Hubbard Hamiltonian. To do so we use the trivial identity

$$c_{i\sigma} = c_{i\sigma}[(1-n_{i\bar{\sigma}}) + n_{i\bar{\sigma}}] \tag{3.10}$$

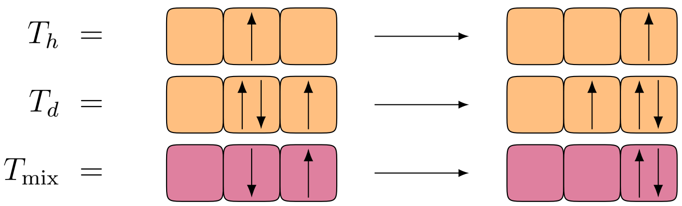
to rewrite

$$H_0 = T_h + T_d + T_{\text{mix}} \tag{3.11}$$

where

$$\begin{aligned} T_h &= -t\sum_{i\sigma} (1-n_{i+1\bar{\sigma}})c^\dagger_{i+1\sigma}c_{i\sigma}(1-n_{i\bar{\sigma}}) \quad \bar{\sigma} = -\sigma \\ T_d &= -t\sum_{i\sigma} n_{i+1\bar{\sigma}}c^\dagger_{i+1\sigma}c_{i\sigma}n_{i\bar{\sigma}} \\ T_{\text{mix}} &= -t\sum_{i\sigma} n_{i+1\bar{\sigma}}c^\dagger_{i+1\sigma}c_{i\sigma}(1-n_{i\bar{\sigma}}). \end{aligned} \tag{3.12}$$

We can give a pictorial representation of T_h, T_d and T_{mix}



It is clear that T_h and T_d commute with H_V while T_{mix} mixes different eigenspaces (in the above picture the different colors, orange for diagonal terms and red for non-diagonal terms). One can convince himself that

$$\begin{aligned} &P_0 H_K P_0 = P_0 T_h P_0 \\ &P_\eta H_K P_\eta = P_\eta T_h P_\eta + T_d \\ &P_\eta H_K P_0 + P_0 H_K P_\eta = T_{\mathrm{mix}}. \end{aligned} \tag{3.13}$$

We define now

$$H = \tilde{H}_K + H_\eta \tag{3.14}$$

where

$$\begin{aligned} &\tilde{H}_K \equiv P_0 H_K P_0 + P_\eta H_K P_\eta + V \\ &H_\eta \equiv P_0 H_K P_\eta + P_\eta H_K P_0 \end{aligned}$$

We now look for a unitary transformation that eliminates (at the lowest order) the effect of non-diagonal terms to obtain an effective Hamiltonian H_{eff} which is completely diagonal, that is

$$P_0 H_{\mathrm{eff}} P_\eta = 0 \tag{3.15}$$

to the required order. We define

$$\begin{aligned} &H(\varepsilon) = \tilde{H}_K + \varepsilon H_\eta \\ &U(\varepsilon) = \exp(i\varepsilon S) \quad \text{with} \quad S = S^\dagger. \end{aligned} \tag{3.16}$$

The previous condition now reads

$$H_{\mathrm{eff}} = e^{i\varepsilon S} H(\varepsilon)\ e^{-i\varepsilon S} = \tilde{H}_K + o(\varepsilon^2) \tag{3.17}$$

expanding in ε one gets

$$\begin{aligned} H_{\mathrm{eff}} \simeq \big(\mathbf{I} + i\varepsilon S\big)\big(\tilde{H}_K + \varepsilon H_\eta\big)\big(\mathbf{I} - i\varepsilon S\big) = \\ H_K + \varepsilon\big(H_\eta + iS\tilde{H}_K - iH_K S\big) + o(\varepsilon^2) \end{aligned}$$

so that requiring terms at order ε to be zero results in

$$[\tilde{H}_K, S] + iH_\eta = 0. \tag{3.18}$$

Inserting (3.18) into (3.17) and imposing $\varepsilon = 1$ (which is the physical case) one gets

$$H_{\text{eff}} = \tilde{H}_K + \frac{i}{2}[S, H_\eta]. \tag{3.19}$$

Now (3.15) can be decomposed into

$$\begin{aligned}
&P_0[\tilde{H}_K, S]P_0 = 0 \rightarrow [P_0HP_0, P_0SP_0] = 0\\
&P_\eta[\tilde{H}_K, S]P_\eta = 0 \rightarrow [P_\eta HP_\eta, P_\eta SP_\eta] = 0\\
&P_0[\tilde{H}_K, S]P_\eta + iP_0H_\eta P_\eta = 0 \rightarrow (P_0SP_\eta)(P_\eta HP_\eta) - (P_0HP_0)(P_0SP_\eta) = iP_0HP_\eta
\end{aligned}$$

which are satisfied if

$$P_0SP_0 = \lambda P_0 \qquad P_\eta SP_\eta = \lambda P_\eta.$$

As λ is arbitrary, one can choose $\lambda = 0$[1]. By denoting now

$$\begin{aligned}
X &\equiv P_0SP_\eta\\
Y &\equiv P_\eta HP_\eta\\
R &\equiv P_0HP_\eta\\
Q &\equiv P_0HP_0
\end{aligned}$$

The last equation to satisfy is indeed

$$X \cdot Y - Q \cdot X = iR \tag{3.20}$$

Limiting ourselves to the eigenspace $\mathcal{E}_\eta$, it is reasonable to assume that Y admits an inverse, so that we can rewrite the previous equation as

$$X = iR \cdot Y^{-1} + Q \cdot X \cdot Y^{-1}$$

which has solution

$$X = i\sum_{n=0}^{\infty} Q^n \cdot R \cdot Y^{-n-1}. \tag{3.21}$$

As the order of magnitude of $Q \cdot Y^{-1}$ is $|t|/U$, this is indeed an expansion in this parameter. So we obtain, to the lowest order

$$X \equiv P_0SP_\eta = iR \cdot Y^{-1} \equiv \frac{i}{U}P_0HP_\eta. \tag{3.22}$$

---

[1] Of course this implies a little loss of generality; an interested reader is invited to read [1] for further details

Using some algebra and this last result one has

$$P_0 H_{\text{eff}} P_0 = P_0 H P_0 - \frac{1}{U} P_0 H P_\eta H P_0 \tag{3.23}$$

$$P_\eta H_{\text{eff}} P_\eta = P_\eta H P_\eta + \frac{1}{U} P_\eta H P_0 H P_\eta. \tag{3.24}$$

It is important now to see that, given the structure of H_0, not all components of P_η will contribute. One has $P_l H_K P_{l'} = 0$ unless $\Delta l = 0, \pm 1$, so in (3.23), which represents the low-energy sector of the Hamiltonian, only P_1 will contribute. It is immediate to see that $P_1 H P_0$ transfers a particle from a single occupied site to another one, creating a doubly occupied site. After this, the $P_0 H P_1$ term transfers aback the particle to an empty site. There are now two different scenarios:

(i) The particle is moved back to the initial site; this is a virtual process, which is not accompanied by any real charge transfer.

(ii) The particle is moved to another site; This is not virtual but is totally quenched at half-filling.

3.5 Perturbation theory $U \gg V \gg t$

As $U \to \infty$ double occupancy is forbidden, so we will not consider the spin degree of freedom. That is equivalent of considering only the subspace identified by the projector

$$\mathbb{P}_0 = \prod_i (1 - n_{i\uparrow} n_{i\downarrow}), \tag{3.25}$$

which eliminates double occupancies. We need now to define the projector onto the state with no nearest or next-to-nearest neighbors (this is analogous to the previously defined *Gutzwiller projector*, which we will call P_0. That is obtained by first eliminating nearest neighbours with the projector

$$\mathbb{P}_1 = \prod_i (1 - n_i n_{i+1}) \tag{3.26}$$

and then next-to-nearest neighbours with

$$\mathbb{P}_2 = \prod_i (1 - n_i n_{i+2}). \tag{3.27}$$

The full projector is therefore obtained by direct multiplication:

$$P_0 = \mathbb{P}_2 \mathbb{P}_1 \mathbb{P}_0. \tag{3.28}$$

We are now able to define the orthogonal complementary P_η, which shall satisfy $P_0+P_\eta = \mathbf{I}$, $P_0P_\eta = P_\eta P_0 = 0$. Calling now P_l the projector onto the space with l nearest or next-to-nearest neighbors one has

$$P_\eta = \sum_{l>0} P_l. \tag{3.29}$$

Splitting the Hamiltonian

$$H = -t\sum_i \left(c^\dagger_{i+1}c_i + \text{h.c.}\right) + V\sum_i\sum_{l=1}^{r_c=2} n_in_{i+l} \equiv H_K + H_V$$

it is trivial that

$$H_VP_0 = P_0H_V = 0 \tag{3.30}$$

while

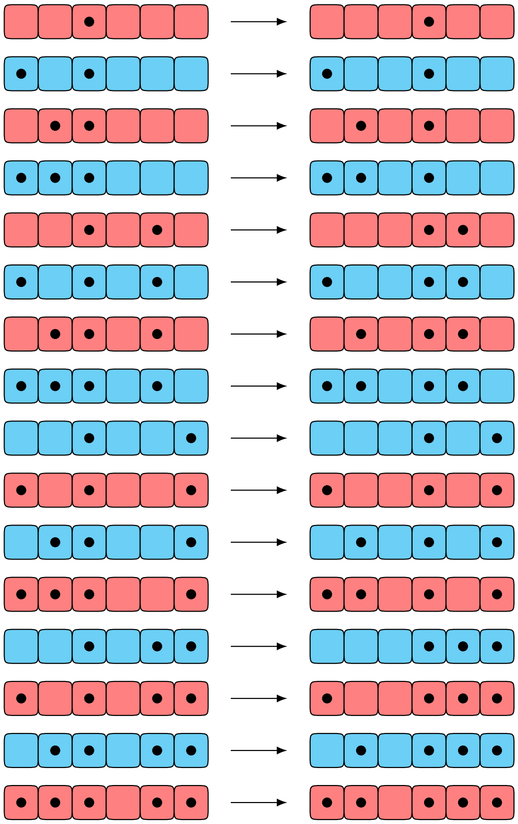
$$H_K = P_0H_KP_0 + P_0H_KP_\eta + P_\eta H_KP_0 + P_\eta H_KP_\eta. \tag{3.31}$$

We can now decompose the Hamiltonian into diagonal and off-diagonal terms by rewriting the creation/annihilation operators via a (less trivial than before) multiplication for one

$$c_i \to c_i\prod_{l=1}^{r_c=2}\big((1-n_{i+l})+n_{i+l}\big)\big((1-n_{i-l})+n_{i-l}\big). \tag{3.32}$$

In this case the expression we find from $c^\dagger_{i+1}c_i$ is quite cumbersome (there are 256 terms). With the help of the symbolic manipulation toolkit FORM [35] we find that the great part of them is irrelevant (always give zero) and just 16 terms contribute. We now list them and examine their action.

$$\begin{array}{r}
(1-n_{i+3})(1-n_{i+2})\ c^\dagger_{i+1}c_i\ (1-n_{i-1})(1-n_{i-2}) \\
(1-n_{i+3})(1-n_{i+2})\ c^\dagger_{i+1}c_i\ (1-n_{i-1})\ n_{i-2} \\
(1-n_{i+3})(1-n_{i+2})\ c^\dagger_{i+1}c_i\ n_{i-1}\ (1-n_{i-2}) \\
(1-n_{i+3})(1-n_{i+2})\ c^\dagger_{i+1}c_i\ n_{i-1}\ n_{i-2} \\
(1-n_{i+3})\ n_{i+2}\ c^\dagger_{i+1}c_i\ (1-n_{i-1})(1-n_{i-2}) \\
(1-n_{i+3})\ n_{i+2}\ c^\dagger_{i+1}c_i\ (1-n_{i-1})\ n_{i-2} \\
(1-n_{i+3})\ n_{i+2}\ c^\dagger_{i+1}c_i\ n_{i-1}\ (1-n_{i-2}) \\
(1-n_{i+3})\ n_{i+2}\ c^\dagger_{i+1}c_i\ n_{i-1}\ n_{i-2} \\
n_{i+3}\ (1-n_{i+2})\ c^\dagger_{i+1}c_i\ (1-n_{i-1})(1-n_{i-2}) \\
n_{i+3}\ (1-n_{i+2})\ c^\dagger_{i+1}c_i\ (1-n_{i-1})\ n_{i-2} \\
n_{i+3}\ (1-n_{i+2})\ c^\dagger_{i+1}c_i\ n_{i-1}\ (1-n_{i-2}) \\
n_{i+3}\ (1-n_{i+2})\ c^\dagger_{i+1}c_i\ n_{i-1}\ n_{i-2} \\
n_{i+3}\ n_{i+2}\ c^\dagger_{i+1}c_i\ (1-n_{i-1})(1-n_{i-2}) \\
n_{i+3}\ n_{i+2}\ c^\dagger_{i+1}c_i\ (1-n_{i-1})\ n_{i-2} \\
n_{i+3}\ n_{i+2}\ c^\dagger_{i+1}c_i\ n_{i-1}\ (1-n_{i-2}) \\
n_{i+3}\ n_{i+2}\ c^\dagger_{i+1}c_i\ n_{i-1}\ n_{i-2}
\end{array}$$



where we have drawn the diagonal and off-diagonal terms in red and blue respectively. We now proceed by dividing the kinetical Hamiltonian into T_A, which includes only the first of the previously listed terms (and its hermitian conjugate), T_B which include the other diagonal terms, and T_{mix} which includes the non-diagonal terms. As the first one represents the translation of a A block, the second is the translation of a B block while the last one mixes the sectors, one finds

$$\begin{aligned}
P_0 H_K P_0 &= P_0 T_A P_0 \\
P_\eta H_K P_\eta &= P_\eta T_A P_\eta + T_B \\
P_\eta H_K P_0 + P_0 H_K P_\eta &= T_{\text{mix}}
\end{aligned} \tag{3.33}$$

so that we can decompose the full Hamiltonian into diagonal and off-diagonal terms

$$H = \big(P_0 H_K P_0 + P_\eta H_K P_\eta + H_V\big) + \big(P_\eta H_K P_0 + P_0 H_K P_\eta\big) \equiv \tilde{H}_K + H_\eta. \tag{3.34}$$

From here one can proceed in the same way as before, but by working out calculations one finds out that in this particular case first-order corrections are equal to zero. Why is that? The fact is that if we act with H_{eff} over the classical ground state $|0\rangle$, one has

$$H_{\text{eff}}\,|0\rangle = H_{\text{eff}}P_\eta\,|0\rangle = \left(H + \frac{1}{V}P_\eta H P_0 H\right)P_\eta\,|0\rangle$$

but it is not possible that the Hamiltonian, acting onto the classical ground state, modifies the number of B blocks from 1 to 0. The only possible scenario with a 1 B block is that of a chain of $L = 10$ sites, but from simple considerations, one can see that if filled with 4 particles with $U \to \infty$, it is impossible to avoid the presence of B blocks. For these reasons P_0 acting onto $HP_\eta\,|0\rangle$ annihilates it. Of course, the same is not true for the previous section as $V = 0$ but shall get this in the limit $V \to \infty$ (where, in principle, every movement is, so to say, forbidden and we have a crystal-like phase). One can then move on to second-order corrections as done in [27], by calculating 2nd order effective Hamiltonian

$$H_{\text{eff}} = H_V + \mathcal{P}H_K\big(1-\mathcal{P}\big)\frac{1}{\epsilon - H_V}\big(1-\mathcal{P}\big)H_K\mathcal{P} + \mathcal{O}(t^4/V^3) \tag{3.35}$$

where $\mathcal{P}$ is the projector onto the classical ground state (whose number of B blocks depend on the density ρ) and ϵ is the classical ground state energy. It is now possible to map this model to a spin chain of M (total number of blocks) elements by associating a spin up to each A block and a spin down to each B block. In the $r_c = 2$ case one finds

$$H_{\text{eff}} = H_V + \frac{t^2}{V}\sum_{i=1}^{M}\left[\big(S_i^+S_i^- + \text{h.c.}\big) + 2\right] + \mathcal{O}(t^4/V^3). \tag{3.36}$$

It was numerically shown in [27] that this effective Hamiltonian is extremely accurate up to $V/t \simeq 10$, starting to be drive away to "real" results at $V/t \simeq 7-5$.

Chapter 4

Machine Learning techniques and DMRG

> *"I am putting myself to the fullest possible use, which is all I think that any conscious entity can ever hope to do."*
>
> Hal 9000 - 2001: A Space Odyssey

In this chapter we are first examining the machine learning (often referred to as ML in the following) techniques that we have used in order to study the phase diagram of the extended Hubbard Model. We also introduce Density Matrix Renormalization Group (DMRG) method, that we used for the calculations of observables and correlators. In particular we will start by looking at the basics of these techniques and then review some known models as the Ising and the Kitaev model.

4.1 Machine Learning

As Machine Learning use grew in everyday life, it caught physicists' attention both because of its mysterious structures (e.g. Neural Networks) and its potential for research [7, 10]. In this master thesis, we are going to focus on techniques whose main objective is phase recognition (as our final goal is indeed to detect and understand different phases of the Extended Hubbard Model). These are divided into supervised and unsupervised methods, where the main difference is that while the former implies some a priori knowledge of the model and its phases, the latter does not. In previous chapters, we have seen that the Extended Hubbard Model is not integrable and that we can study only some extreme limits (i.e. $U \gg t$ with $V = 0$, $t = 0$, $U \gg V \gg t$ etc.). That (together with the author's belief that supervised techniques are more useful to understand Machine Learning rather than Physics) is the reason why we will uniquely use unsupervised techniques. In particular, we will focus on *Principal Component Analysis* (PCA), *k-means clustering* and *Learning by confusion*.

4.1.1 Principal Component Analysis

The first important remark that has to be done regarding PCA is that it is improper to call it a Machine Learning technique as it is rather a data analysis tool (as we will see it is a simple algebraic operation, and there is no learning algorithm).
Given a matrix, the PCA identifies the directions along which the data have maximum variance. In particular, it identifies mutually orthogonal directions in the data space, which are called Principal Components, along which the linear correlation in the data vanishes. Directions that show the least variance are then discarded, while the ones which exhibit large variance are considered to contain the most relevant information.
To do this one starts by stacking the data into a $n \times m$ matrix X (in our case n will be the number of set of data and m the dimension of each set of data) and then normalizing it, which implies modifying it so that it has zero mean column-wise. As real data typically do not have zero mean, normalization simply implies subtracting the mean column-wise. It is then necessary to compute the symmetric $m \times m$ *covariance matrix* $L = X^T X$. The diagonal entries of this matrix (i.e. L_{ii}) represent the variance of the i-th element over the entire data, while the off-diagonal elements (i.e. L_{ij}) are the covariance between the i-th and j-th element. Diagonalizing this matrix (which is always possible since it is symmetric) one finds that the eigenvectors corresponding to the greatest eigenvalues are indeed the directions in the dataset along which the data change most. Finally, to understand the physical meaning of Principal Components, two different approaches are possible:

1. Once the k eigenvectors corresponding to the k most relevant eigenvalues are selected, they are stacked into a matrix $V = [v_1, \ldots, v_k]$. The data are now projected onto the Principal Components by the linear transformation $\tilde{X} = XV$.

2. After selecting a small number of Principal Components $\{v_i\}$, imagining that each row of the original matrix X corresponds to a point in the phase diagram, one can plot the projection onto each eigenvector by computing $u = Xv$ and plotting the components of u in the phase diagram (the phase diagram point identified with the i-th row of X corresponds to the i-th component of u).

The reader who feels confused and cannot understand the difference between the two methods shall not be discouraged as it is, at first sight, obscure. When dealing with the application to studied physical models we will encounter both of them, which shall clarify the difference between the two.

4.1.2 k-means clustering

K-means clustering is a machine learning method for finding clusters and cluster centers in a set of unlabelled data [24]. The idea is to divide the dataset into K different clusters (K is fixed *a priori* by the user) with, of course, K different cluster centers (called *centroids*). These are initially fixed autonomously by the algorithm [34] and then moved (as cluster boundaries) to minimize variance within the cluster. Finally, a possible choice to evaluate the success of the algorithm is to calculate the *Silhouette score*. This is done by computing the euclidean distance between each point and its centroid (which we call a) and then between the point and the second closest centroid (which we call b). The Silhouette is the average of $S = (b - a)/\max(a, b) \in [-1, 1]$. Notice that a point close to its centroid will have $S \lesssim 1$, while points at the boundary between two clusters will have $S \simeq 0$.

4.1.3 Learning by confusion

Convolutional Neural Networks (CNN) have had astonishing growth recently, succeeding in complicated tasks such as pattern recognition, multi-media compression, and image classification [21, 48, 49]. We will not describe in-depth Convolutional Neural Networks and Neural Networks in general, as it exulates the main aim of this thesis. We address the interested readers to [17].

First of all one should have a set of $n \times m$ matrices to start with (each matrix can represent an image or, in our case, the data set of physical observables). This is randomly split into a training, validation and test batch (the use of these will soon be clarified). Every element of the set has a label which classifies it (the number of the possible labels can vary from case to case). After dividing the set, the training of the network begins. During the training every element of the training and the validation set and is given to the network, which tries to predict the right label. In order to guess the label some *filters* (also called *kernels*, they are small square matrices) "run" over the input matrix (see Fig. 4.1). Results of this convolution are stored into a new map which is sometimes called the *feature map*. The shape of the feature map depends on the shape of the input matrix and the size of the kernel. After this first convolution, other convolutions or linear transformations may be applied until the final transformation finally produces the result. Each transformation is called a *layer* and is composed of *neurons* (There is no precise rule on how to build the network, only empirical rules set by experience; see

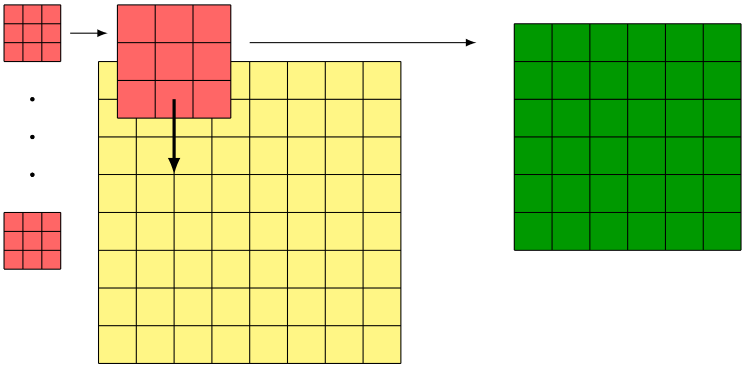


Figure 4.1: Example of convolution; here a 3×3 kernel is applied over a 8×8 matrix, resulting in a 6×6 feature map.

[17] for more details). The last layer produces the output of the network, which is the predicted label. The values defining the kernels and the linear transformations (which were set randomly at the beginning) are now modified in order to reduce the difference between the predicted and "real" labels of the training set (in technical terms in order to minimize the *loss* function [17]) and the operation is repeated onto the validation set. The magnitude of the change in the parameter is controlled by the *learning rate* l, which is an arbitrary external parameter (commonly set $l \sim 10^{-2}$ by empirical rules). The operation is then repeated for an arbitrary number of times, called the *number of epochs* n_{epochs}. Finally one should check the result of the training onto the test set (this is tipically done by checking the accuracy a, which is the percentage of right labelling performed by the network).
In [31] an unsupervised method involving Neural Networks to study 2 phases systems was introduced. The main idea is to "confuse" the network by mislabeling the training data and then checking the accuracy over the testing data. Imagine as an example a 2 phase system which is regulated by a parameter x (which can either be a temperature or a Hamiltonian parameter). In particular we assume that the phase transition occurs at $x = x_c$. The system is simulated for various values of x and observables which shall indicate the transition are stored into $n \times m$ matrices (one matrix for each value of x). Once the data are collected the procedure is articulated as follows:

1. Start by assigning to all the training data the label 1; now split the set randomly into a training set, a validation set and a test set. Now train the network for a fixed number of epochs n_{epochs} and with a given learning rate l. Once the training is done check the accuracy a onto the test set. At this first stage the accuracy shall be high as the network sees the data as uniform.

2. Re-label the training data, now assigning to all data corresponding to some value

of the parameter x less than a given $\tilde{x}$ the label 0 and 1 to all the others and repeat what previously done; the accuracy shall now naturally decrease.

3. Increase $\tilde{x}$ and compute again the accuracy of the network until $\tilde{x}$ reaches the maximum simulated value (at this last stage the network shall show a new maximum in the accuracy as the training batch is uniform again).

Plotting the accuracy of the network we now expect it to show a local maximum). when the labeling correctly divides two different phases, i.e. when $\tilde{x} = x_c$. This is traditionally called the W-shape of the accuracy.

4.2 DMRG

Density Matrix Renormalization Group (DMRG) is the state-of-the-art best numerical method to study 1D systems. Its origin consists in the failure of the Real Space Renormalization Group (RSRG). As they both use the Renormalization Group idea of integrating out some degrees of freedom without modifying the physics, the difference is that while the former operates on wave functions, the latter acts on density matrices. The proof of the failure of RSRG was given in [46].
The DMRG procedure is structured as follows [38, 39]:

(i) Take a system of finite length l (a *block*) and suppose its Hilbert space is M-dimensional with states $\{\ |i_l\rangle\}$ In particular, during the first iteration, the block is composed of just one site so that $M = 2$ for standing spin 1/2 particles (i.e. particles which cannot move, which is not the Hubbard case; for moving particles $M = 4$) ; as the block length grows the exact dimension of the Hilbert space would become too large to handle and is therefore reduced (as we will see) so that M is no longer the exact dimension of the Hilbert space but is the dimension of a reduced Hilbert space. The block Hamiltonian H_l (and other local operators representation) is composed by the matrix elements $\langle i_l| H_l |j_l\rangle$.

(ii) Imagine now adding a site to the previous block and forming a new block of length $l + 1$. Taking the basis $\{\ |\sigma\rangle\}$ for the 1-site Hilbert space (which we take as N-dimensional), we can write the new block Hamiltonian H_{l+1} in the product basis $\{\ |i_l\rangle\,|\sigma\rangle\} \equiv \{\ |i_l\sigma\rangle\}$.

(iii) We now mimic the Thermodynamic limit by embedding this block into an *environment* composed of a copy of it, arriving therefore to a *superblock* of length $2l + 2$. As we want to compute the density matrix, it is necessary to find the superblock ground state $|\psi\rangle$. That can be written as

$$|\psi\rangle = \sum_{i^S}\sum_{\sigma^S}\sum_{i^E}\sum_{\sigma^E} \psi_{i^S\sigma^S i^E\sigma^E} |i^S\rangle |\sigma^S\rangle |i^E\rangle |\sigma^E\rangle \equiv \sum_{j^S j^E} \psi_{j^S j^E} |j^S\rangle |j^E\rangle$$

where the index S, E indicates the system or the environment. The dimensions of the Hilbert spaces containing $|j^S\rangle$, $|j^E\rangle$ are $N^S = M^S N_{\text{site}}$ and $N^E = M^E N_{\text{site}}$ respectively (with N_{site} being the dimension of the Hilbert space of one site). It is easy to see that N^S, N^E fast become too big to be handled by normal computational resources. It is therefore necessary to cut that space removing some degrees of freedom. This is done through some different truncation procedures (e.g. *Optimization of expectation values*, *Optimization of the wave function* , or *Optimization of entanglement* as reported in [38]) and reduces the Hilbert space dimension to $M^S < N^S$ and $M^E < N^E$. Once the truncation is done it is possible to find the ground state $|\psi\rangle$ through large sparse matrix diagonalization of H_{2l+2}.

(iv) It is possible to form the reduced density matrix by tracing out the environment $\rho = \mathrm{Tr}_E |\psi\rangle \langle\psi|$ and to determine the new eigenbasis $|w_\alpha\rangle$, ordered by descending eigenvalues w_α. The "largest" eigenstates (the cutoff is determined a priori) are kept and form a new basis for the system S whose dimension is M^S. By computing the matrix elements between the product basis elements $|i_l^S\sigma\rangle$ and the "largest weights" basis $|j_{l+1}\rangle$ we find the change of basis rectangular basis T. The same is done for the environment.

(v) Finally we carry out the reduced basis transformation $H'_{l+1} = T^\dagger H_{l+1} T$ and restart until some final length is achieved.

This procedure is called the *infinite-system* DMRG. This algorithm alone gives good results in general but presents some problems in systems such as the one under study. In our case, the soft-shoulder potential (together with the filling under consideration) has physical effects which are not taken into account by intermediate Hamiltonians (as they do not include all the chain). For this reason, *finite-system* DMRG has been used. The procedure consists in:

(i) The *infinite-system* DMRG is stopped at some *superblock* of length L.

(ii) DMRG is now applied again, but now the length of the superblock is fixed, which implies that the growth of the system is done at the expense of the shrinkage of the environment.

(iii) When the environment size reaches a minimum size the growth direction is reversed (the environment grows at the expense of the system).

A complete growth and shrinkage is called a *sweep* and takes about two to four times the CPU time necessary to complete the *infinite-system* DMRG.
Numerical calculations used in this work were performed by massive use of the iTensor library [13]. This library works out the DMRG by parametrizing the Hilbert space by Matrix Product State (MPS) [39].
Finally, it has to be said that from the beginning of DMRG it has been known that it offers much more accurate results with open boundary conditions (OBC) for Periodic Boundary Conditions (PBC) [47]. The reasons for these problems were summarized in [44] along with possible solutions. In the following, we are always going to consider PBC as they have more physical advantages (removing finite-size and boundary effects) than computational disadvantages.

4.3 Applications to physics models

We will now briefly review the two models under study and then examine how Machine Learning has been used in order to study their phase diagrams.

4.3.1 The Ising model

The Hamiltonian of the model reads

$$H = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j \tag{4.1}$$

where σ_i is the discrete spin (can only take $\sigma = \uparrow, \downarrow$ as values) at site i and $\langle ij \rangle$ stands for the sum over neighboring sites. Due to its simplicity, it is commonly used as the first example in Statistical Mechanics and Statistical Field Theory books. It is a rare example of an integrable system (for $d = 1, 2$). In particular, the case with $d = 2$, it represents the simplest example of 2nd order phase transition between a high-temperature paramagnetic (disordered) phase and a low-temperature ferromagnetic (ordered) phase. The transition temperature was analytically found by Osanger in 1944 [32]

$$T_c = \frac{2J}{k_B \ln\left(1 + \sqrt{2}\right)}. \tag{4.2}$$

4.3.2 The Kitaev model

The Hamiltonian of the non-interacting case reads

$$H^{NI} = J \sum_i \left(a^\dagger_{i+1} a_i + \text{h.c.}\right) + \Delta \sum_i \left(a_i a_{i+1} + \text{h.c.}\right) + \mu \sum_i n_i \tag{4.3}$$

where $a_i^\dagger$, a_i are creation and annihilation operators for spinless fermions at site i and $n_i = a_i^\dagger a_i$. The interacting Hamiltonian is

$$H^I = H^{NI} + \ V \sum_i n_i n_{i+1}. \tag{4.4}$$

If we consider periodic boundary conditions the former is exactly solvable by means of Fourier transform and Bogoliubov transformation [19]

$$H^{NI} = \sum_k E(k) \eta_k^\dagger \eta_k$$

where η_k are Bogoliubov operators and

$$E(k) = \sqrt{(J\cos(k) + \mu/2)^2 + (\Delta \sin(k))^2} \tag{4.5}$$

while the latter is not solvable. Naturally, they exhibit two different phase diagrams. In particular, the non-interacting phase diagram depends on the values of μ and Δ and exhibits three different phases: a trivial phase for $\mu < -2$ and $\mu > 2$ and two superconducting phases (which differ for their topological winding number) for $-2 < \mu < 2$ and $\Delta > 0$ or $\Delta < 0$.

4.3.3 Applications of Machine Learning

We will now consider what was done in [10, 42]. We have seen that the Ising model is an example of a 2-phase system. For this reason it is a perfect model onto which to apply Learning by confusion. This was done in [10] using a chain of $L = 30$ spins, using 100 spin configurations sampled using Monte Carlo methods at temperatures T varying from $T_1 = 1$ to $T_2 = 3.5$ in equidistant steps as input. The results are plotted in Fig. 4.2 and show a good agreement with Osanger prediction, which was $T_c = 2.23$ for the selected value of J.

On the other hand k-means clustering has proved itself to be useful when dealing with richer phase diagrams. In [42], for the non-interacting case, the single-particle standard and anomalous correlation functions

$$c(k) = \sum_{ij} e^{ik(i-j)} \langle a_i^\dagger a_j \rangle$$
$$f(k) = \sum_{ij} e^{ik(i-j)} \langle a_i a_j \rangle$$

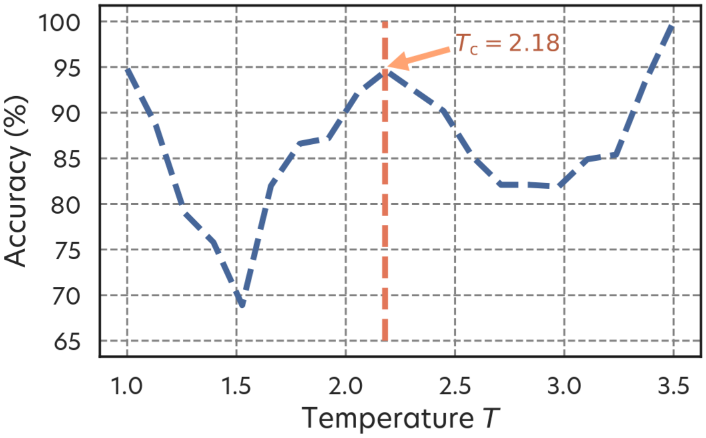


Figure 4.2: Plot of the accuracy as function of the guessed T_c. The shape resembles the predicted W-shape; image taken from [10].

where numerically found for $N = 12800$ different points of the parameter space (with $\Delta \in [-2, 2]$ and $\mu \in [-8, 8]$) for $L = 100$ different values of k ($k_n = 2\pi/L$, $n = 0, \ldots, L-1$) and stacked into a design matrix X. The authors then applied k-means clustering onto X, varying the number of clusters in order to obtain maximum average Silhouette score. Once they found the maximum for $n_{clusters} = 4$, they plotted the score of every element of the parameter space, finding a very good agreement with the exact phase diagram (results are depicted in Fig. 4.3). The same procedure was followed for the interacting case (but this time in the μ, V space, with $\mu \in [0, 5]$ and $V \in [-4, 4]$), obtaining again a good representation of the expected phase diagram. Results are reported in Fig. 4.4

Finally we look at how PCA was used in both works. We previously mentioned the fact that there are two different approaches when using PCA, with the first one prescribing to stack the most k relevant eigenvectors (where k depends on some arbitrary cutoff for the relative eigenvalue weight) into a matrix V and then studying the projection of the full design matrix X onto the principal components $\tilde{X} = XV$, while the second suggests to study each of the most relevant k eigenvectors separately, multiplying each line of X (recall that each line of X corresponds to a point in the parameter space) and then plotting the result in the parameter space itself (e.g. with a heatmap). In [10] PCA was applied onto the previously mentioned spin configurations considering the $k = 2$ most relevant components. The result is plotted in Fig. 4.5. The data are clearly divided into 3 clusters (two low-temperature and one high-temperature) and it can be shown [45] that the first principal component corresponds to the magnetization. Furthermore if one draws a decision boundary (perpendicular to the first principal component) which

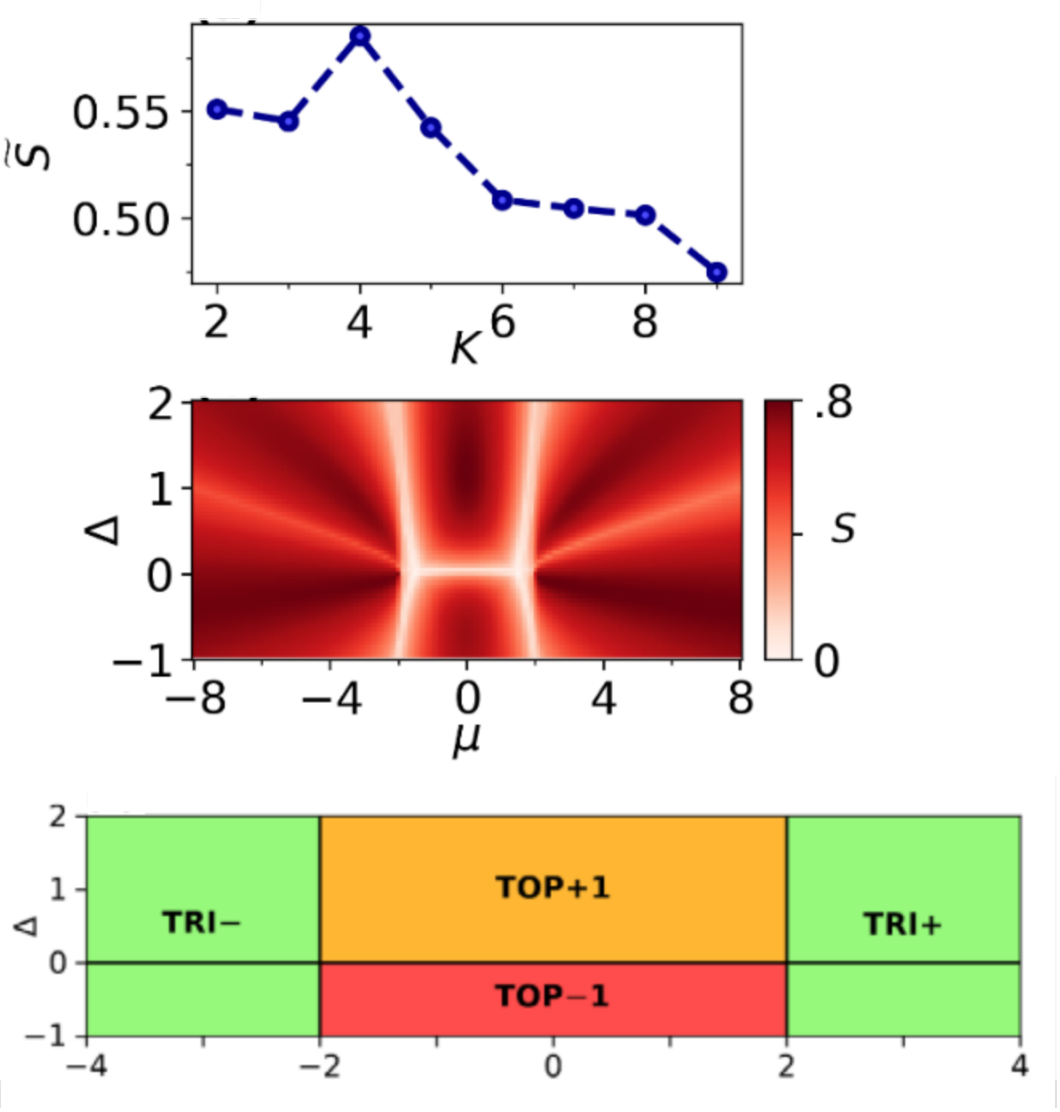


Figure 4.3: Study of the phase diagram of the non-interacting Kitaev model with k-means clustering. Moving vertically from top to bottom, in the first image the average Silhouette score $\bar{S}$ in function of the number of cluster is plotted, with a clear maximum for $K = 4$. In the second image the Silhouette score is given for the different points in the phase diagram. The clustering in parameter space clearly resembles the exact phase diagram, which is represented in the last image. Images from [42].

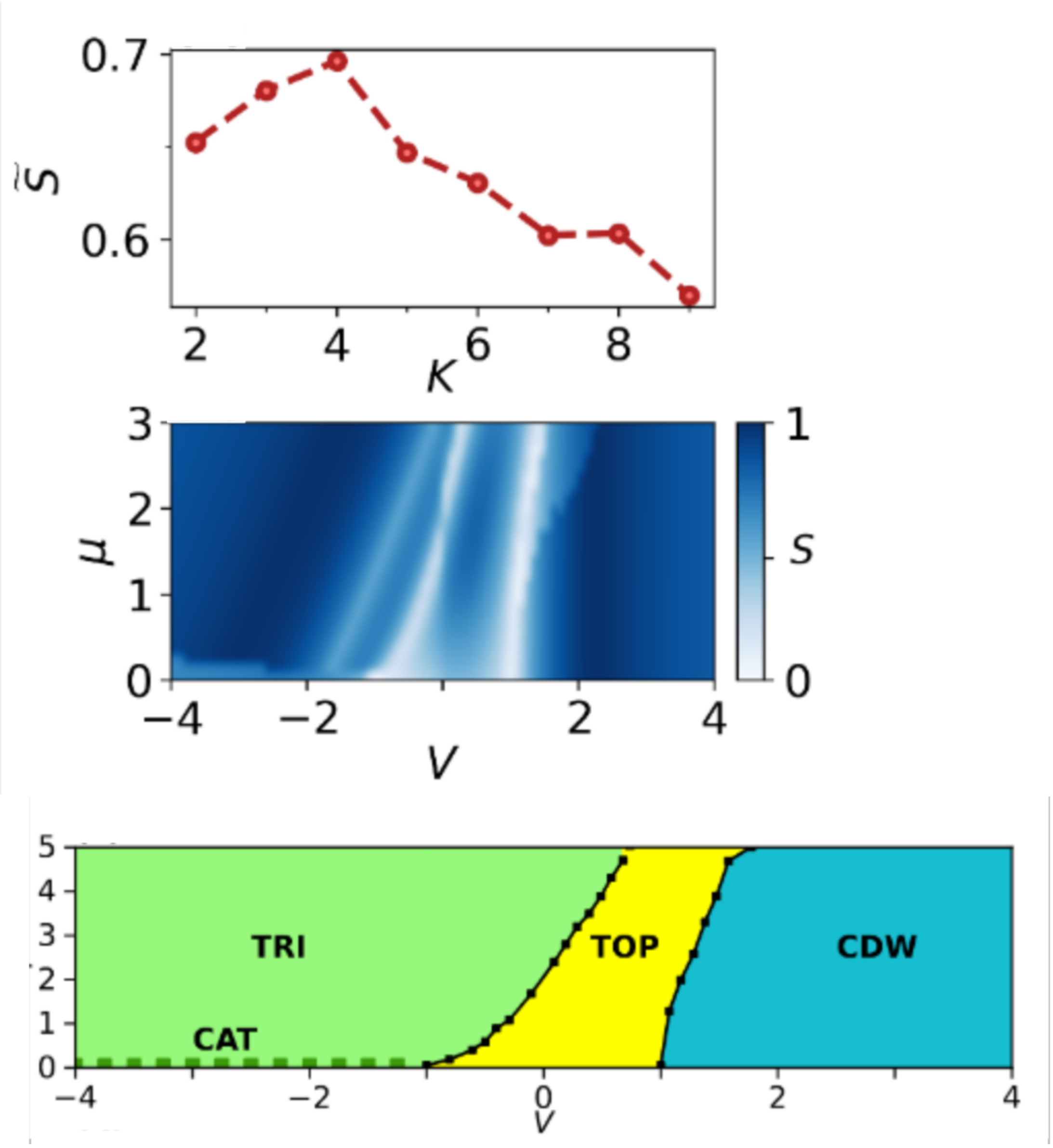


Figure 4.4: Study of the phase diagram of the interacting Kitaev model with k-means clustering. Moving vertically from top to bottom, in the first image the average Silhouette score $\bar{S}$ in function of the number of cluster is plotted, with a clear maximum for $K = 4$. In the second image the Silhouette score is given for the different points in the phase diagram. The clustering in parameter space clearly resembles the expected phase diagram, which is represented in the last image. Images from [42].

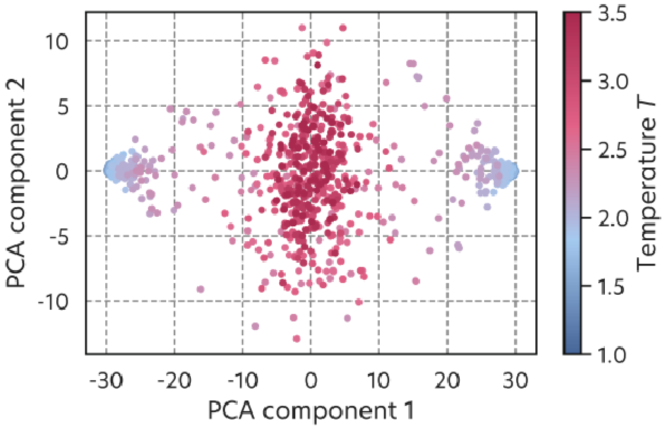


Figure 4.5: PCA applied to the Ising model. It is manifest the separation into two low-temperature and one high-temperature clusters. Image from [10].

separates the high-temperature cluster and a low-temperature cluster a rough estimate of the critical transition temperature can be obtained as $T \simeq 2.3$, compatible with the Osanger prediction. On the other hand in [42] PCA was applied to the matrix X containing the single particle correlation functions (in the non-interacting case) but the second prescription was followed. In this case the first $k = 4$ eigenvectors were selected. The result (which is plotted in Fig. 4.6) was that principal components manifestly underlined some important features of the phase diagram. Looking at Fig. 4.6 (a) we can see that the first eigenvector projection reveals the points with topological winding number $\nu = 0$, the second the ones with $\nu = \pm 1$, the third phase transition lines and the fourth has a high projection onto the points with $\nu = -1$. Later the projection onto the same eigenvectors of the non-interacting case was studied in the interacting one (i.e. design matrix of the interacting Hamiltonian but with eigenvectors of the non-interacting one) obtaining surprising results. In fact it can be seen in Fig. 4.6 (b) that the first and fourth eigenvectors projections reveals the regions with non-trivial winding number, while the third one puts in evidence a phase transition line.

It should be stressed that the main difference between the two approaches is that while the first analysis happens in the principal component space, while the second is the plot of a sort of "score" in the parameter space.

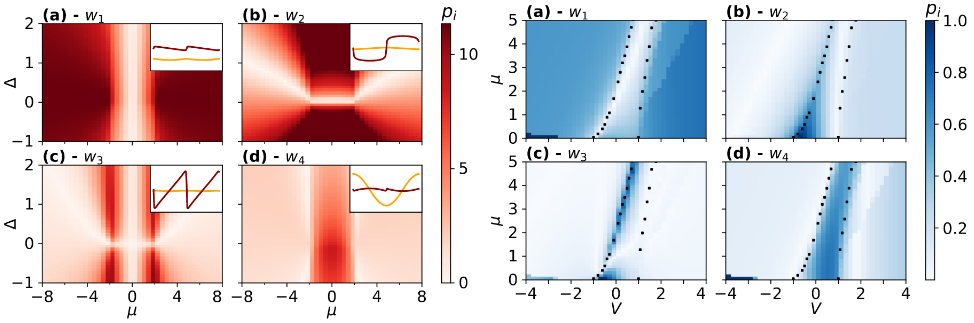


Figure 4.6: PCA application to the non-interacting (left panel) and interacting (right panel) Kitaev model. For each panel, plot (a), (b), (c), (d) represent the projection onto the first four principal components showing some key aspects such as different phases or phase transition lines. Image from [42].

Chapter 5

ML techniques for the Extended Hubbard Model

> *"And I knew exactly what to do. But in a much more real sense, I had no idea what to do."*
>
> ---
>
> Micheal Scott -
> The Office

We are now going to analyze the results obtained with numerical simulations and the use of Machine Learning techniques. First of all, we have to examine the observables required to understand the different phases of the system. Then we are going to build the design matrix X onto which we will apply the previously illustrated Machine Learning techniques and finally we will study the results and build our proposal for the phase diagram.

5.1 Observables

5.1.1 Charge structure factor

The static charge structure factor is defined as

$$S_c(k) = \frac{1}{L} \sum_{l,j} e^{ik(l-j)} g_{2c}(l-j) \tag{5.1}$$

where

$$g_{2c}(l-j) = \langle n_l n_j \rangle - \langle n_l \rangle \langle n_j \rangle \, . \tag{5.2}$$

In the thermodynamic limit, this is the Fourier transform of the density-density correlation function. To calculate it, one can use the field theory description of Chapter 2 and get

$$\langle\rho(r)\rho(0)\rangle = \frac{1}{\pi^2}\left\langle\nabla\varphi_\rho(r)\nabla\varphi_\rho(0)\right\rangle + \frac{1}{(2\pi\alpha)^2}\left[e^{-i2k_Fx}\left\langle e^{2i(\varphi_\rho(r)-\varphi_\rho(0))}\right\rangle + \text{h.c.}\right]. \tag{5.3}$$

To calculate this we need to compute

$$\mathcal{R}(r_1-r_2) = \left\langle(\varphi_\rho(r_1)-\varphi_\rho(r_2))^2\right\rangle. \tag{5.4}$$

Fourier transforming it, we get

$$\mathcal{R}(r_1-r_2) = \frac{1}{(\beta L)^2}\sum_{\mathbf{q}_1\mathbf{q}_2}\left\langle\varphi_\rho(\mathbf{q}_1)\varphi_\rho(\mathbf{q}_2)\right\rangle\ \left(e^{i\mathbf{q}_1r_1}-e^{i\mathbf{q}_1r_2}\right)\left(e^{i\mathbf{q}_2r_1}-e^{i\mathbf{q}_2r_2}\right) \tag{5.5}$$

where $r=(x,u\tau)$ are the coordinates in space-time (it really is space-imaginary time as τ is the Wick rotated analogue of the real time t), with u velocity, $\mathbf{q}=(k,\omega_n/u)$, L is the space period while β is the imaginary time τ one [15]. The quantity

$$\left\langle\varphi_\rho(\mathbf{q}_1)\varphi_\rho(\mathbf{q}_2)\right\rangle$$

in the right-hand side can be computed using path-integral formulation (see [15])

$$\left\langle\varphi_\rho(\mathbf{q}_1)\varphi_\rho(\mathbf{q}_2)\right\rangle = \frac{1}{Z}\int\mathcal{D}\varphi_\rho(x,\tau)\mathcal{D}\theta(x,\tau)\ \varphi_\rho(\mathbf{q}_1)\varphi_\rho(\mathbf{q}_2)\ e^{-S}$$

where, for Luttinger liquids

$$\begin{aligned}
-S &= \int_0^\beta d\tau\int_0^L dx\left[\frac{i}{\pi}\nabla\theta(x,\tau)\ \partial_t\varphi_\rho(x,\tau) - \frac{1}{2\pi}\left(uK(\nabla\theta)^2+\frac{u}{K}(\nabla\varphi_\rho)^2\right)\right]\\
&= -\frac{1}{2\pi}\frac{1}{\beta L}\sum_{\mathbf{q}}\left(\theta^*_{\rho\mathbf{q}},\varphi^*_{\rho\mathbf{q}}\right)\begin{pmatrix}k^2uK & ik\omega_n\\ ik\omega_n & k^2\frac{u}{K}\end{pmatrix}\begin{pmatrix}\theta_{\rho\mathbf{q}}\\ \varphi_{\rho\mathbf{q}}\end{pmatrix}\\
&\equiv -\frac{1}{2}\frac{1}{\beta L}\sum_{\mathbf{q}}\left(\theta^*_{\rho\mathbf{q}},\varphi^*_{\rho\mathbf{q}}\right)M^{-1}\begin{pmatrix}\theta_{\rho\mathbf{q}}\\ \varphi_{\rho\mathbf{q}}\end{pmatrix}.
\end{aligned} \tag{5.6}$$

One can then perform the sum over theta (as the quantity to average does not depend on it) and cancels the, so to say, "theta part" of the partition function in the denominator. This leaves us with the "phi part" of the action

$$\begin{aligned}
S_{\varphi_\rho} &= \frac{1}{\beta L}\sum_{\mathbf{q}=(k,\omega_n/u)}\left[\frac{\omega_n^2}{u}+uk^2\right]\varphi^*_\rho(\mathbf{q})\ \varphi_\rho(\mathbf{q})\\
&= \frac{1}{2\pi K}\int_0^L dx\int_0^\beta d\tau\left[\frac{1}{u}(\partial_t\varphi_\rho)^2+u(\partial_x\varphi_\rho)^2\right].
\end{aligned} \tag{5.7}$$

From this, it is possible to find [15]

$$\langle \varphi_\rho(\mathbf{q}_1)\varphi_\rho(\mathbf{q}_2)\rangle = \frac{\pi K \delta_{\mathbf{q_1},-\mathbf{q_2}} L\beta}{\omega_n^2/u + uk^2}. \tag{5.8}$$

Inserting this last result back into (5.5) one gets

$$\mathcal{R}(r_1 - r_2) = \frac{1}{\beta}\sum_{\omega_n}\int \frac{dk}{2\pi}\frac{2\pi K}{\omega_n^2/u + uk^2}\big(1 - \cos(kx + \omega_n \tau)\big). \tag{5.9}$$

One shall now use some analytic tricks (e.g. inserting $e^{-\alpha k}$, inserting Bose occupation function $f_B(z) = 1/(e^{\beta z} - 1)$ etc. see [15] for more details) and get

$$\mathcal{R}(r) = K \log\left[\frac{x^2 + (u|\tau| + \alpha)^2}{\alpha^2}\right] \tag{5.10}$$

with $r = r_1 - r_2$. Using this and the fact that for quadratic Hamiltonians

$$\left\langle e^{i\sum_j (A_j \varphi_\rho(r_j) + B_j \varphi_\rho(r_j))}\right\rangle = e^{-\frac{1}{2}\langle(\sum_j (A_j \varphi_\rho(r_j) + B_j \varphi_\rho(r_j))^2\rangle}$$

one gets

$$\langle \rho(r)\rho(0)\rangle = \frac{K}{2\pi^2}\frac{y_\alpha^2 - x^2}{(x^2 + y_\alpha^2)^2} + \frac{2}{(2\pi\alpha)^2}\cos(2k_F x)\left(\frac{\alpha}{r}\right)^{2K} \tag{5.11}$$

where $y_\alpha = u\tau + \alpha\ \mathrm{Sign}(\tau)$.
We shall now stop and focus on this last expression. Similar to the Fermi liquid correlation function the $q \sim 0$ part of the correlator decays as $1/x^2$ (just as a free fermion). The difference emerges in the second term, the $2k_F x$ one. While in the Fermi liquid case that decays as $1/r^2$, in this case, it decays as non-universal power law which depends on the interactions through the coefficient K. Furthermore, this term will be of particular interest in the following because Fourier-transforming it a peak emerges at $k = 2k_F$. The value of k_F depends, as we will see, on the nature of the microscopic entities composing the system, i.e. particles or clusters.

5.1.2 Spin structure factor

The static spin structure factor is defined as

$$S_s(k) = \frac{1}{L}\sum_{l,j} e^{ik(l-j)} g_{2s}(l - j) \tag{5.12}$$

where

$$g_{2s}(l - j) = \langle S_l^z S_j^z\rangle - \langle S_l^z\rangle\langle S_j^z\rangle\,. \tag{5.13}$$

This can be calculated similarly to the charge structure (swapping $\varphi_\rho \to \varphi_\sigma$) factor but shall exhibit a well-defined peak at the one-particle density $k = \rho\pi$ (which gives $k = 2\pi/5$ for $\rho = 2/5$, which is the case under examination) when double occupancy is forbidden, while shall emerge for lower values of k if double occupied states start their formation.

5.1.3 Local density

The local density is the most intuitive observable from a physical point of view. It is defined as

$$n(x_i) = \langle 0| \, n_i \, |0\rangle \, . \tag{5.14}$$

This shall give us some insights about the liquid or crystal nature of our system. The latter in fact shall be characterized by a quasi-total localization of particles, while the former shall correspond to a quasi-uniform distribution over the sites (the liquid phase is a momentum eigenstate while the crystal phase is a position eigenstate). This also gives a good check on numerical simulation via particle number conservation

$$\sum_i n(x_i) = N.$$

5.1.4 Double occupancy and nearest and next-to-nearest neighbors numbers

We define double occupancy ($\mathcal{D}$) and nearest and next-to-nearest neighbors ($\mathcal{N}$) numbers as

$$\mathcal{D} = \sum_i \; n_{i\uparrow} n_{i\downarrow} \tag{5.15}$$

$$\mathcal{N} = \sum_i \sum_{l=1}^{r_c=2} \; n_i n_{i+l}. \tag{5.16}$$

These are more similar to order parameters and have clear computational advantages w.r.t. the other observables.

For different values of U and V we performed numerical simulations using iTensors to calculate the above mentioned observables on a chain of $L = 30$ sites and with PBC (Periodic Boundary Conditions). We set the number of sweeps varying from 20 to 30 and with the maximum bond dimension varying from 100 to 400. No significant changes in the results were detected for a different number of sweeps and different maximum bond dimension.

5.2 k-means

We simulated the model for 100 equally spaced values of $U \in [0, 25]$ and 40 equally spaced values of $V \in [0, 10]$, creating a grid of 4000 points in the parameter space. For each simulation we followed what was done in [6], numerically evaluating $S_c(k)$ and $S_s(k)$ for 15 different values of k, $k_n = 2\pi n/L$ with $n = 0, \dots, L/2$ and $L = 30$ (half of chain is enough in order to collect information about the system because of PBC) and we also evaluated the local density $n(x_i)$ onto each of the 30 sites composing the chain. We finally arranged all the observables into a 4000×60 matrix X composed as follows:

$$X = \begin{pmatrix} S_c^{(1)}(k_0) & \dots & S_c^{(1)}(k_{14}) & S_s^{(1)}(k_0) & \dots & S_s^{(1)}(k_{14}) & n^{(1)}(x_1) & \dots & n^{(1)}(x_{30}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ S_c^{(N)}(k_0) & \dots & S_c^{(N)}(k_{14}) & S_s^{(N)}(k_0) & \dots & S_s^{(N)}(k_{14}) & n^{(N)}(x_1) & \dots & n^{(N)}(x_{30}) \end{pmatrix}$$

where, again, $k_n = 2\pi n/L$, and the superscript is indicating that the various observables have been computed in $N = 4000$ different points of the phase diagram. We varied the number of clusters to find the maximum value of the Silhouette score. As shown in Fig. 5.1, the Silhouette score exhibits two maxima at $n = 2$ and $n = 8$. We study the two hypotheses separately.

$n = 2$

The labels assigned by the algorithm are plotted in Fig. 5.2. As we will see later on by looking at the centroids in the $n = 8$ case, the algorithm first detects the difference between the crystal and the liquid phase. In particular, the boundaries of the crystal phase can be identified with $U \simeq 5$ and a line with a slope similar to the one which separates the two classical ground states, i.e. $U = 3V/2$. From the Silhouette score plot, we can see that the average value is quite high (as expected) but has a visibly decreases for the points assigned to the liquid phase and values of $V \gtrsim 7$.

$n = 8$

In this case, the picture is richer. The first things that pop to the eye are that the crystal phase previously detected is again revealed and that the lower Silhouette region that we previously noted (liquid $V \gtrsim 7$) is now divided into different regions, but well separated by the other one ($V \lesssim 7$), as depicted in Fig. 5.3. We shall now look at properties of the centroids to understand the physical content of the clusters. We recall that the centroid can be considered as the most representative point of a cluster found by the algorithm. Since each centroid found by the kmeans algorithm is an array of 60 elements, we can

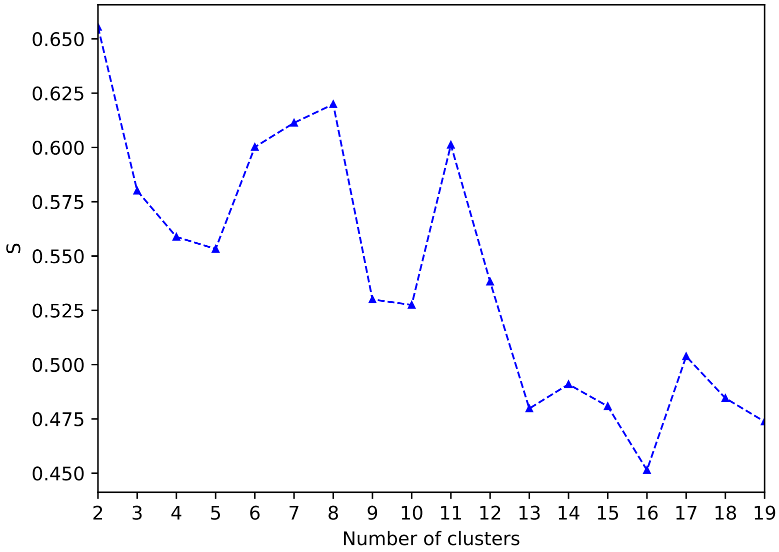


Figure 5.1: Silhouette score in function of the number of the cluster. We see two clear maxima at $n = 2$ and $n = 8$ with an average silhouette $S = 0.655$ and $S = 0.620$ respectively.

divide it into a charge structure factor part, a spin structure factor part a and a local density part.

Fig. 5.4 shows the charge structure factor (i.e. the first 15 elements of the array) of the eight centroids. Recalling that the peak of it shall underline the number of elementary units composing the system (i.e. the number of blocks) we immediately see from Fig. 5.4 that there are two regimes with $M = 8$ ($tt - CLL_d$, $tt - CLL'_d$), three with $M = 9$ (CLL_{nn}, $t - CLL_d$, CLL'_{nn}), one with $M = 10$ (CLL_d), while the TLL and the *locked* phase do not exhibit a clear peak. To understand the difference between phases with the same number of blocks we turn to the density part of the centroids (for the time being these names are only labels, we will give them physical content later).

Starting from the high-U part of the phase diagram we can see in Fig. 5.5 that the local density part of the centroid has a perfectly uniform distribution in the TLL phase, then, raising V, starts developing periodic peaks in CLL_{nn} phase, and finally becomes very peaked in the CLL'_{nn} phase. In particular, we can see that the occupation distribution in the CLL'_{nn} phase resembles the $B\ A\ A$ blocks alternation of the classical ground state. Moving now to the low-U, high-V part of the phase diagram we see that here it is difficult to catch a block structure of the system due to the nonregular density shape as depicted

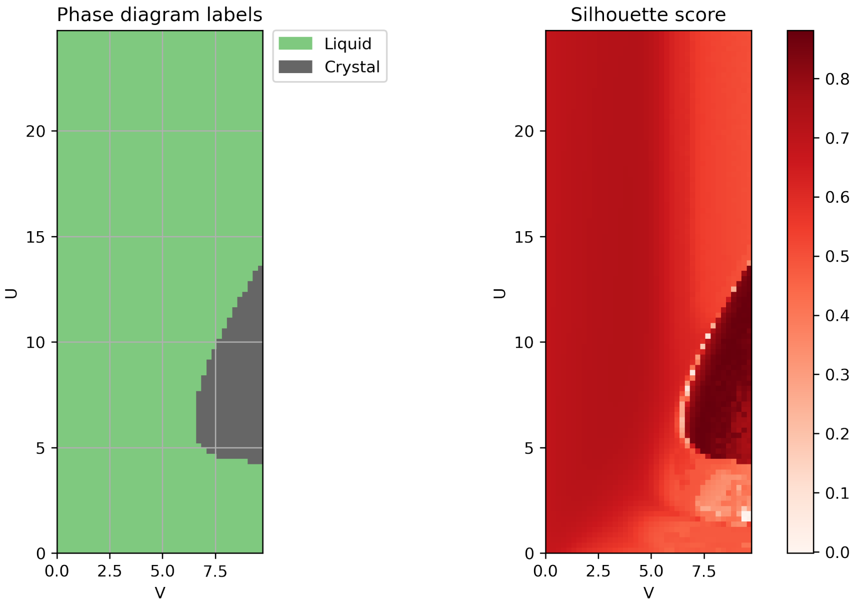


Figure 5.2: Left panel: labels assigned by k-means algorithm with $n = 2$ clusters. Right panel: values of the corresponding Silhouette score (darker for higher values).

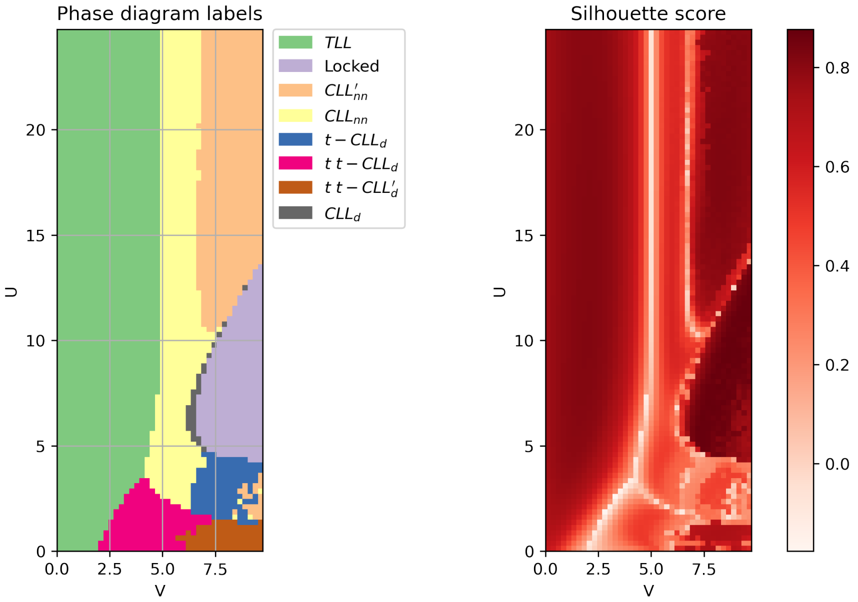


Figure 5.3: Left panel: labels assigned by k-means algorithm with $n = 8$ clusters. Right panel: values of the corresponding Silhouette score (darker for higher values). For explanation about the different phases, see text.

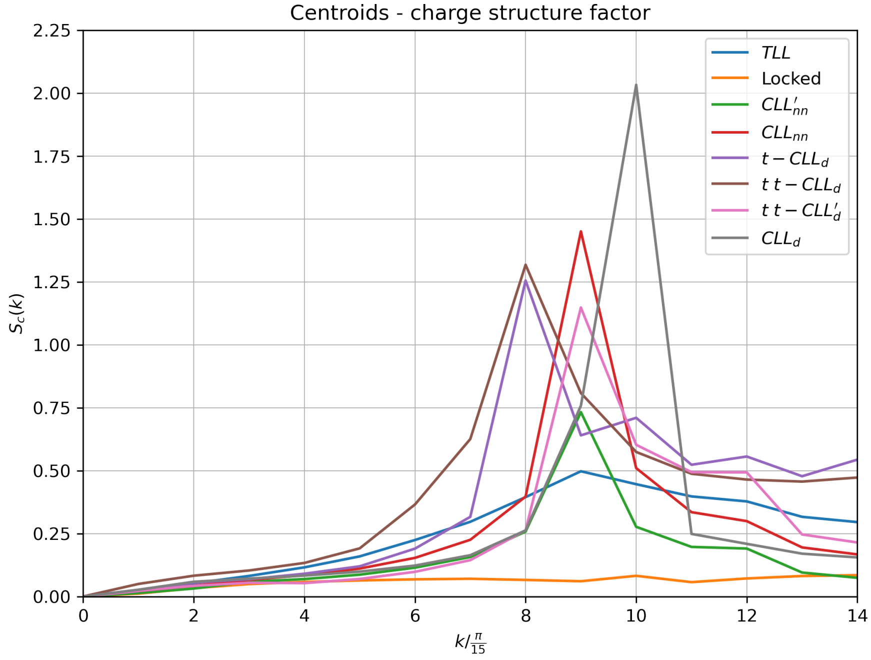


Figure 5.4: Charge structure factor part of the eight centroids. The nomenclature is explained in the text.

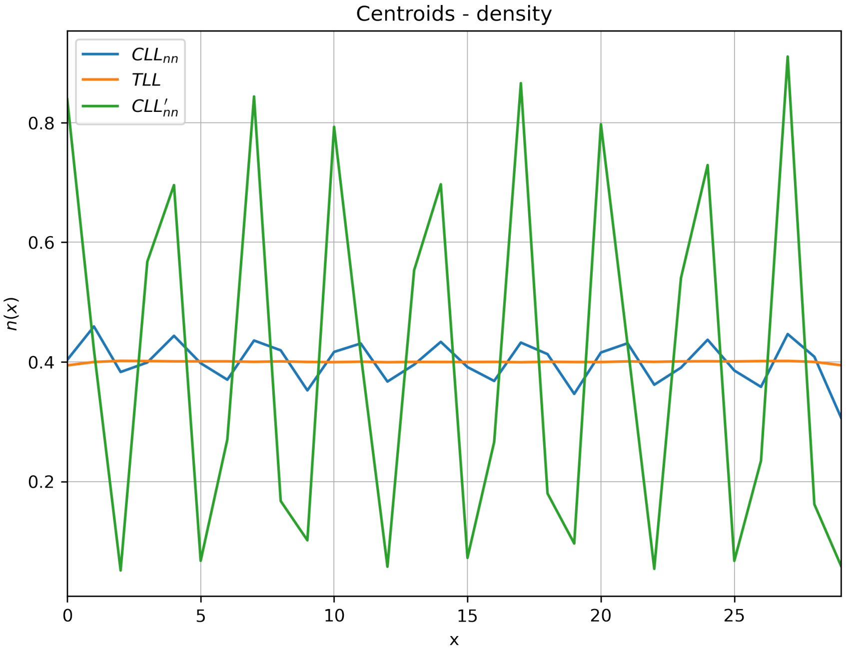


Figure 5.5: Density part of the TLL, CLL_{nn} and CLL'_{nn} centroids.

in Fig. 5.6. The only pattern we can identify is that again when moving to the $V \gtrsim 7$ region the density shifts from a quasi-uniform distribution to a periodically peaked one. Finally, we examine the $U \sim 7$, $V \sim 6-7$ region. Here the system moves from the previously examined CLL_{nn} phase to a very "tiny" phase which we call CLL_d and then falls into the locked (or crystal) phase. From Fig. 5.7 we see that the difference between the CLL_d and the locked regions of the phase diagram is precisely the different nature of the system (i.e. liquid/crystal). While the former presents some peaks in the density but maintains a quasi-liquid structure, the latter is almost perfectly localized in B', A blocks. Due to the perfectly localized nature of this phase, the charge structure factor looks completely flat w.r.t. the other centroids. It should be stressed that it just *looks* flat due to scale reasons, in fact by zooming onto it one can see that it has indeed a sinusoidal shape with a very low peak at $k = 2\pi/3 = 10\pi/15$. Moreover, that shape resembles the one assumed by the spin structure factor in this phase as depicted in Fig. 5.8. Finally, we plot the Spin structure factor part of the centroids for completeness in Fig. 5.9. We see that TLL, CLL_{nn} and CLL_{nn} phases exhibit a well defined peak at $k = 6\pi/15$, CLL_d has a low peak at $k = 4\pi/15$ while $t - CLL_d$, $tt - CLL_d$, $tt - CLL'_d$ and *locked* phases do not exhibit a clear peak, with the latter having a quasi-perfect sinusoidal form.

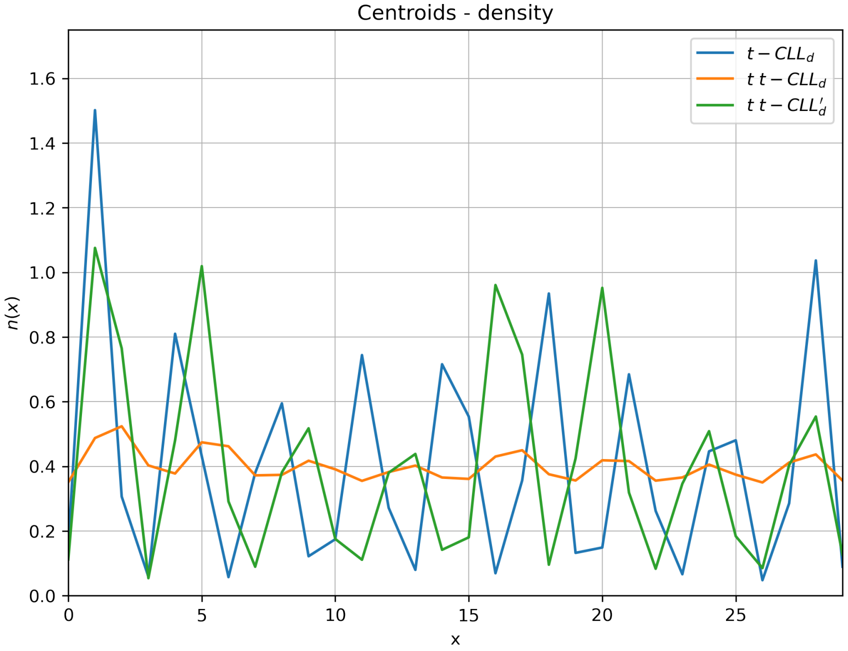


Figure 5.6: Density part of the $t - CLL_d$, $tt - CLL_d$ and $tt - CLL'_d$ centroids.

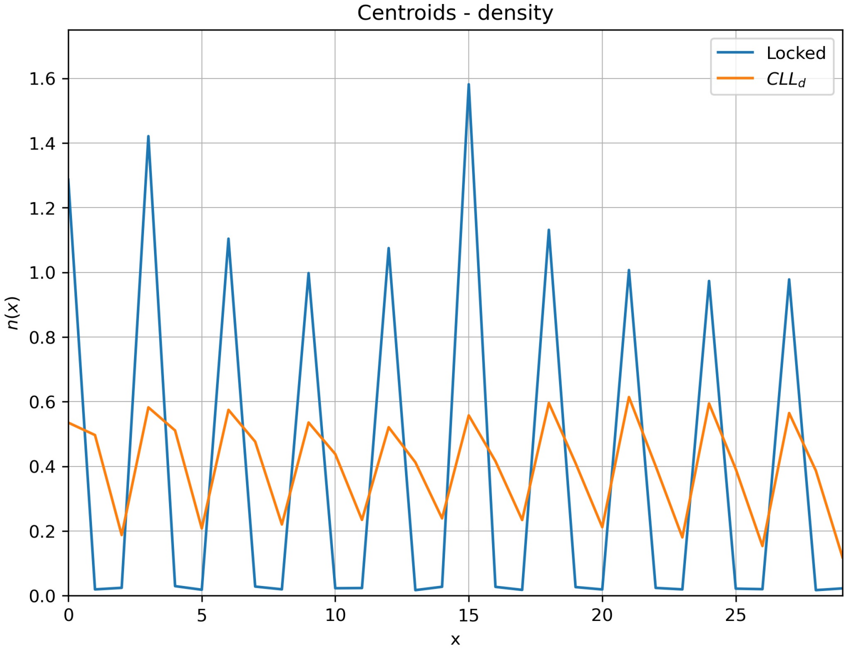


Figure 5.7: Density part of the *locked* and CLL_d centroids.

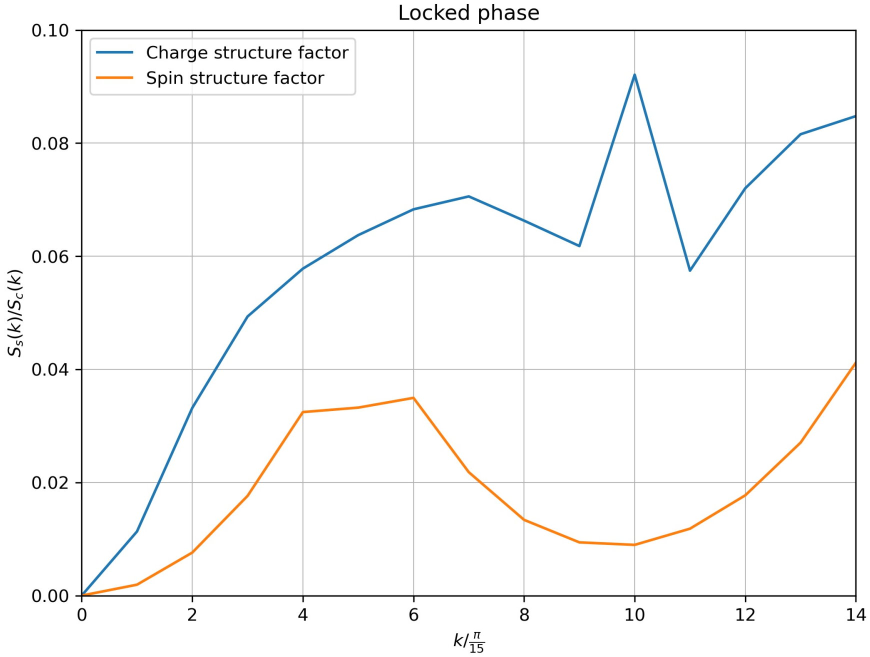


Figure 5.8: Charge and Spin structure factors parts of locked phase centroids.

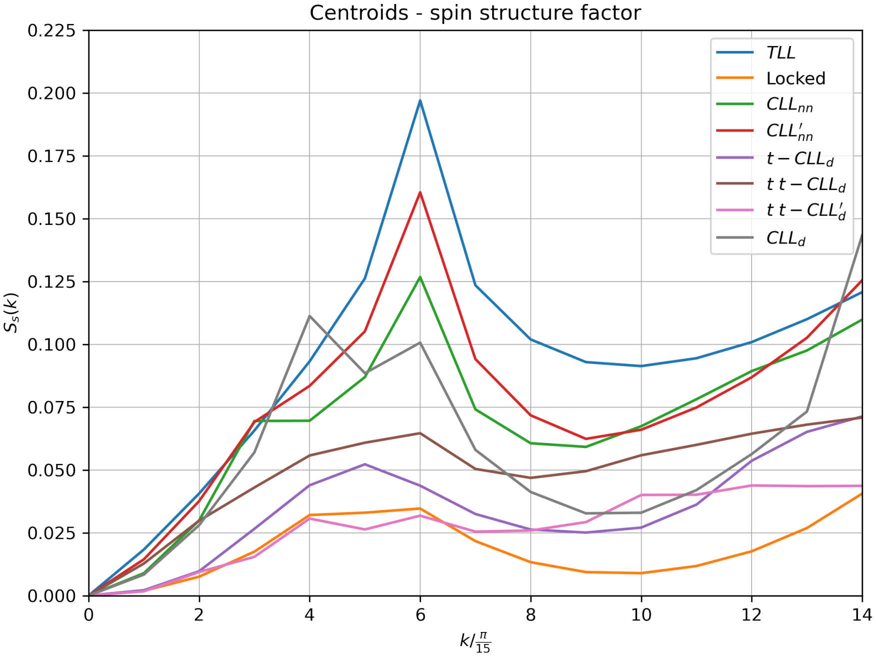


Figure 5.9: Spin structure factor part of the centroids.

5.3 PCA

Eigenvector Number	**Weight**
1	45.7 %
2	24.8 %
3	10.6 %
4	6.0 %
5	5.1 %
6	1.8 %
7	1.6 %
8	1.0 %
Sum	96.6 %

Table 5.1: PCA relative weights

We now apply PCA (Principal Component Analysis) to the same matrix onto which we applied k-means clustering. The relative weight of the first eight [1] eigenvectors are reported in Table 5.1 and Fig. 5.10 while Fig. 5.11, 5.12, 5.13, 5.14 represent the projections of the data corresponding to the different points in the phase diagram onto the eight eigenvectors.
Some comments are in order:

- All the eigenvectors clearly distinguish the *locked* phase from the others.
- The first eigenvector divides the phase diagram at $V \sim 7$ (liquid/crystal distinction).
- No eigenvector is able to catch the TLL to CLL_{nn} transition.
- All eigenvectors projections clearly show the transitions which take place at high values of V for increasing U ($tt-CLL'_d \rightarrow t-CLL_d \rightarrow$ *locked*) previously identified by the k-means algorithm.
- The last two eigenvectors identify the CLL_d phase.

---

[1] There is no precise rule that establishes how many eigenvectors one has to consider; we decided to consider the first eight as the k-means the algorithm identified eight different clusters. The ninth eigenvalue weight is equal to 0.7%, so there is no abrupt decrease (i.e. difference in order of magnitude).

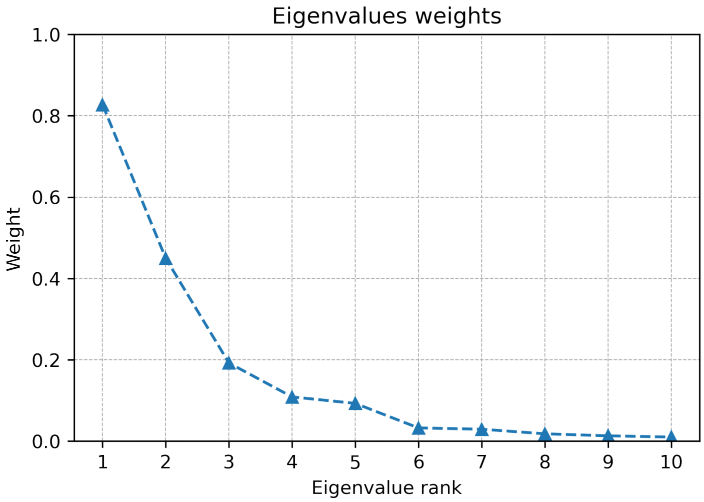


Figure 5.10: Principal components eigenvalues weights. It is manifest that there is no abrupt change of scale between the 8th and 9th eigenvalue.

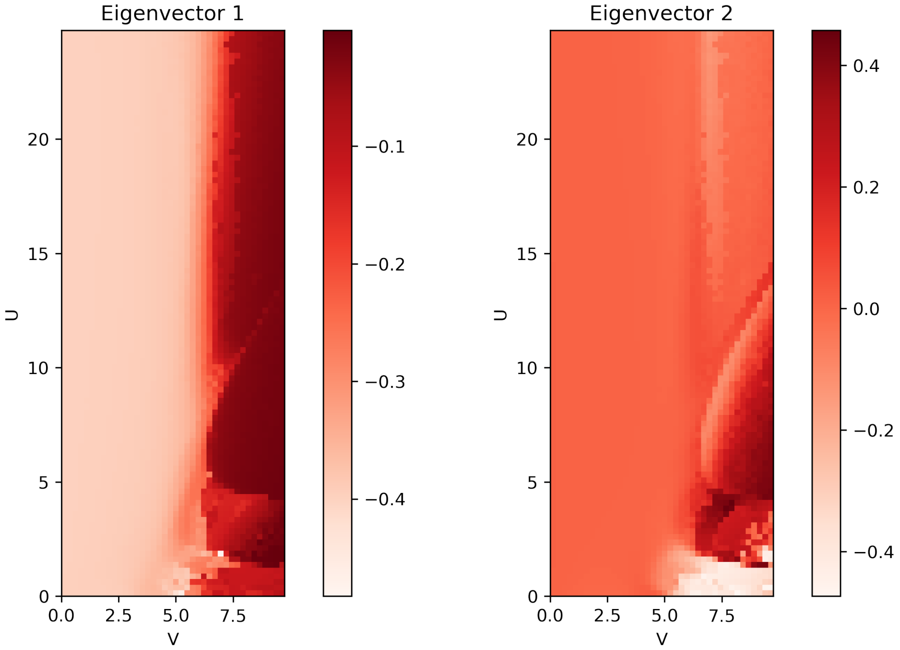


Figure 5.11: Eigenvectors 1-2.

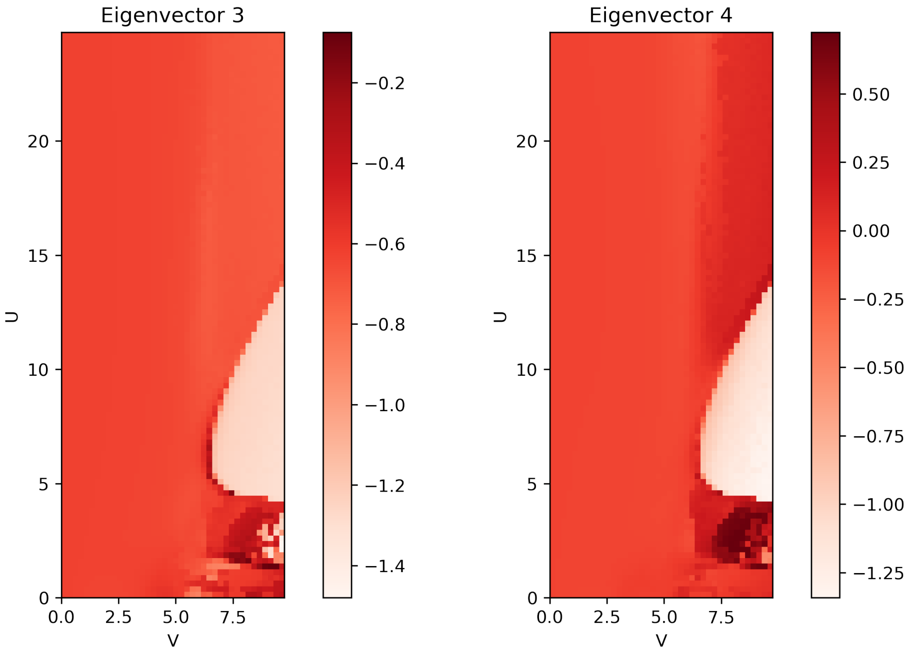


Figure 5.12: Eigenvectors 3-4.

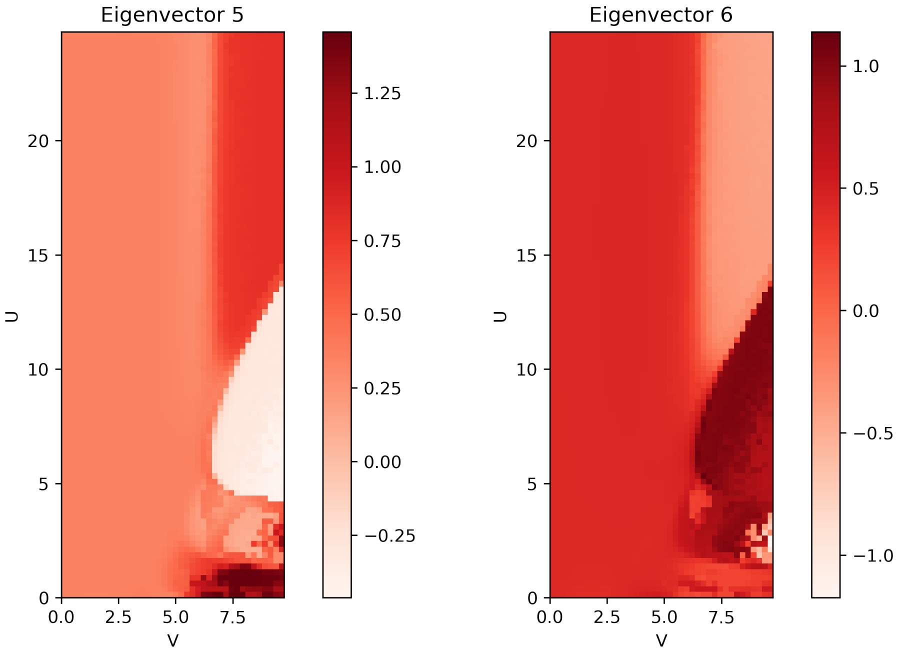


Figure 5.13: Eigenvectors 5-6.

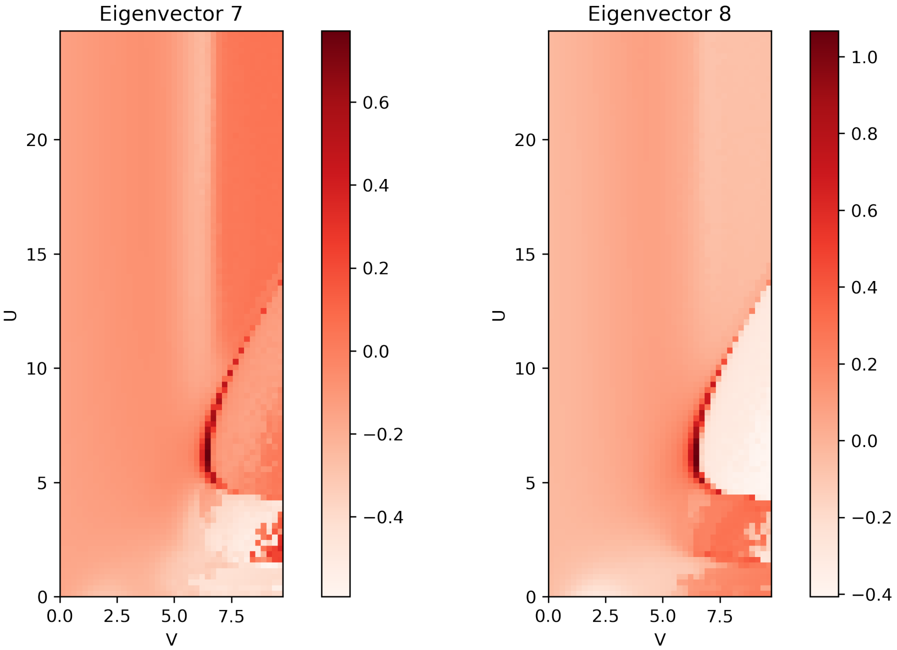


Figure 5.14: Eigenvectors 7-8.

5.4 Learning by Confusion

As we mentioned before, Learning by Confusion is a machine learning technique that was firstly introduced to identify phase transitions in two-states systems [31]. Here we decided to apply it in the sector with $U = 20$ and $V \in [0, 10]$, which shall be a three-states framework. In particular we generated data for 1000 different values of V. As explained in chapter 4 the method consists in training a neural network repeatedly on the same data but with different labels created by assigning a "fake" transition point in a range. In this case we considered a convolutional neural network so our data will be arranged into a 2×30 matrix where the first row is composed of the charge structure factor and the spin structure factor concatenated while the second row is the local density.
Going into details, we start by labelling every point with the same class (i.e. $'1'$). Then we train and test a binary classification convolutional neural network and finally evaluate the accuracy a onto the test set. As expected that is equal to 100%. Once the accuracy is computed we re-label the set by putting the "fake" phase transition at $V = 0.01$, and repeat. In this case we expect the network to perform slightly worse than at the previous step as the labelling is no longer uniform and points belonging to the same phase are labelled differently. This is done until we reach again a uniform set (i.e. all $'0$). At each step the dataset is randomly split into a training set, validation set and test

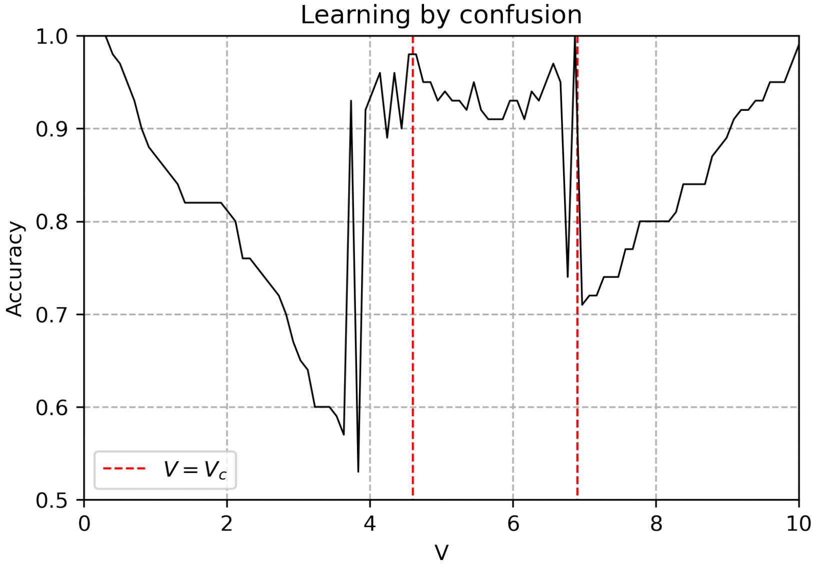


Figure 5.15: Learning by confusion with binary classification at $U = 20$ and varying V. It is clear that this is not the previously mentioned W-shape, but has two maxima at $V_c = 4.6$ and $V_c = 6.9$ (red dotted lines).

set. The accuracy is plotted in Fig. 5.15. We see that the shape of the accuracy is not the previously mentioned W one as it looks more like a *"bridge"* (i.e. a high-accuracy region between two maxima) between two critical values $V_c = 4.6$ and $V_c = 6.9$. This is, in the author's opinion, a first clue of the three-state nature of the system. In fact, in this interval, the network is able to catch that the $'0'$ is associated to the TLL phase (and part of the CLL_{nn} which still confuses it) while $'1'$s are associated to CLL'_{nn}. A second clue that confirms our hypothesis is given by the accuracy plot that we obtain if we repeat the procedure twice by dividing the set into two. Precisely, we applied the procedure proposed in [31] to the two intervals $V \in [0, 6]$ and $V \in [6, 10]$. These are plotted in Fig. 5.16 and clearly show the W-shape, indicating that in these ranges the system encounters two different phases. However, it should be pointed out that the phase transition points found in Fig. 5.15 and Fig. 5.16 are different.

A 3-state generalization is needed. We therefore define a 3-state classifier and change the number of labels to three (i.e. $'0'$, $'1'$ and $'2'$). First, we label the entire set as $'2'$, then, as we did before, we compute the accuracy, move the $1 \to 2$ transition and repeat. Once the procedure is completed we introduce the $'0'$ by labeling the first 10 elements of the set as $'0'$ and repeat the previously explained procedure with labels $'1'$ and $'2'$. From this point the generalization is straightforward. The accuracy heatmap is represented in

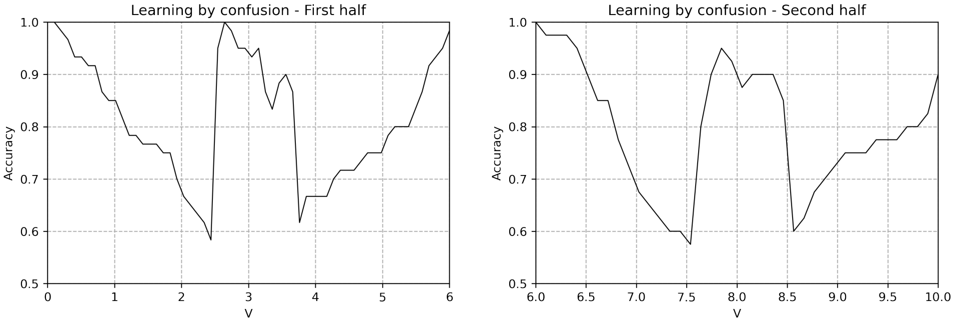


Figure 5.16: Split set accuracies plot

Fig. 5.17. Every line represents a different position of the end of $'0'$ labels, while every

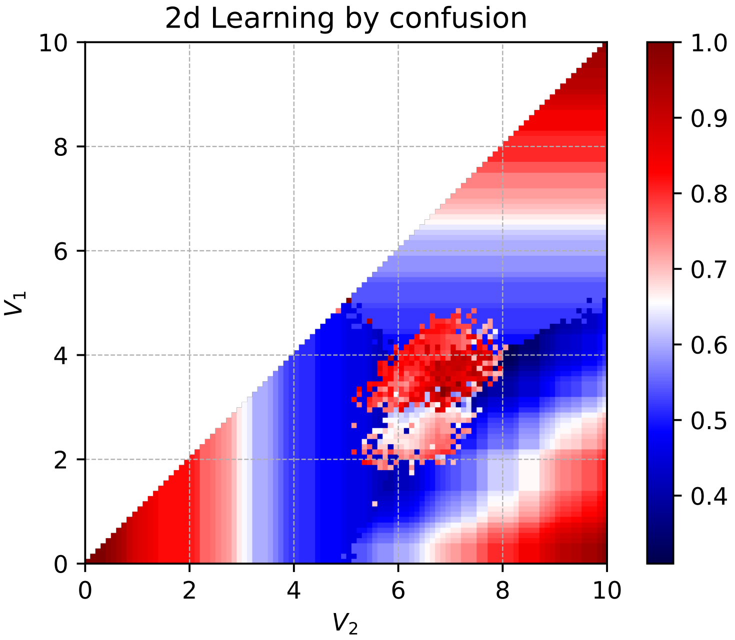


Figure 5.17: Accuracy heat map for the 2-dimensional Learning by confusion with fixed $U = 20$ and increasing value of V. V_1 and V_2 are the first and second phase transitions points respectively. Left corner corresponds to uniform $'1'$ labelling, right corner to uniform $'2'$ labelling, while top corner to uniform $'0'$ labelling.

point corresponds to a different position of the end of $'1'$ labels. The color associated with every point represents the accuracy achieved by the network. As expected the three corners (which correspond to a uniform set) exhibit perfect accuracy. As a confirmation of what we thought, other local maxima are achieved for first phase transition localized at $V_1 \sim 2-4$ and second phase transition at $V_2 \sim 6-8$, in particular at $V_1 = 3.8$ and $V_2 = 7.2$ the accuracy reach its maximum value of 0.98.

We can affirm that this technique confirms what is found by the k-means algorithm. We then repeated the procedure in order to study the $tt-CLL'_d \rightarrow t-CLL_d \rightarrow locked$ phase transitions. We collected data for 1000 values of $U \in [0, 10]$ with fixed $V = 7.5$. Results are plotted in Fig. 5.18. In this case the accuracy plot shows indeed the traditional W-shape, with peak at $V_c = 4.7$, which is extremely compatible with the $t-CLL_d \rightarrow locked$ phase transition identified by k-means clustering. On the other hand the neural network could not distinguish the difference between the $tt - CLL'_d$ and $t - CLL_d$ phases.

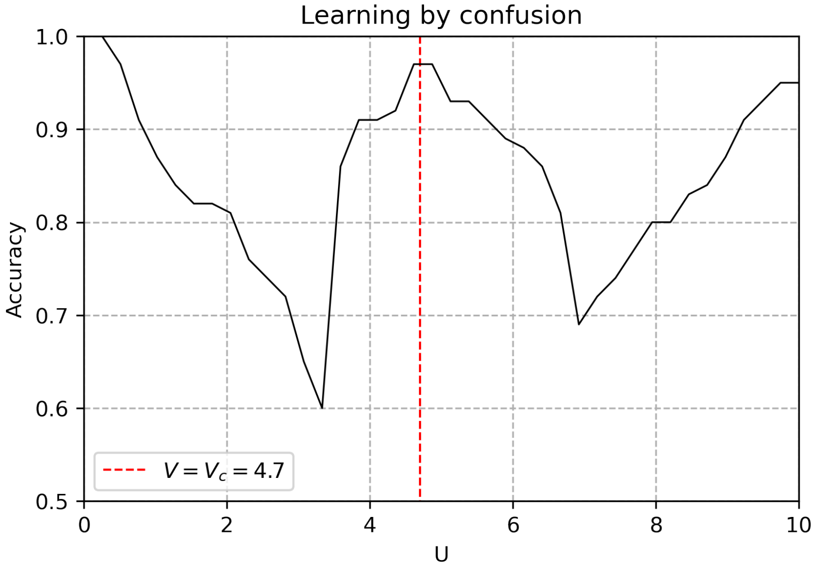


Figure 5.18: Accuracy plot for Learning by confusion with fixed $V = 7.5$ and increasing value of U. We can identify the traditional W-shape with peak at $U = 4.7$ (red dotted line).

5.5 Order parameters

We plot in Fig. 5.19 the number of doubly occupied sites $\mathcal{D}$ and the number of nearest and next-to-nearest neighbors $\mathcal{N}$ for the number of sites L. It is quite manifest that while $\mathcal{N}/L$ changes smoothly as V increases, $\mathcal{D}/L$ changes abruptly for high values of U. In particular it goes from $\mathcal{D}/L \sim 0.13$ (which is equivalent to $\mathcal{D} \simeq 4$ for $L = 30$) for $U \sim 0$, to $\mathcal{D}/L \sim 0.1$ (which is equivalent to $\mathcal{D} \simeq 3$ for $L = 30$) for $U \sim 2-3$, to finally

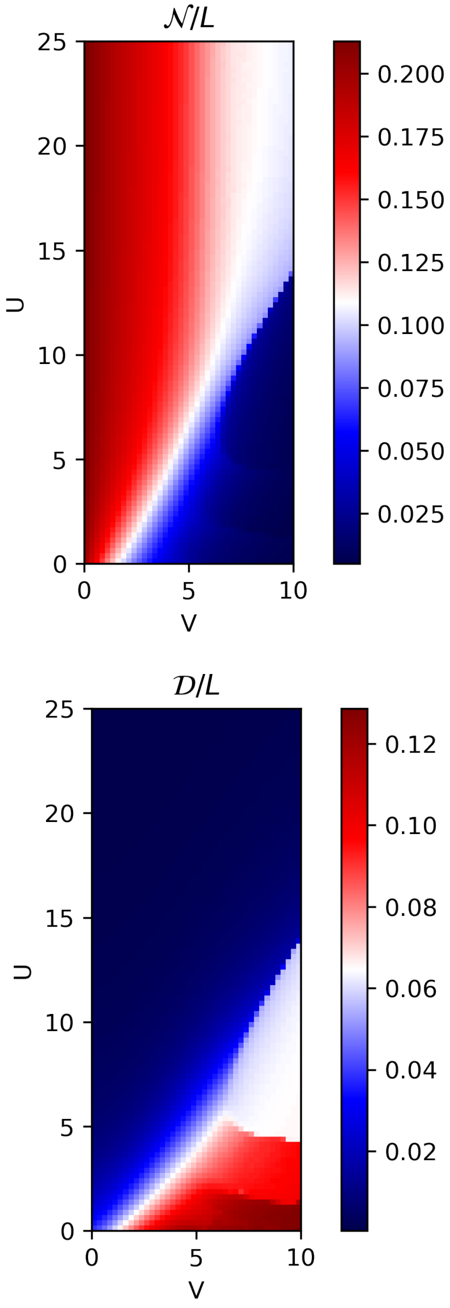


Figure 5.19: Upper panel: nearest and next-to-nearest neighbor number plot. Lower panel: double occupancy number plot. Right panel:

$\mathcal{D}/L \sim 0.07$ (which is equivalent to $\mathcal{D} \simeq 2$ for $L = 30$) for $U \gtrsim 5$. We are now ready to make our hypothesis for the phase diagram of the system.

5.6 Phases of the system

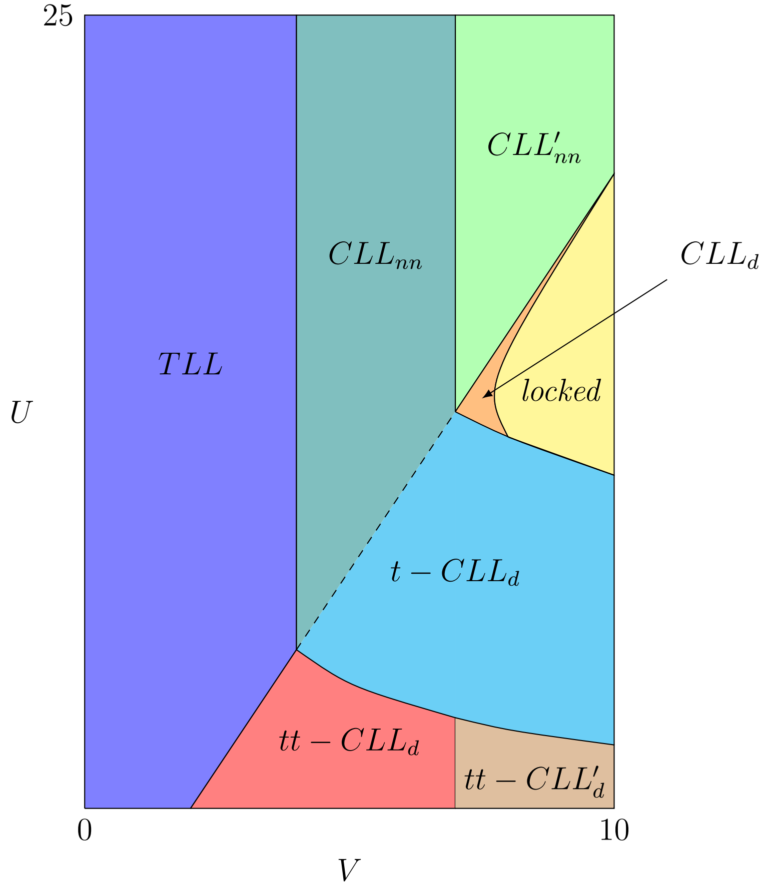


Figure 5.20: The hypothesis of phase diagram for the extended Hubbard model.

Combining the information given by each analysis technique we end up with the hypothesis represented in Fig. 5.20. We shall now examine more carefully the supposed phases of the system.

Tomonaga Luttinger Liquid - TLL

As we did in the k-means analysis we start by looking at the high-U regime, i.e. $U \sim 20$. From k-means we have seen that for $V \in [0,4]$ the density is uniform $n(x) = \frac{2}{5}\ \forall x$, the charge density structure factor does not exhibit a clear peak. This agrees with what was said when dealing with bosonization, that is for low values of V the system is a Tomonaga Luttinger Liquid (see Fig. 5.21). As V grows the system encounters a phase transition to CLL_{nn} (which we will examine later). For small values of U the transition takes place for $V \sim U$, as found by classical arguments in Eq. (2.86). When U grows this limit is no longer valid, as the fundamental hypothesis for perturbation theory was that V was small to safely expand the cosine. For $U \gtrsim 5$ we found that the phase transition threshold is fixed at $V \sim 4$ (that was firstly found by the k-means algorithm and then confirmed by Learning by confusion).

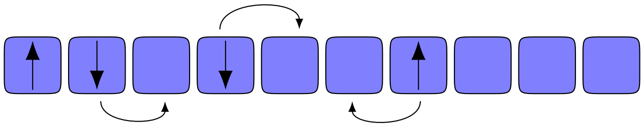


Figure 5.21: Pictorial representation of Tomonaga Luttinger Liquid (TLL). There is no block structure and single-particles mobility is extremely high.

Nearest neighbor liquid of clusters - CLL_{nn}

As pointed out in Sect. 3.2, in order to apply phenomenological bosonization it is not necessary for fundamental entities to be particles, as they can also be clusters of particles. When $U \gg V$ (in particular if we consider the classical ground state as guiding start when $U > 3V/2$) the clusters which compose the liquid are of type B and A (as depicted in Fig. 5.22). In this case, it is easy to see from simple geometrical considerations that one should expect $M/L = 3/10$, so the charge structure factor shall exhibit a peak at $k_c^{CLL_{nn}} = 3\pi/5 = 9\pi/15$, the double occupancy shall be zero and the density shall be quasi-uniform (liquid phase). This is in total agreement with what found by k-means algorithm (both the density and the structure factor) and in [27]. As V grows the mobility is progressively reduced until the critical value of $V \sim 7$, where the system undergoes a new phase transition to a crystal. This second transition localizes the particles and was found by k-means, PCA, and Learning by confusion. Fixing V at an intermediate value (e.g. $V \sim 5-6$) and decreasing U the system undergoes another transition to the $t-CLL_d$ phase, which we will examine later. It should be stressed that the line of this phase transition was not uniquely determined by our analysis, as k-means and $\mathcal{D}$ data were in light contradiction (the first suggests that $t - CLL_d$ is realized only for $V \gtrsim 7$ while looking at $\mathcal{D}$ plot it seems that the phase transition takes place at $U \sim V$).

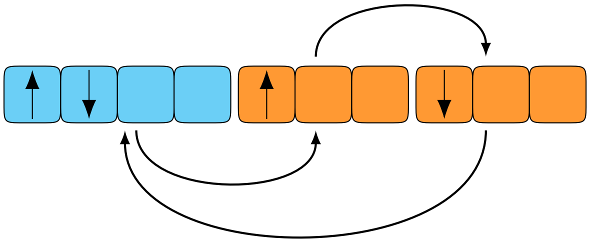


Figure 5.22: Pictorial representation of Cluster Luttinger Liquid of nearest neighbours (CLL_{nn}). The system is arranged in A and B blocks, which are the elementary units of it.

Crystal *"nn"* - CLL'_{nn}

When V grows mobility is reduced (every movement risks to cost V in terms of energy) but not eliminated. Furthermore as $U < 3V/2$, B blocks are favored w.r.t. B' blocks. This is in agreement with strong coupling perturbation theory and with [27] (even if in [27] the threshold of the liquid-crystal transition was set to higher values, i.e. $V \sim 10$). This phase, represented in Fig. 5.23, shall exhibit a more peaked density and a flatter charge structure factor, in complete agreement with what was found by the k-means algorithm.

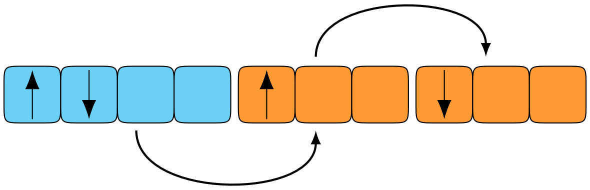


Figure 5.23: Pictorial representation of crystal CLL'_{nn}. The block strucutre is identical to CLL_{nn}, but mobility is reduced due to the high value of V.

Doublon liquid of clusters - CLL_d

When dealing with smaller values of U, i.e. $U \sim 5 - 10$, double occupancy is no longer forbidden and that is the origin of the emergence of new phases. In particular, when

$U \sim 5 - 10$ and $U \sim 3V/2$, the system undergoes a transition between CLL_{nn} (or $t-CLL_d$) to CLL_d. Similar to CLL_{nn} this is a cluster liquid phase. The main difference is that in this case the blocks that compose the system are no longer A and B but A and B', as depicted in Fig. 5.24. The difference of length of B' blocks w.r.t. B blocks (the former are r_c+1 long while the latter are r_c+2) implies a different number of blocks. In particular, the peak of the charge structure factor shall now be localized at $k_c^{CLL_d} = 2\pi/3$. It must be underlined that this phase is expected to emerge when $V < 2U/3$, but we must remember that when dealing with perturbation theory we found that when $V \gg t$ (in particular for $V/t \gtrsim 7$) the crystalline solution should also be valid. This competition between the crystalline and liquid solution is one of the main aspects of interest in this work and is the main reason for CLL_d smallness in the phase diagram. CLL_d was found by k-means and confirmed by the last eigenvectors of PCA, whose projections underlined this phase.

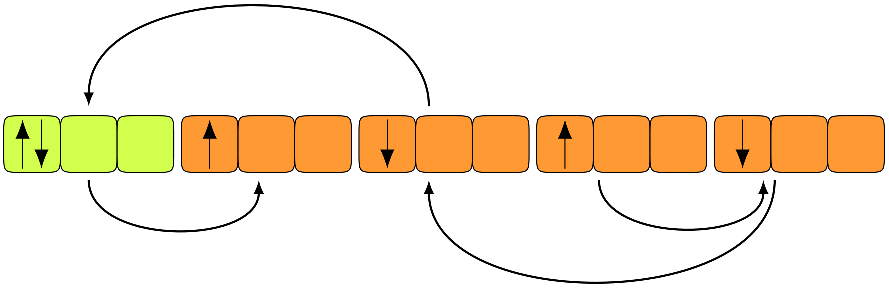


Figure 5.24: Pictorial representation of Cluster Luttinger Liquid of doublons (CLL_d). The system is composed of blocks A and B' and mobility is not quenched.

Crystal *"locked"* phase

As $V \to \infty$, double occupied sites are favored w.r.t to nearest neighbors, so the system arranges in blocks B' and A. The difference between this state and the $U \gg V \gg t$ one is that now mobility is completely quenched. It is easy to see that moving a particle from a doubly occupied state would have an energy cost of V or $2V$. Particles are therefore *locked* into a perfect crystalline phase, which is represented in Fig. 5.25. This shall correspond to a flat charge structure factor and perfectly peaked density. The substantial difference between this state and all the others is identified by all the techniques we used.

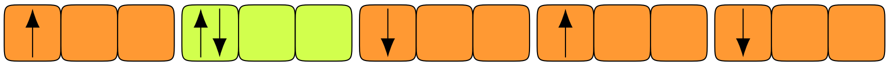


Figure 5.25: Representation of the "locked" phase; the block structure is identical to CLL_d but now mobility is almost reduced to zero.

$tt - CLL_d$ cluster liquid

For small values of U, V, i.e. $U \sim 0-2$, $V \sim 2-7$, the single particle mobility is not negligible. From k-means clustering, we see that in this region of the phase diagram the system arranges itself into $M = 8$ blocks. This, combined with the value of the order parameter $\mathcal{D} \sim 4$, suggests the structure in the figure below. The fact that $L = 30$ and $M = 8$ implies that there should be 6 blocks of length 4 and 2 blocks of length 3. The two blocks of length 2 shall be B' blocks, which are energetically favored due to the small value of U. The other two doubly occupied sites are arranged into two length 4 blocks. Finally, the remaining particles are arranged into the other blocks. The extra site in this blocks w.r.t. to A blocks favor the single particle mobility, which can be done *"for free"* (i.e. there is no cost V). A representation of this phase is depicted in Fig. 5.26.

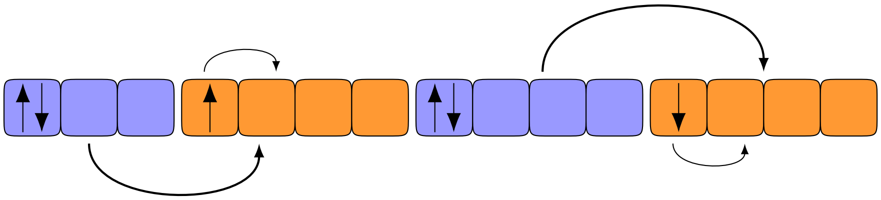


Figure 5.26: Representation of $tt - CLL_d$ phase. As U and V have intermediate values, there is not a clear adavantage of B blocks w.r.t. to B' blocks (or viceversa). That, combined to the fact that single particle mobility is still relevant (single particle movement is here represented with a thinner arrow), is the cause of this new phase of the system.

$tt - CLL'_d$ *locked* liquid

As V crosses the previously seen critical value of $V \sim 7$, mobility is reduced but the underlying structure is not modified, as shown in Fig. 5.27. That is underlined by both PCA and k-means (the number of blocks does not change, the density shape does).

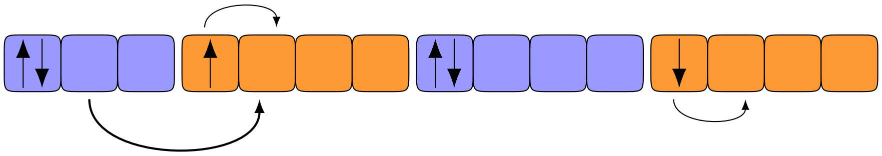


Figure 5.27: $tt-CLL'_d$ representation; the block structure is identical to the $tt-CLL_d$ one, but now mobility is reduced due to the high value of V.

$t-CLL_d$ cluster liquid

When U grows to higher values (i.e. $U \sim 2-4$) the number of blocks changes to $M=9$ and the doubly occupied sites become $\mathcal{D}=3$. This suggest a structural change w.r.t. the $tt-CLL_d$ phase. For $L=30$ in fact, one can suppose that the system arranges itself into 6 blocks of length 3 and 3 blocks of length 4. Three of the former shall be B' blocks while the other shall be A blocks. On the other hand, the length 4 blocks shall contain just one particle to favor single particle mobility. This structure is represented in Fig. 5.28.

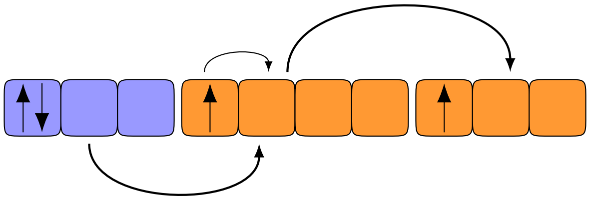


Figure 5.28: Representation of $t-CLL_d$ phase. As U increases double occupancies are energetically less favourable and reduces to $\mathcal{D}=3$.

Chapter 6

Conclusions

"In case I'm wrong. It has happened."

Dr. House

We shall now briefly review the main outcomes of our work.

Competition between the different terms of the Hamiltonian and frustration of particles play a fundamental role in building the phases of the system and make them extremely hard to determine. Given this, Machine Learning has proved itself to be extremely useful to tackle this task. Moreover, Machine Learning techniques used in this work demonstrated to be complementary, as PCA and k-means clustering are more useful to get a general impression of the phase diagram while Learning by Confusion is useful in order to determine precisely the transition points.

Thanks to k-means clustering we saw that the phase diagram of the extended Hubbard model shall be richer than what predicted in [6]. In particular we found that CLL_{nn} and CLL_d appear for $U \gtrsim 5$ and $V \gtrsim 5$, but the latter occupies a very thin space of the phase diagram, as the system falls into the *locked* crystalline phase as soon as we move from the classical phase transition line $U \simeq 3V/2$. Moreover, we found that Tomonaga Luttinger Liquid of doublons TLL_d is not a phase of the system, as the low-U region of the phase diagram presents three different phases: $tt - CLL_d$, $tt - CLL'_d$ and $t - CLL_d$. This region was the most difficult to analyze, as, due to the high value of V and low value of U, it is possible that, during Density Matrix Renormalization Group simulations, the system can "fall" into local energy minima difficult to overcome. A simple example to think of can be the one of a doubly occupied site from which it is difficult for a particle to exit (as soon as it goes out from the doubly occupied site the energy increases of V). Finally we found that for $V \gtrsim 7$, which was not analyzed in [6] due to difficult numerical convergence, all the phases undergo a sort of crystallization, as particles become more

and more localized. This crystallization is indicated by the prime subscript (e.g. CLL_{nn} to CLL'_{nn}).
As no physical work is, in a sense, never finished, we indicate now three possible directions along which one could operate in order to confirm what found:

1. First of all one should try to simulate the system for bigger systems. In this thesis, we just examined the case of $L = 30$ due to limited computational resources (and time). When enlarging the chain one should always keep in mind that it is important to maintain a length that allows both the classical ground state. As the blocks A and B' are $r_c + 1$ long while blocks B are $r_c + 2$ long, the total length of the system shall be a multiple of $3r_c + 4$ and $r_c + 1$. In our case, that is $r_c = 2$, that is being a multiple of 10 and 3 (i.e. $L = 60, 90, \dots$). Enlarging the system should limit finite-size effects.

2. The second possibility to verify our hypothesis is to modify boundary conditions. In our study, we limited ourselves to PBC (periodic boundary conditions), which limit finite-size effects but can reduce the DMRG precision. One should therefore generalize the study to anti-periodic and open boundary conditions, which better fits the density matrix renormalization group.

3. Finally one can go to higher values of the maximum bond dimension. We limited ourselves to 400 due to limited computational resources (that should be sufficient for $1d$ systems) but increasing it to e.g. 1000 should make the simulation more precise. It shall be stressed that this is simply a simulation parameter and has no physical relevance (differently to the other refinements).

Acknowledgements

First of all, I must thank prof. Elisa Ercolessi, who introduced me to Statistical Mechanics and, in particular, this beautiful topic. Her suggestions have been fundamental throughout the entire work, which would not have been possible without her help (I also seize the opportunity to thank her for having given me her copy of the *Hubbard model and Anyon Superconductivity*).

The realization of this work was possible also thanks to Federico dell'Anna (who helped me with iTensor), Simone Tibaldi (who guided me through the secret magic of deep learning) and Sunny Pradhan (who helped me both with theoretical and numerical work and whose ability to use Linux still amazes me).

Even if they were not directly involved in this thesis I also want to thank all the professors that have taught me during my Bachelor's and Master's degree. In particular, I want to thank prof. Maltoni and prof. Remiddi; your passion and your work represent a real and deep inspiration.

Physics is not a simple subject and can sometimes be discouraging. Paraphrasing what Guido Tonelli said last July about the qualities of a good researcher, I think that a good physicist shall have the ability to get up after being pulled down by difficulties. If I managed to do so during these years is only because I was lucky enough to find classmates that made these years a lot easier.

I can't find the right words to thank my friends Domingo, Luca, Giovanni, Viola, Francesca and Clara, so I'll only say that you're like a second family to me, and thank you for bearing my being know-it-all.

I want to thank my family for their support and presence and I have to thank particularly my parents for encouraging me to follow my dreams and my brother for our nice chats about topics we are fond of.

Finally, I want to thank Serena; you cheered me up when I was down and believed in me even when I could not.

Bibliography

[1] Balachandran, A. P., Ercolessi, E., Morandi, G., and Srivastava, A. M. *Hubbard model and anyon superconductivity*. Vol. 38. World Sci.Lect.Notes Phys. 1990.

[2] Bardeen, J., Cooper, L. N., and Schrieffer, J. R. "Microscopic Theory of Superconductivity". In: *Phys. Rev.* 106 (1 Apr. 1957), pp. 162–164. DOI: `10.1103/PhysRev.106.162`.

[3] Bardeen, J., Cooper, L. N., and Schrieffer, J. R. "Theory of Superconductivity". In: *Phys. Rev.* 108 (5 Dec. 1957), pp. 1175–1204. DOI: `10.1103/PhysRev.108.1175`.

[4] Blaiszik, B. *Charting ML Publications in Science*. 2021. URL: `https://github.com/blaiszik/ml_publication_charts#charting-ml-publications-in-science`.

[5] Botzung, T. "Study of strongly correlated one-dimensional systems with long-range interactions". PhD thesis. 2019. URL: `http://www.theses.fr/2019STRAF062`.

[6] Botzung, T., Pupillo, G., Simon, P., Citro, R., and Ercolessi, E. *One-dimensional extended Hubbard model with soft-core potential*. 2019. DOI: `10.48550/ARXIV.1909.12168`.

[7] Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., Vogt-Maranto, L., and Zdeborová, L. "Machine learning and the physical sciences". In: *Reviews of Modern Physics* 91.4 (Dec. 2019). DOI: `10.1103/revmodphys.91.045002`.

[8] Cooper, L. N. "Bound Electron Pairs in a Degenerate Fermi Gas". In: *Phys. Rev.* 104 (4 Nov. 1956), pp. 1189–1190. DOI: `10.1103/PhysRev.104.1189`.

[9] Dantas, V. and Andrade, E. C. "Disorder, Low-Energy Excitations, and Topology in the Kitaev Spin Liquid". In: *Physical Review Letters* 129.3 (July 2022). DOI: `10.1103/physrevlett.129.037204`.

[10] Dawid, A. et al. *Modern applications of machine learning in quantum sciences*. 2022. DOI: `10.48550/ARXIV.2204.04198`.

[11] Epstein, A., Etemad, S., Garito, A., and Heeger, A. "Metal-insulator transition in an organic solid: Experimental realization of the one-dimensional Hubbard model". In: *Solid State Communications* 9.21 (1971), pp. 1803–1808. DOI: https://doi.org/10.1016/0038-1098(71)90094-9.

[12] Fazzini, S. "Non-local orders in Hubbard-like low dimensional systems". PhD thesis. Politecnico di Torino, Feb. 2018. URL: http://hdl.handle.net/11583/2703392.

[13] Fishman, M., White, S. R., and Stoudenmire, E. M. *The ITensor Software Library for Tensor Network Calculations*. 2020. arXiv: 2007.14822.

[14] Giamarchi, T. *Quantum Physics in One Dimension*. International Series of Monographs on Physics. Clarendon Press, 2004, p. 75. URL: https://books.google.it/books?id=1MwTDAAAQBAJ.

[15] Giamarchi, T. *Quantum Physics in One Dimension*. International Series of Monographs on Physics. Clarendon Press, 2004. URL: https://books.google.it/books?id=1MwTDAAAQBAJ.

[16] Glocke, S., Klümper, A., and Sirker, J. "Half-filled one-dimensional extended Hubbard model: Phase diagram and thermodynamics". In: *Physical Review B* 76.15 (Oct. 2007). DOI: 10.1103/physrevb.76.155121.

[17] Gurney, K. *An Introduction to Neural Networks*. USA: Taylor and Francis, Inc., 1997.

[18] Hubbard, J. "Electron correlations in narrow energy bands". In: *Proc. R. Soc. Lond. A* 276 (1963), pp. 238–257. DOI: https://doi.org/10.1098/rspa.1963.0204.

[19] Kitaev, A. Y. "Unpaired Majorana fermions in quantum wires". In: *Physics-Uspekhi* 44.10S (Oct. 2001), pp. 131–136. DOI: 10.1070/1063-7869/44/10s/s29.

[20] Korepin, E. F. G. K. *The one-dimensional Hubbard model*. 2003. URL: http://insti.physics.sunysb.edu/~korepin/PDF_files/Hubbard.pdf.

[21] Krizhevsky, A., Sutskever, I., and Hinton, G. E. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems*. Ed. by F. Pereira, C. Burges, L. Bottou, and K. Weinberger. Vol. 25. Curran Associates, Inc., 2012. URL: https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

[22] Lieb, E. H. and Wu, F. "The one-dimensional Hubbard model: a reminiscence". In: *Physica A: Statistical Mechanics and its Applications* 321.1 (2003). Statphys-Taiwan-2002: Lattice Models and Complex Systems, pp. 1–27. DOI: https://doi.org/10.1016/S0378-4371(02)01785-5.

[23] Long, M. W., Castleton, C. W. M., and Hayward, C. A. "The one-dimensional Hubbard model with first-, second- and third-nearest-neighbour hopping in the strong-coupling limit". In: *Journal of Physics: Condensed Matter* 6.2 (Jan. 1994), pp. 481–493. DOI: 10.1088/0953-8984/6/2/019.

[24] MacKay, D. J. C. *Information Theory, Inference, and Learning Algorithms.* Copyright Cambridge University Press, 2003.

[25] Malard, M. "Sine-Gordon Model: Renormalization Group Solution and Applications". In: *Brazilian Journal of Physics* 43.3 (2013), pp. 182–198. DOI: 10.1007/s13538-013-0123-4.

[26] Mandal, S. and Jayannavar, A. M. *An introduction to Kitaev model-I.* 2020. DOI: 10.48550/ARXIV.2006.11549.

[27] Mattioli, M., Dalmonte, M., Lechner, W., and Pupillo, G. "Cluster Luttinger Liquids of Rydberg-Dressed Atoms in Optical Lattices". In: *Physical Review Letters* 111.16 (Oct. 2013). DOI: 10.1103/physrevlett.111.165302.

[28] Mott, N. F. "The Basis of the Electron Theory of Metals, with Special Reference to the Transition Metals". In: *Proceedings of the Physical Society. Section A* 62.7 (July 1949), pp. 416–422. DOI: 10.1088/0370-1298/62/7/303.

[29] Mott, N. F. and Peierls, R. "Discussion of the paper by de Boer and Verwey". In: *Proceedings of the Physical Society* 49.4S (Aug. 1937), pp. 72–73. DOI: 10.1088/0959-5309/49/4s/308.

[30] Ng, K.-K. and Yang, M.-F. "Frustration-induced incommensurate solids in the extended Bose-Hubbard model". In: *Phys. Rev. B* 103 (12 Mar. 2021), p. 125151. DOI: 10.1103/PhysRevB.103.125151.

[31] Nieuwenburg, E. P. L. van, Liu, Y.-H., and Huber, S. D. "Learning phase transitions by confusion". In: *Nature Physics* 13.5 (Feb. 2017), pp. 435–439. DOI: 10.1038/nphys4037.

[32] Onsager, L. "Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition". In: *Phys. Rev.* 65 (3-4 Feb. 1944), pp. 117–149. DOI: 10.1103/PhysRev.65.117.

[33] Pathria, R. *Statistical Mechanics.* International series of monographs in natural philosophy. Elsevier Science & Technology Books, 1972, p. 195. URL: https://books.google.it/books?id=ejUVnQEACAAJ.

[34] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

[35] Ruijl, B., Ueda, T., and Vermaseren, J. *FORM version 4.2*. 2017. DOI: 10.48550/ARXIV.1707.06453.

[36] Salfi, J., Mol, J., Rahman, R., Klimeck, G., Simmons, M., Hollenberg, L., and Rogge, S. "Quantum Simulation of the Hubbard Model with Dopant Atoms in Silicon". In: *Nature Communications* 7 (Apr. 2016), p. 11342. DOI: 10.1038/ncomms11342.

[37] Scalapino, D. J. "Numerical Studies of the 2D Hubbard Model". In: *Handbook of High Temperature Superconductivity*. Springer, 2006.

[38] Schollwöck, U. "The density-matrix renormalization group". In: *Reviews of Modern Physics* 77.1 (Apr. 2005), pp. 259–315. DOI: 10.1103/revmodphys.77.259.

[39] Schollwöck, U. "The density-matrix renormalization group in the age of matrix product states". In: *Annals of Physics* 326.1 (Jan. 2011), pp. 96–192. DOI: 10.1016/j.aop.2010.09.012.

[40] Schulz, H. J., Cuniberti, G., and Pieri, P. *Fermi liquids and Luttinger liquids*. 1998. DOI: 10.48550/ARXIV.COND-MAT/9807366.

[41] Tarruell, L. and Sanchez-Palencia, L. "Quantum simulation of the Hubbard model with ultracold fermions in optical lattices". In: (2018). DOI: 10.48550/ARXIV.1809.00571.

[42] Tibaldi, S., Magnifico, G., Vodola, D., and Ercolessi, E. *Unsupervised and supervised learning of interacting topological phases from single-particle correlation functions*. 2022. DOI: 10.48550/ARXIV.2202.09281.

[43] Tremblay, A. .-. S. *Strongly correlated superconductivity*. 2013. DOI: 10.48550/ARXIV.1310.1481.

[44] Verstraete, F., Porras, D., and Cirac, J. I. "Density Matrix Renormalization Group and Periodic Boundary Conditions: A Quantum Information Perspective". In: *Physical Review Letters* 93.22 (Nov. 2004). DOI: 10.1103/physrevlett.93.227205. URL: https://doi.org/10.1103%2Fphysrevlett.93.227205.

[45] Wang, L. "Discovering phase transitions with unsupervised learning". In: *Phys. Rev. B* 94 (19 Nov. 2016), p. 195105. DOI: 10.1103/PhysRevB.94.195105.

[46] White, S. R. and Noack, R. M. “Real-space quantum renormalization groups”. In: *Phys. Rev. Lett.* 68 (24 June 1992), pp. 3487–3490. DOI: `10.1103/PhysRevLett.68.3487`.

[47] White, S. R. “Density-matrix algorithms for quantum renormalization groups”. In: *Phys. Rev. B* 48 (14 Oct. 1993), pp. 10345–10356. DOI: `10.1103/PhysRevB.48.10345`.

[48] Zhang, Q., Zhang, M., Chen, T., Sun, Z., Ma, Y., and Yu, B. *Recent Advances in Convolutional Neural Network Acceleration*. 2018. DOI: `10.48550/ARXIV.1807.08596`.

[49] Zhang, W. and Zeng, Z. *Research Progress of Convolutional Neural Network and its Application in Object Detection*. 2020. DOI: `10.48550/ARXIV.2007.13284`.

[50] Ziman, J. M. *Principles of the Theory of Solids*. 2nd ed. Cambridge University Press, 1972. DOI: `10.1017/CBO9781139644075`.