NEURAL-NETWORK QUANTUM STATES FOR A TWO-LEG BOSE-HUBBARD LADDER UNDER A SYNTHETIC MAGNETIC FIELD

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

PHYSICS

By

Kadir Çeven

July 2023

Neural-network quantum states for a two-leg Bose-Hubbard ladder under a synthetic magnetic field
By Kadir Çeven
July 2023

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Mehmet Özgür Oktel(Advisor)

Ahmet Keleş(Co-Advisor)

Ceyhun Bulutay

Rifat Onur Umucalılar

Approved for the Graduate School of Engineering and Science:

Orhan Arıkan
Director of the Graduate School

ABSTRACT

NEURAL-NETWORK QUANTUM STATES FOR A TWO-LEG BOSE-HUBBARD LADDER UNDER A SYNTHETIC MAGNETIC FIELD

Kadir Çeven
M.S. in Physics
Advisor: Mehmet Özgür Oktel
Co-Advisor: Ahmet Keleş
July 2023

This thesis explores novel quantum phases in a two-leg Bose-Hubbard ladder, achieved using neural-network quantum states. The remarkable potential of quantum gas systems for analog quantum simulation of strongly correlated quantum matter is well-known; however, it is equally evident that new theoretical bases are urgently required to comprehend their intricacies fully. While simple one-dimensional models have served as valuable test cases, ladder models naturally emerge as the next step, enabling studying higher dimensional effects, including gauge fields. Utilizing the paper [Çeven *et al.*, Phys. Rev. A **106**, 063320 (2022)], this thesis investigates the application of neural-network quantum states to a two-leg Bose-Hubbard ladder in the presence of strong synthetic magnetic fields. This paper showcased the reliability of variational neural networks, such as restricted Boltzmann machines and feedforward neural networks, in accurately predicting the phase diagram exhibiting superfluid-Mott insulator phase transition under strong interaction. Moreover, the neural networks successfully identified other intriguing many-body phases in the weakly interacting regime. These exciting findings firmly designate a two-leg Bose-Hubbard ladder with magnetic flux as an ideal testbed for advancing the field of neural-network quantum states. By expanding these previous results, this thesis contains various essential aspects, including a comprehensive introduction and analysis of the vanilla Bose-Hubbard model and the two-leg Bose-Hubbard ladder under magnetic flux, an in-depth overview of neural-network quantum states tailored for bosonic systems, and a thorough presentation and analysis of the obtained results using neural-network quantum states for these two Bose-Hubbard models.

Keywords: Bose-Hubbard model, two-leg ladder flux system, superfluid phase,

Mott insulator phase, synthetic magnetic field, neural-network quantum states, machine learning, artificial neural networks, restricted Boltzmann machine, feed-forward neural network.

ÖZET

SENTETİK MANYETİK AKI ALTINDAKİ İKİ-BACAKLI BOSE-HUBBARD MERDİVENİ İÇİN SİNİR-AĞ KUANTUM DURUMLARI

Kadir Çeven
Fizik, Yüksek Lisans
Tez Danışmanı: Mehmet Özgür Oktel
İkinci Tez Danışmanı: Ahmet Keleş
Temmuz 2023

Bu tez sinir-ağ kuantum durumlarını kullanarak iki-bacaklı Bose-Hubbard merdiveninde özgün kuantum fazlarını araştırmaktadır. Kuantum gaz sistemlerinin, güçlü korelasyonlu kuantum maddelerinin analog kuantum simülasyonu için olağanüstü potansiyeli iyi bilinmektedir; fakat aynı şekilde, bunların sebep olduğu karmaşıklıkları tam olarak anlamak için yeni teorik temellere acil olarak ihtiyaç duyulduğu açıktır. Basit bir boyutlu modeller önemli test durumları olarak bilinirken, doğal olarak merdiven modelleri bir sonraki adım olarak ortaya çıkmakta ve ayar alanları da dahil olmak üzere daha yüksek boyutlu etkilerin incelenmesini sağlamaktadır. [Çeven *et al.*, Phys. Rev. A **106**, 063320 (2022)] makalesinden istifade eden bu tez, güçlü sentetik manyetik alanlar varlığında iki-bacaklı Bose-Hubbard merdivenine sinir-ağı kuantum durumlarının uygulanmasını araştırmaktadır. Bu makale, kısıtlı Boltzmann makineleri ve ileri beslemeli sinir ağları gibi varyasyonel sinir ağlarının, güçlü etkileşim altında süperakışkan-Mott yalıtkanı faz geçişini gösteren faz diyagramını doğru bir şekilde tahmin etmedeki güvenilirliğini gösterdi. Ayrıca, bu sinir ağları, zayıf etkileşimli rejimde ortaya çıkan diğer ilginç çok parçacıklı fazları başarıyla buldu. Bu heyecan verici tespitler, manyetik akılı iki-bacaklı Bose-Hubbard merdivenini sinir-ağ kuantum durumları alanını ilerletmek için ideal bir test platformu olduğunu kesin bir şekilde işaret etmektedir. Bu tez, önceki sonuçları geliştirerek, sıradan Bose-Hubbard modeli ve manyetik akı altında iki-bacaklı Bose-Hubbard merdiveni hakkında kapsamlı bir giriş ve analiz, bosonik sistemler için özel olarak tasarlanmış sinir-ağ kuantum durumlarına derinlemesine bir genel bakışı ve bu her iki Bose-Hubbard modeli için sinir-ağ kuantum durumlarını kullanılarak elde edilen sonuçların kapsamlı bir sunumu ve analizini içeren çeşitli temel yönleri barındırmaktadır.

Anahtar sözcükler: Bose-Hubbard modeli, iki-bacaklı merdiven akı sistemi, süperakışkan fazı, Mott yalıtkanı fazı, sentetik manyetik alan, sinir-ağ kuantum durumları, makine öğrenmesi, yapay sinir ağları, kısıtlı Boltzmann makinesi, ileri beslemeli sinir ağı.

Acknowledgement

I am immensely grateful to Prof. M. Özgür Oktel for guiding my academic journey and providing unwavering support and guidance throughout nearly three years of research and exploration. This endeavor would not have been possible without his academic guideway and continuous help.

Furthermore, I extend my most profound appreciation to Prof. Ahmet Keleş, who introduced and taught me the cutting-edge applications of artificial neural networks in quantum many-body physics. His invaluable support and mentorship have played a pivotal role in my specialization in this fascinating field. I am genuinely grateful for his patience and expertise.

In addition, I would like to express my sincere thanks to Prof. Ceyhun Bulutay and Prof. Rifat Onur Umucalılar for their time and expertise in reading and reviewing my work. Their insightful feedback and constructive criticism have undoubtedly enhanced the quality and depth of my research.

Moreover, I would like to acknowledge the support provided by the Scientific and Technological Research Council of Turkey (TÜBİTAK) during my master's studies. I am honored to have been selected as a 2210-A National MSc/MA scholar. This recognition has provided financial assistance and served as a testament to the value of my research efforts. I am sincerely grateful for TÜBİTAK's support in facilitating my academic pursuits.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Quantum simulation has become a powerful tool, enabling the exploration of practical applications in quantum information and computation and the behavior of strongly correlated quantum matter [1, 2]. These advancements have led to significant progress in using quantum gas systems as analog quantum simulation platforms, providing fresh insights into challenging problems in condensed matter physics. In particular, these systems can potentially explain enigmatic phenomena in superconducting cuprates, fractional quantum Hall systems, and frustrated quantum magnets [3–6]. However, although pioneering cold-atom experiments have explored low-energy quantum correlations, they have also emphasized the requirement for dependable theoretical frameworks capable of accurately evaluating experimental outcomes [7].

To address this challenge, a successful approach involves investigating toy models in reduced-dimensional systems like two-leg ladders. These experimental setups offer several advantages. First, they allow for deploying accurate theoretical and numerical approaches developed for quasi-one-dimensional systems. Second, they are large enough to exhibit complex many-body phases, such as vortex and chiral phases in the superfluid state, similar to the phases observed in superconductors under magnetic fields and the Mott insulator phase arising from strong interactions [8, 9]. Two-leg ladder systems have been successfully realized

in various cold-atom setups [10–17], including those with synthetic dimensions [14, 18–20]. Their theoretical studies have significantly advanced alongside critical experimental progress [21–28]. By examining emerging numerical techniques in these well-controlled toy systems, researchers can uncover their limitations and contribute to the realization of quantum simulation in fully two-dimensional systems. Moreover, this exploration aids in formulating new theoretical frameworks that can effectively anticipate the fundamental principles governing the underlying physical phenomena.

Among the most effective unbiased numerical methods for studying strongly correlated matter, variational and projection Monte Carlo techniques are widely used, particularly in high-dimensional systems. These techniques involve stochastic minimization of variational wave functions derived from well-established physical principles, enabling the investigation of such models as the Hubbard, $t-J$, and Heisenberg models [29–32]. In 2017, Carleo and Troyer introduced a family of wave functions derived from artificial neural networks, employing a variational Monte-Carlo scheme inspired by the rapid advancements in machine learning and artificial intelligence [33, 34]. Their groundbreaking work demonstrated that these wave functions, referred to as the neural-network quantum states, have the potential to approximate ground states and dynamics of strongly correlated quantum systems. Remarkably, this approximation can be achieved with polynomial resources within the exponential Hilbert space [33, 35]. Moreover, their accuracy can be systematically enhanced, reaching a level of competitiveness comparable to advanced techniques like tensor networks and projected entangled pair states (PEPS). Follow-up studies revealed that neural-network quantum states exhibit volume law entanglement and have expressive capabilities similar to tensor-network quantum states [36–43]. The optimization of neural-network quantum states can be efficiently achieved using established machine learning and variational Monte-Carlo techniques. These optimization techniques tackle fundamental obstacles linked to tensor networks, including the difficulty of tensor contraction [44] and the exponential increase in the matrix product state (MPS) bond dimension with the size of the transverse system [45–47]. The utilization of neural-network quantum states has been effectively expanded to diverse

deep-learning architectures, demonstrating their utility in numerous problems in condensed matter physics [48–53].

In this thesis, by drawing inspiration from the simultaneous advancements in cold atomic systems and variational quantum Monte-Carlo simulations, the exploration of novel quantum phases in a two-leg Bose-Hubbard ladder under an artificial magnetic field is pursued using the neural-network quantum states. While significant theoretical interest and recent experimental progress exist, the application of neural networks in bosonic systems remains relatively limited. Several investigations into the superfluid-Mott insulator transition in the vanilla Bose-Hubbard models have been conducted, using restricted Boltzmann machines [54] and feedforward neural networks [55, 56]. As previously done in [57], a particular focus in this thesis has been placed on evaluating the efficiency of these approaches in the presence of artificial magnetic fields, which brake time-reversal symmetry and introduce frustration to the many-body system. Here, the primary objective is to contribute to investigating alternative numerical frameworks. Doing so aims to enhance future studies focusing on synthetic magnetic fields in optical lattice experiments. Moreover, an emphasis is placed on the two-leg flux ladder system, which exhibits a remarkable diversity of many-body phases. This system is presented as a candidate for future algorithmic developments within the field of neural network studies, further expanding the potential of this prototypical many-body system.

The organization of this thesis is as follows[1]:

Chapter 2 begins by introducing the vanilla Bose-Hubbard model, along with its phases and symmetries. The model is then analyzed using both mean-field theory and strong-coupling perturbation theory. Additionally, the chapter presents a two-leg Bose-Hubbard ladder under magnetic flux, exploring its novel phases in weakly interacting regimes through a mean-field approach [58]. The two-leg flux ladder system is further studied using the Gutzwiller variational approach and the strong-coupling expansion.

[1]Several chapters have been adapted from the paper [57].

Chapter 3 provides an introduction to neural-network quantum states for bosonic systems, focusing on two different ansatzes defined by distinct neural-network architectures: a restricted Boltzmann machine and a feedforward neural network. The chapter also delves into the measurement, sampling, and optimization processes associated with these ansatzes, offering a detailed discussion.

Finally, Chapter 4 presents the results obtained with neural-network quantum states for the two distinct Bose-Hubbard models. It concludes with a comprehensive analysis of these results, concluding the findings.

Chapter 2

Bose-Hubbard model

The Bose-Hubbard model, initially proposed by Gersch and Knollman in 1963 [59] and thoroughly analyzed to a much greater extent by Fisher *et al.* in 1989 [8], provides a framework to understand the physical behavior of spinless interacting bosonic particles within a lattice structure. In this chapter, two distinct Bose-Hubbard models are examined. The first model investigated is the vanilla Bose-Hubbard model, which exhibits two well-known quantum phases: the superfluid and Mott insulator. This model is analyzed using the mean-field theory and strong-coupling expansion, employing second-order perturbation theory. Additionally, the symmetries associated with this model are addressed. The second model explored is the two-leg ladder Bose-Hubbard model, which exhibits three novel superfluid phases under the influence of a synthetic magnetic flux in the weakly interacting regime. Conversely, in the strong interaction regime, the superfluid and Mott insulator phases still compete, and the resulting phase diagram captures distinct re-entrant phase transitions beyond a critical value of the magnetic flux. In addition to the analytical methods mentioned earlier, the Gutzwiller variational approach is employed to study this two-leg flux ladder system.

2.1 Vanilla Bose-Hubbard model

The standard or, so to say, *vanilla* Bose-Hubbard model in absolute zero ($T = 0$) can be investigated by the following Hamiltonian $\hat{\mathcal{H}}$:

$$\hat{\mathcal{H}} = -\sum_{i,j} J_{i,j}\left(\hat{a}_i^\dagger \hat{a}_j + \text{h.c.}\right) + \frac{U}{2}\sum_{i=1}^{N_\text{s}} \hat{n}_i(\hat{n}_i - 1) - \mu \sum_{i=1}^{N_\text{s}} \hat{n}_i, \tag{2.1}$$

where $J_{i,j}$ is the matrix element of the hopping amplitude between the lattice sites (i, j), which determines the boson mobility on the lattice, U is the on-site interaction strength responsible for how attractive or repulsive the bosons can be on the same lattice site, μ is the chemical potential, and N_s is the number of lattice sites. Here, $\hat{a}_i^\dagger$ and $\hat{a}_i$ denote the bosonic creation and annihilation operators at the lattice site i, and thus $\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i$ is the bosonic number operator for the same site.

For convenience, the hopping in this model is generally considered to be between the nearest neighbors $\langle i, j\rangle$, so that the Hamiltonian in Eq. (2.1) transforms into

$$\hat{\mathcal{H}} = -J\sum_{\langle i,j\rangle} \left(\hat{a}_i^\dagger \hat{a}_j + \text{h.c.}\right) + \frac{U}{2}\sum_{i=1}^{N_\text{s}} \hat{n}_i(\hat{n}_i - 1) - \mu \sum_{i=1}^{N_\text{s}} \hat{n}_i, \tag{2.2}$$

which is modelled in Fig. 2.1.

2.1.1 Mott insulator phase

In the atomic limit where the hopping terms are considered to be zero, the overall Hamiltonian in Eq. (2.2) can be written by defining a local Hamiltonian $\hat{h}_i = \frac{U}{2}\hat{n}_i(\hat{n}_i - 1) - \mu\hat{n}_i$, as follows:

$$\hat{\mathcal{H}}_\text{MI} = \sum_{i=1}^{N_\text{s}} \hat{h}_i, \tag{2.3}$$

so that

$$\hat{h}_i \left|\psi\right\rangle_i = \epsilon(n_i) \left|\psi\right\rangle_i, \tag{2.4}$$

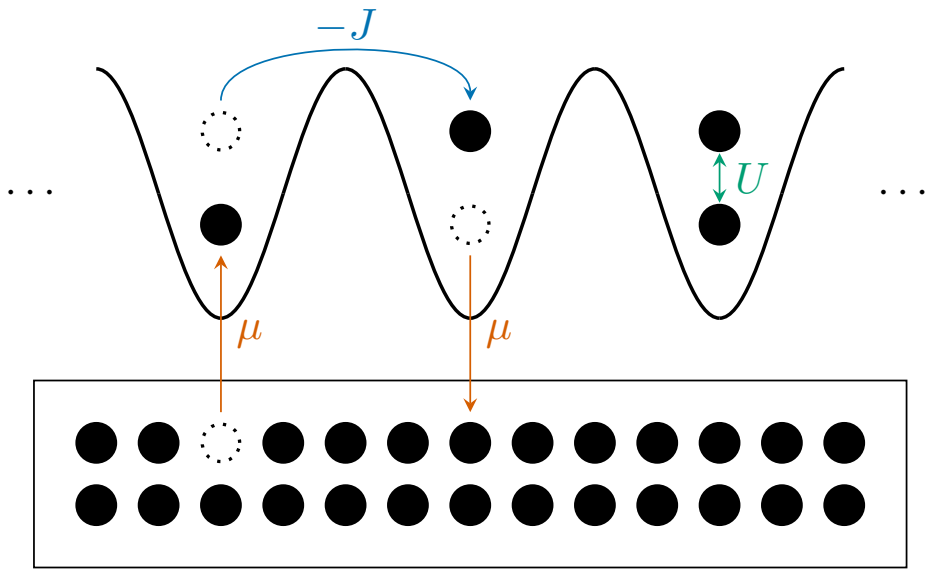

Figure 2.1: Schematic of the 1D Bose-Hubbard model. The rectangular box below represents the particle reservoir for the system.

where $\epsilon(n_i) \coloneqq \frac{U}{2} n_i(n_i - 1) - \mu n_i$ is the local ground-state energy for the lattice site i and the $|\psi\rangle_i$ is the ground state for the site i, which is independent of other lattice sites. In such a limit, the total energy can be minimized with a uniform distribution of bosons along the lattice because of the dominance of the on-site interactions between the bosons. Such a phase is called the *Mott insulator (MI) phase*. Thus, the overall ground state $|\Psi_{\mathrm{MI}}\rangle$ is

$$|\Psi_{\mathrm{MI}}\rangle = \bigotimes_{i=1}^{N_\mathrm{s}} \frac{1}{\sqrt{n_0}} \left(\hat{a}_i^\dagger\right)^{n_0} |0\rangle_i = \bigotimes_{i=1}^{N_\mathrm{s}} |n_0\rangle_i = |n_0\, n_0\, \ldots\, n_0\rangle\,, \tag{2.5}$$

where n_0 is the uniform number of bosons in each site, which suggests that the bosons are localized in real space but delocalized in the momentum space. This ground state also leads to the ground-state energy $E(N)$ equal to $N_\mathrm{s}\epsilon(n_0)$ where $N = n_0 N_\mathrm{s}$ is the total number of bosons.

Finding out how the choice of the on-site interaction strength U and the chemical potential μ determines the value of n_0 can be done by applying a particle or hole excitation to the system since there is a gap between the total energy $E(N \pm 1)$ with these excitations and the ground-state energy $E(N)$, i.e. $E(N \pm 1) \neq E(N)$. Supposing $\mu > 0$ and $U > 0$, for a hole excitation in the site

i causes the following inequality:

$$\begin{aligned} \epsilon(n_i) &\geq \epsilon(n_i - 1) && (2.6) \\ \frac{U}{2} n_0(n_0 - 1) - \mu n_0 &\geq \frac{U}{2}(n_0 - 1)(n_0 - 2) - \mu(n_0 - 1), && (2.7) \end{aligned}$$

which results as

$$\frac{\mu}{U} \geq n_0 - 1 \ . \tag{2.8}$$

Similarly, a particle excitation in the site i leads to

$$\epsilon(n_i + 1) \geq \epsilon(n_i), \tag{2.9}$$

which gives out

$$n_0 \geq \frac{\mu}{U} \ . \tag{2.10}$$

By combining Eqs. (2.8) and (2.10), it can be obtained that $n_0(\mu, U)$ is restricted by

$$n_0(\mu, U) = \begin{cases} n_0 & \text{if } n_0 - 1 \leq \frac{\mu}{U} \leq n_0 \text{ and } n_0 \in \mathbb{Z}^+ \\ 0 & \text{if } \frac{\mu}{U} \leq 0 \text{ and } n_0 = 0 \end{cases} , \tag{2.11}$$

which is also shown in Fig. 2.2.

This uniformity also causes fixing the particle density since

$$\langle \hat{\rho} \rangle \coloneqq \frac{1}{N_\mathrm{s}} \sum_{i=1}^{N_\mathrm{s}} \langle \hat{n}_i \rangle = n_0 = \frac{N}{N_\mathrm{s}} \in \mathbb{Z}^+, \tag{2.12}$$

which also indicates that the particle fluctuation vanishes, as shown below

$$\left\langle \hat{n}_i^2 \right\rangle - \langle \hat{n}_i \rangle^2 = 0 \ . \tag{2.13}$$

Using the result of particle density in Eq. (2.12), it can be said that the system in such a limit is incompressible since the density does not depend on μ, which leads to

$$\kappa \coloneqq \frac{\partial \langle \hat{\rho} \rangle}{\partial \mu} = 0, \tag{2.14}$$

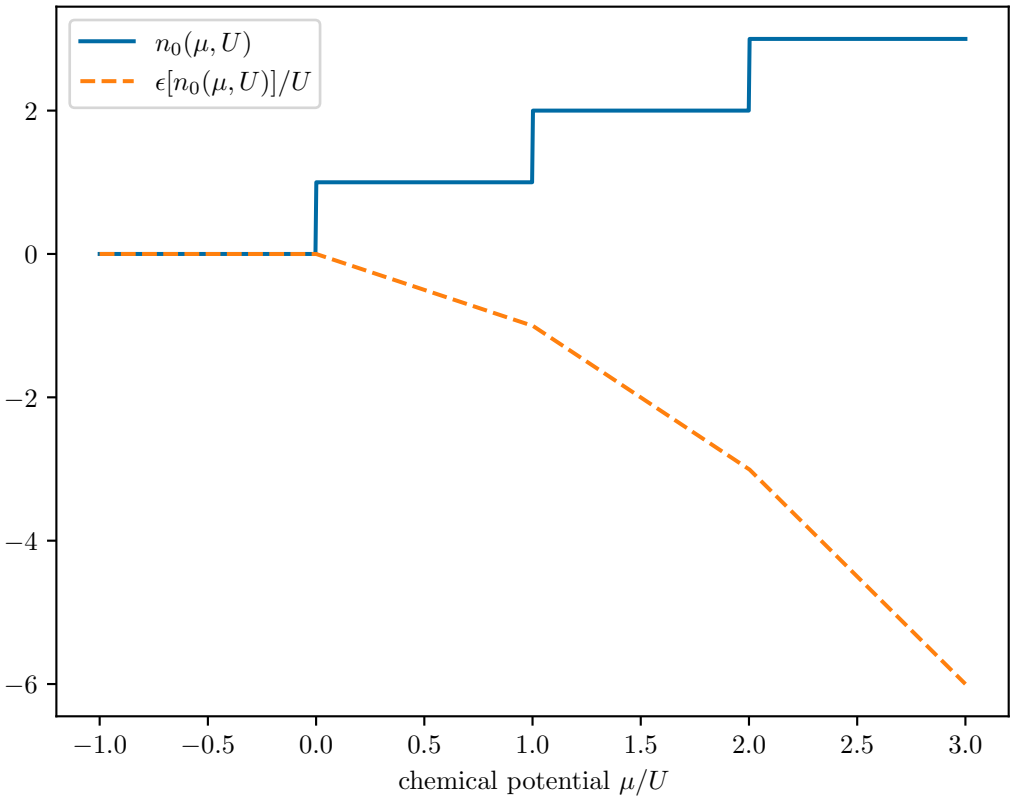

Figure 2.2: Uniform number $n_0(\mu, U)$ of bosons (blue line), and the local ground-state energy in the on-site interaction strength U (orange dashed line) as a function of the chemical potential μ and U for the perfect Mott insulator phase.

where κ is the compressibility of the system.

The expectation values of the bosonic field operators also vanish, as follows:

$$\langle \hat{a}_i \rangle = 0, \tag{2.15}$$

which is an important measure to distinguish the MI phase from the following phase.

2.1.2 Superfluid phase

In the other limit, where the on-site interactions completely disappear, and the bosons can move freely throughout the lattice, the Hamiltonian becomes

$$\hat{\mathcal{H}}_{\mathrm{SF}} = -J \sum_{\langle i,j \rangle} \left(\hat{a}_i^\dagger \hat{a}_j + \mathrm{h.c.} \right) - \mu \sum_{i=1}^{N_\mathrm{s}} \hat{n}_i, \tag{2.16}$$

which can be diagonalized in the momentum space by defining the equivalent bosonic field operators in this space as

$$\hat{a}_{\mathbf{k}} \;\coloneqq\; \frac{1}{\sqrt{N_\mathrm{s}}} \sum_i e^{i\mathbf{k}\cdot\mathbf{r}_i} \hat{a}_i \tag{2.17}$$

$$\hat{a}_{\mathbf{k}}^\dagger \;\coloneqq\; \frac{1}{\sqrt{N_\mathrm{s}}} \sum_i e^{-i\mathbf{k}\cdot\mathbf{r}_i} \hat{a}_i^\dagger, \tag{2.18}$$

under the assumption of the periodic boundary conditions and, thus

$$\hat{n}_k \coloneqq \frac{1}{N_\mathrm{s}} \sum_{i,j}^{N_\mathrm{s}} e^{i\mathbf{k}\cdot(\mathbf{r}_j - \mathbf{r}_j)} \hat{a}_i^\dagger \hat{a}_j, \tag{2.19}$$

which also suggests

$$\sum_{\mathbf{k}} \hat{n}_{\mathbf{k}} = \sum_i^{N_\mathrm{s}} \hat{n}_i \ . \tag{2.20}$$

These definitions also come with the following two orthonormality relations

$$\frac{1}{N_\mathrm{s}} \sum_{i=1}^{N_\mathrm{s}} e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}_i} \;=\; \delta(\mathbf{k}-\mathbf{k}') \tag{2.21}$$

$$\frac{1}{N_\mathrm{s}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{r}_j-\mathbf{r}_i)} \;=\; \delta(\mathbf{r}_j-\mathbf{r}_i) \ . \tag{2.22}$$

Using these definitions for a d-dimensional hypercubic lattice, the Hamiltonian

in Eq. (2.16) is obtained as

$$\begin{aligned}\hat{\mathcal{H}}_{\mathrm{SF}} &= -J\frac{1}{N_{\mathrm{s}}}\sum_{\langle i,j\rangle}\sum_{\mathbf{k},\mathbf{k}'}\left(e^{i(\mathbf{k}\cdot\mathbf{r}_i-\mathbf{k}'\cdot\mathbf{r}_j)}\hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}'}+\mathrm{h.c.}\right) \qquad (2.23)\\ &\quad -\mu\frac{1}{N_{\mathrm{s}}}\sum_{i=1}^{N_{\mathrm{s}}}\sum_{\mathbf{k},\mathbf{k}'}e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}_i}\hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}'}\\ &= -J\frac{1}{N_{\mathrm{s}}}\sum_{\langle i,j\rangle}\sum_{\mathbf{k},\mathbf{k}'}\left(e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}_i}e^{-i\mathbf{k}'\cdot\mathbf{d}_{ij}}\hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}'}+\mathrm{h.c.}\right) \qquad (2.24)\\ &\quad -\mu\sum_{\mathbf{k}}\hat{n}_{\mathbf{k}},\end{aligned}$$

where $\mathbf{d}_{ij} \coloneqq \mathbf{r}_j - \mathbf{r}_i$ is the nearest-neighbor vector, which has the same norm $\|\mathbf{d}_{ij}\| = a$ between all the nearest neighbors. These vectors can be simply rewritten as the primitive lattice vectors, *i.e.* $\mathbf{d}_{ij} \to \mathbf{d}_\ell$, and the transformation $\sum_{\langle i,j\rangle} \to \sum_{i=1}^{N_{\mathrm{s}}}\sum_{\ell=1}^{d}$ can be applied to the Hamiltonian in Eq. (2.24), which turns out to be

$$\begin{aligned}\hat{\mathcal{H}}_{\mathrm{SF}} &= -J\frac{1}{N_{\mathrm{s}}}\sum_{i=1}^{N_{\mathrm{s}}}\sum_{\ell=1}^{d}\sum_{\mathbf{k},\mathbf{k}'}\left(e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}_i}e^{-i\mathbf{k}'\cdot\mathbf{d}_\ell}\hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}'}+\mathrm{h.c.}\right) \qquad (2.25)\\ &\quad -\mu\sum_{\mathbf{k}}\hat{n}_{\mathbf{k}}\\ &= \sum_{\mathbf{k}}\left[-2J\sum_{\ell=1}^{d}\cos(\mathbf{k}\cdot\mathbf{d}_\ell)-\mu\right]\hat{n}_{\mathbf{k}}\ . \qquad (2.26)\end{aligned}$$

Defining the single-particle energy $\varepsilon_{\mathbf{k}}$ as

$$\begin{aligned}\varepsilon_{\mathbf{k}} &\coloneqq -2J\sum_{\ell=1}^{d}\cos(\mathbf{k}\cdot\mathbf{d}_\ell) \qquad (2.27)\\ &= -2J\sum_{\ell=1}^{d}\cos(k_\ell a), \qquad (2.28)\end{aligned}$$

the Hamiltonian in Eq. (2.26) becomes

$$\hat{\mathcal{H}}_{\mathrm{SF}} = \sum_{\mathbf{k}}(\varepsilon_{\mathbf{k}}-\mu)\hat{n}_{\mathbf{k}}, \qquad (2.29)$$

which can be minimized by setting the wave vector $\mathbf{k}$ to zero, *i.e.* the Bloch state of the lowest band. Such a phase is called the *superfluid phase (SF)*. Thus, its

ground state $|\Psi_{\mathrm{SF}}\rangle$ is defined as

$$|\Psi_{\mathrm{SF}}\rangle = \frac{1}{\sqrt{N!}}\left(\hat{a}^{\dagger}_{\mathbf{k}=0}\right)^{N}|0\rangle \tag{2.30}$$

$$= \frac{1}{\sqrt{N!}}\left(\frac{1}{\sqrt{N_{\mathrm{s}}}}\sum_{i=1}^{N_{\mathrm{s}}}\hat{a}^{\dagger}_{i}\right)^{N}|0\rangle\,, \tag{2.31}$$

which indicates the bosons' delocalization in real space and their localization in the momentum space. In this phase, the ground-state energy equals the total energy with particle-hole excitations, *i.e.* $E(N \pm 1) = E(N)$, so there is no energy gap between them.

Unlike the MI phase, the SF phase has no fixed particle density, *i.e.* $\langle\hat{\rho}\rangle \in \mathbb{R}^{+}$, and the particle fluctuation does not vanish anymore. Moreover, the system in such a phase is compressible, *i.e.* $\kappa \neq 0$, and $\langle\hat{a}_i\rangle$ or simply the SF order parameter is no longer zero.

2.1.3 Symmetries

A system's properties that remain invariant under transformations such as translations, rotations, or reflections are called the symmetries of the system. They are essential to analyze because they provide some significant constraints on the system's behavior.

For instance, the symmetries in the Bose-Hubbard model allow one to simplify the model by reducing its degrees of freedom. Besides, as Noether suggests in her first theorem [60], each continuous symmetry corresponds to a conserved quantity. Thus, the Bose-Hubbard model also has some conserved quantities due to the following symmetries.

2.1.3.1 Particle number conservation

The time derivative of an operator $\hat{A}$ can be written as

$$\frac{\partial \left\langle \hat{A} \right\rangle}{\partial t} = \frac{i}{\hbar} \left\langle \left[\hat{\mathcal{H}}, \hat{A} \right] \right\rangle \ . \tag{2.32}$$

Using Eq. (2.32) for the total bosonic number operator $\hat{N} \coloneqq \sum_{i=1}^{N_\mathrm{s}} \hat{n}_i$, its time derivative turns out to be zero because the Bose-Hubbard Hamiltonian commutes with it, which corresponds to a global-$U(1)$ symmetry, *i.e.* the total number of boson is conserved. This symmetry exists if only if no boson is created or annihilated due to some external sources.

2.1.3.2 Discrete translational symmetry

If the system is under a discrete translation transformation $i \to i + \ell$ where i is the lattice site index and ℓ is the shift amount, *i.e.* a spatial shift of the entire lattice by ℓ in any direction, the Hamiltonian and all observables turn out to be invariant. This symmetry is typically broken if the system is under a magnetic field in a staggered configuration [11, 61–64]. It is also noted that this symmetry may be broken in the numerical solutions due to finite system size.

2.1.3.3 $U(1)$ gauge symmetry

The Hamiltonian is invariant under the following local-$U(1)$ gauge transformation

$$\left.\begin{aligned} \hat{a}_i &\to e^{i\theta} \hat{a}_i \\ \hat{a}_i^\dagger &\to e^{-i\theta} \hat{a}_i^\dagger \end{aligned}\right\} \Rightarrow \hat{n}_i \to \hat{n}_i \Rightarrow \hat{\mathcal{H}} \to \hat{\mathcal{H}} \tag{2.33}$$

unless an external field is applied to the lattice.

Specifically, this gauge symmetry is broken for the SF phase due to the presence of the complex-valued SF order parameter $\langle \hat{a}_i \rangle \propto e^{i\phi}$ [3, 8, 65].

2.1.3.4 Time-reversal symmetry

Each bosonic field operator can be rewritten as

$$\hat{a}_i = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{q}_i + i\frac{1}{m\omega}\hat{p}_i\right) \tag{2.34}$$

$$\hat{a}_i^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{q}_i - i\frac{1}{m\omega}\hat{p}_i\right), \tag{2.35}$$

where $\hat{q}_i$ is the position operator and $\hat{p}_i$ is the momentum operator for the lattice site i.

Since a time reversal transformation causes $t \to -t$, then its related transformation operator, *i.e.* the anti-unitary time-reversal operator $\hat{\Theta}$ commutes with the position operator $\hat{q}_i$, and does not commute with the momentum operator $\hat{p}_i$, *i.e.* $\hat{\Theta}\hat{p}_i\hat{\Theta}^{-1} = -\hat{p}_i$ [66]. Using these commutation relations, it turns out that

$$\hat{\Theta}\hat{a}_i\hat{\Theta}^{-1} = \sqrt{\frac{m\omega}{2\hbar}}\hat{\Theta}\left(\hat{q}_i + i\frac{1}{m\omega}\hat{p}_i\right)\hat{\Theta}^{-1} \tag{2.36}$$

$$= \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{\Theta}\hat{q}_i\hat{\Theta}^{-1} - i\frac{1}{m\omega}\hat{\Theta}\hat{p}_i\hat{\Theta}^{-1}\right) \tag{2.37}$$

$$= \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{q}_i + i\frac{1}{m\omega}\hat{p}_i\right) \tag{2.38}$$

$$= \hat{a}_i\ . \tag{2.39}$$

A similar result for $\hat{a}_i^\dagger$ comes out to $\hat{\Theta}\hat{a}_i^\dagger\hat{\Theta}^{-1} = \hat{a}_i^\dagger$, so $\hat{\Theta}\hat{\mathcal{H}}\hat{\Theta}^{-1} = \hat{\mathcal{H}}$, *i.e.* the Hamiltonian is invariant under time reversal as long as the system is subjected to no external field.

2.1.4 Mean-field theory

Since these two phases compete in J-U-μ phase space, a phase transition must happen at some point. If the total number of bosons is kept fixed, the phase fluctuation determines a transition of the Berezinskii-Kosterlitz-Thouless (BKT) type [67, 68]. Once the particle-hole excitations make the total number of bosons vary, the transition is controlled by the particle fluctuation [8].

To determine where these transitions take place in the phase space, several analytical or numerical methods can be utilized. For such purposes, the simplest method is called the *mean-field (MF) theory*.

In the MF theory, the hopping terms are decoupled by assuming that the correlations between bosons are merely considered with an average value, *i.e.* a mean field, so that the bosonic field operators are written as

$$\hat{a}_i \coloneqq \psi + \delta\hat{a}_i \tag{2.40}$$

$$\hat{a}_i^\dagger \coloneqq \psi^* + \delta\hat{a}_i^\dagger, \tag{2.41}$$

where ψ is the complex expectation value $\langle\hat{a}_i\rangle$ of the annihilation operator — the SF order parameter — that is now independent of its lattice site, and $\delta\hat{a}_i^{(\dagger)} \coloneqq \hat{a}_i^{(\dagger)} - \psi^{(*)}$ is the small fluctuation term.

Using these definitions for the field operators, the hopping operator can be decoupled, as follows:

$$\hat{a}_i^\dagger\hat{a}_j = \psi^*\hat{a}_j + \psi\hat{a}_i^\dagger - |\psi|^2 + \delta\hat{a}_i^\dagger\delta\hat{a}_j \tag{2.42}$$

$$\approx \psi^*\hat{a}_j + \psi\hat{a}_i^\dagger - |\psi|^2, \tag{2.43}$$

which leads to

$$\hat{\mathcal{H}}_{\mathrm{MF}} = -J\sum_{\langle i,j\rangle}\left(\psi^*\hat{a}_j + \psi\hat{a}_i^\dagger - |\psi|^2\right) + \frac{U}{2}\sum_{i=1}^{N_\mathrm{s}}\hat{n}_i(\hat{n}_i-1) - \mu\sum_{i=1}^{N_\mathrm{s}}\hat{n}_i \tag{2.44}$$

$$= \sum_{i=1}^{N_\mathrm{s}}\left[-zJ\left(\psi^*\hat{a}_i + \psi\hat{a}_i^\dagger - |\psi|^2\right) + \frac{U}{2}\hat{n}_i(\hat{n}_i-1) - \mu\hat{n}_i\right] \tag{2.45}$$

$$= \sum_{i=1}^{N_\mathrm{s}}\left(\hat{h}_i^{(0)} + \hat{v}_i + zJ|\psi|^2\right), \tag{2.46}$$

where z is the coordination number of the lattice, $\hat{h}_i^{(0)} \coloneqq \frac{U}{2}\hat{n}_i(\hat{n}_i-1) - \mu\hat{n}_i$, and $\hat{v}_i \coloneqq -zJ\left(\psi^*\hat{a}_i + \psi\hat{a}_i^\dagger\right)$.

Since the perfect MI ground state $\left|n_0^{(0)}\right\rangle$ is the eigenstate of the unperturbed single-site Hamiltonian $\hat{h}_i^{(0)}$, then $\hat{h}_i^{(0)}$ exhibits

$$\hat{h}_i^{(0)}\left|n_0^{(0)}\right\rangle = \epsilon_{n_0}^{(0)}\left|n_0^{(0)}\right\rangle, \tag{2.47}$$

where $\epsilon_{n_0}^{(0)} \coloneqq \frac{U}{2} n_0(n_0-1) - \mu n_0$ is its eigenvalue.

For the rest of the calculations in the MF theory, only the single-site Hamiltonian is considered.

2.1.4.1 Perturbation theory

One way to find the MF ground states and the corresponding phase boundaries is to apply the perturbation theory by assuming that the unperturbed Hamiltonian is $\hat{h}_i^{(0)}$, and $\hat{v}_i$ is the perturbation.

The calculation of the first-order MF energy correction $\Delta\epsilon_{n_0}^{(1)}$ is straightforward:

$$\begin{aligned} \Delta\epsilon_{n_0}^{(1)} &= \left\langle n_0^{(0)} \middle| \hat{v}_i \middle| n_0^{(0)} \right\rangle && (2.48) \\ &= -zJ\left(\psi^* \sqrt{n_0}\delta_{n_0,n_0-1} + \psi\sqrt{n_0+1}\delta_{n_0,n_0+1}\right) && (2.49) \\ &= 0\ . && (2.50) \end{aligned}$$

Similarly, $\Delta\epsilon_{n_0}^{(p)}$ vanishes for every μ, U and ψ if p is a positive odd integer.

For the first-order MF state correction, *i.e.*

$$\left| n_0^{(1)} \right\rangle = \sum_{m \neq n} \frac{\left\langle m_0^{(0)} \middle| \hat{v}_i \middle| n_0^{(0)} \right\rangle}{\epsilon_{n_0}^{(0)} - \epsilon_{m_0}^{(0)}} \left| m_0^{(0)} \right\rangle , \tag{2.51}$$

the matrix element $\left\langle m_0^{(0)} \middle| \hat{v}_i \middle| n_0^{(0)} \right\rangle$ must be determined, which results as

$$\left\langle m_0^{(0)} \middle| \hat{v}_i \middle| n_0^{(0)} \right\rangle = -zJ\left(\psi^* \sqrt{n_0}\delta_{m_0,n_0-1} + \psi\sqrt{n_0+1}\delta_{m_0,n_0+1}\right) . \tag{2.52}$$

Eq. (2.52) suggests that there are only two possible m_0 values that do not equal to n_0, which is $m_0 = n_0 \pm 1$. Thus, the first order correction $\left| n_0^{(1)} \right\rangle$ in the MF state $|n_0\rangle$ becomes

$$\left| n_0^{(1)} \right\rangle = -zJ\left[\frac{\psi^*\sqrt{n_0}}{U(n_0-1)-\mu} \left| (n_0-1)^{(0)} \right\rangle - \frac{\psi\sqrt{n_0+1}}{Un_0-\mu} \left| (n_0+1)^{(0)} \right\rangle \right] . \tag{2.53}$$

In a similar way, the second- and fourth-order energy corrections can be calculated as

$$\Delta\epsilon_{n_0}^{(2)} = (zJ)^2|\psi|^2\left[\frac{n_0}{U(n_0-1)-\mu}-\frac{n_0+1}{Un_0-\mu}\right] \tag{2.54}$$

$$\Delta\epsilon_{n_0}^{(4)} = (zJ)^4|\psi|^4\left\{\frac{n_0(n_0-1)}{[U(n_0-1)-\mu]^2[U(2n_0-3)-2\mu]}\right. \tag{2.55}$$

$$-\frac{(n_0+1)(n_0+2)}{(Un_0-\mu)^2[U(2n_0+1)-2\mu]} \tag{2.56}$$

$$-\left[\frac{n_0}{U(n_0-1)-\mu}-\frac{n_0+1}{Un_0-\mu}\right] \tag{2.57}$$

$$\left.\times\left[\frac{n_0}{[U(n_0-1)-\mu]^2}+\frac{n_0+1}{(Un_0-\mu)^2}\right]\right\}. \tag{2.58}$$

Therefore, the MF state energy can be rewritten as

$$\epsilon_{n_0} = a_0 + a_2|\psi|^2 + a_4|\psi|^4 + \mathcal{O}\big(|\psi|^6\big), \tag{2.59}$$

as in Landau theory [69], where $a_0 \coloneqq \epsilon_{n_0}^{(0)}$, $a_p \coloneqq \Delta\epsilon_{n_0}^{(p)}/|\psi|^p$, and p is a positive even integer. For completeness, $a_2 \leftarrow a_2 + zJ$.

To minimize the MF energy in the fourth order of ψ, its partial derivative with respect to the complex conjugate ψ^* of the SF order parameter must be zero, as shown below

$$\frac{\partial\epsilon_{n_0}^{(4)}}{\partial\psi^*} = a_2\psi + 2a_4|\psi|^2\psi \tag{2.60}$$

$$= \psi\big(a_2 + 2a_4|\psi|^2\big) = 0, \tag{2.61}$$

which only gives two arguments of the energy minimum:

$$\psi = 0 \quad \& \quad |\psi| = \sqrt{\frac{-a_2}{2a_4}}\,. \tag{2.62}$$

To ensure that they are minima, the following derivative must be positive:

$$\frac{\partial^2\epsilon_{n_0}^{(4)}}{\partial\psi\partial\psi^*} = a_2 + 4a_4|\psi|^2 > 0\,. \tag{2.63}$$

Since Landau theory suggests that the highest power of the SF order parameter must be positive so that the minimization of the energy functional does not require an infinite order parameter, there are only two possible solutions for $|\psi|$:

$$|\psi| = \begin{cases} 0 & \text{if } a_2 > 0 \\ \sqrt{\frac{-a_2}{2a_4}} & \text{if } a_2 < 0 \end{cases}, \tag{2.64}$$

which obviously indicates that the MI phase exists if $a_2 > 0$, and the SF phase exists if $a_2 < 0$.

Since there must be a second-order phase transition occurring at $a_2 = 0$, then

$$a_2 = zJ + (zJ)^2 \left[\frac{n_0}{U(n_0 - 1) - \mu} - \frac{n_0 + 1}{Un_0 - \mu} \right] = 0 \ . \tag{2.65}$$

By defining $\tilde{J} \coloneqq J/U$ and $\tilde{\mu} \coloneqq \mu/U$, Eq. (2.65) thus becomes

$$z\tilde{J} + \left(z\tilde{J}\right)^2 \frac{\tilde{\mu} + 1}{[(n_0 - 1) - \tilde{\mu}](n_0 - \tilde{\mu})} = 0, \tag{2.66}$$

which has one valid solution for $\tilde{J}$:

$$\tilde{J} = -\frac{[(n_0 - 1) - \tilde{\mu}](n_0 - \tilde{\mu})}{z(\tilde{\mu} + 1)} \ . \tag{2.67}$$

From $\tilde{J}(\tilde{\mu}, z; n_0)$, $\tilde{\mu}\left(z\tilde{J}; n_0\right)$ can be obtained by solving the following quadratic equation:

$$\tilde{\mu}^2 + \tilde{\mu}\left(z\tilde{J} - 2n_0 + 1\right) + z\tilde{J} + n_0(n_0 - 1) = 0, \tag{2.68}$$

which gives two solutions:

$$\tilde{\mu}_\pm = \frac{1}{2}\left[2n_0 - 1 - z\tilde{J} \pm \sqrt{\left(z\tilde{J}\right)^2 - 2z\tilde{J}(2n_0 + 1) + 1} \right], \tag{2.69}$$

where $\mu_{+(-)}$ designates the upper (lower) boundaries of the Mott lobes. As shown in Fig. 2.3, the shape of the Mott lobes is concave.

As an example, for $n_0 = 1$, the critical $\tilde{\mu}$, *i.e.* $\tilde{\mu}_\mathrm{c}$ occurs when $\tilde{\mu}_+ = \tilde{\mu}_-$ at $\left(z\tilde{J}\right)_\mathrm{c}$, which results as

$$\left(z\tilde{J}\right)^2 - 6z\tilde{J} + 1 = 0 \Rightarrow \left(z\tilde{J}\right)_\mathrm{c} = 3 - 2\sqrt{2} \approx 0.172 \ . \tag{2.70}$$

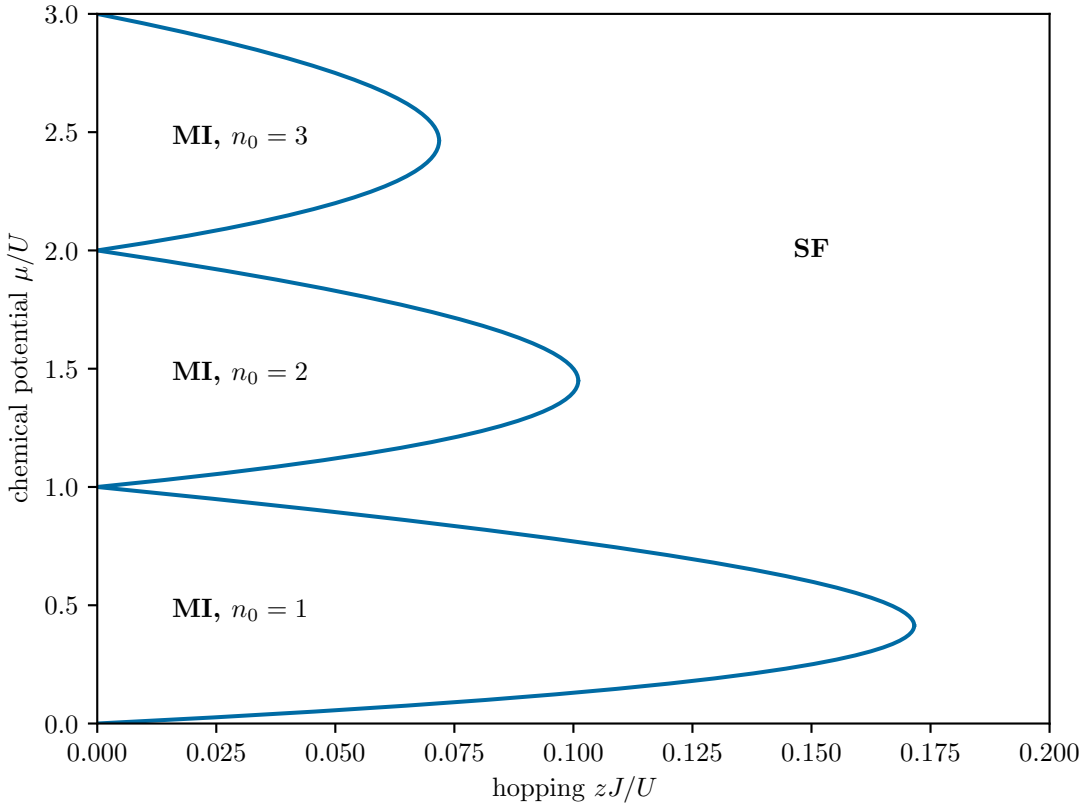

Figure 2.3: Mean-field phase diagram of the Bose-Hubbard model, obtained with the perturbative approach.

The expectation value of the particle density operator can also be determined, as follows:

$$\langle\hat{\rho}\rangle = -\frac{\partial\epsilon^{(4)}_{n_0}}{\partial\mu} = n_0 - \frac{\partial}{\partial\mu}\left(a_2|\psi|^2 + a_4|\psi|^4\right), \tag{2.71}$$

which results as $\langle\hat{\rho}\rangle = n_0$ for the MI phase because $|\psi| = 0$. However, since $|\psi| = \sqrt{\frac{-a_2}{2a_4}}$ for the SF phase, it now turns out to be

$$\begin{aligned}
\langle\hat{\rho}\rangle &= n_0 - \frac{\partial}{\partial\mu}\left(-\frac{a_2^2}{2a_4} + \frac{a_2^2}{4a_4}\right) && (2.72)\\
&= n_0 + \frac{\partial}{\partial\mu}\left(\frac{a_2^2}{4a_4}\right) && (2.73)\\
&= n_0 + \frac{a_2}{2a_4}\frac{\partial a_2}{\partial\mu} - \frac{a_2^2}{4a_4^2}\frac{\partial a_4}{\partial\mu}\ . && (2.74)
\end{aligned}$$

Here, it is noted that the SF-MI phase transition with a fixed $\langle\hat{\rho}\rangle = n_0$ occurs

when

$$\frac{\partial a_2}{\partial \mu} - \frac{a_2}{2a_4}\frac{\partial a_4}{\partial \mu} = 0 \ . \tag{2.75}$$

Thus, since this transition at the tips of the Mott lobes takes place when $a_2 = 0$, Eq. (2.75) then becomes $\frac{\partial a_2}{\partial \mu} = 0$.

Regarding compressibility, for the MI phase, $\kappa = 0$ since $\langle \hat{\rho} \rangle = n_0$. Whereas, for the SF phase, κ no longer vanishes but turns out to be

$$\kappa = \frac{\partial^2}{\partial \mu^2}\left(\frac{a_2^2}{4a_4}\right) \ . \tag{2.76}$$

2.1.4.2 Self-consistency solution

Instead of using the perturbation theory, the ground state can be found by optimizing the SF order parameter ψ via a self-consistency procedure.

To do that, the matrix elements of the single-particle Hamiltonian denoted in Eq. (2.46) must be determined, which are

$$\begin{aligned} \langle n|\hat{\mathcal{H}}_{\mathrm{MF},i}|m\rangle \ &= \ \left[\frac{U}{2}m(m-1) - \mu m + zJ|\psi|^2\right]\delta_{n,m} \\ &\quad -zJ\psi^*\sqrt{m}\,\delta_{n,m-1} - zJ\psi\sqrt{n}\,\delta_{n-1,m}, \end{aligned} \tag{2.77}$$

which can be shown as

$$\hat{\mathcal{H}}_{\mathrm{MF},i} \coloneqq \begin{pmatrix} zJ|\psi|^2 & -zJ\psi^* & 0 & 0 \\ -zJ\psi & -\mu + zJ|\psi|^2 & -\sqrt{2}zJ\psi^* & 0 \\ 0 & -\sqrt{2}zJ\psi & U - 2\mu + zJ|\psi|^2 & -\sqrt{3}zJ\psi^* \\ 0 & 0 & -\sqrt{3}zJ\psi & 3U - 3\mu + zJ|\psi|^2 \end{pmatrix} \tag{2.78}$$

if the maximum occupation number is set to be 3, *i.e.* $n_{\max} \to 3$. Similarly, the bosonic field operators can also be represented by the following matrices:

$$\hat{a}_i \coloneqq \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & \sqrt{3} \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \& \quad \hat{a}_i^\dagger \coloneqq \begin{pmatrix} 0 & 0 & 0 & 0 \\ \sqrt{1} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix} . \tag{2.79}$$

In the beginning, ψ is initialized randomly. Then, the expectation value $\langle\Psi_0|\hat{a}_i|\Psi_0\rangle$ of the annihilation operator is determined with the first normalized eigenvector $|\Psi_0\rangle$ of the single-particle Hamiltonian matrix in Eq. (2.78) for the current ψ. Ideally, this expectation value should equal ψ, so the true value of ψ can be found by repeating this calculation process, as shown in Algorithm 1. As an example, $|\psi|^2$ and $\langle\hat{n}_i\rangle$ are plotted for the different values of zJ/U and μ/U in Fig. 2.4.

Algorithm 1 Self-consistency procedure for the mean-field theory

Require: $\varepsilon \ll 0$, $n_{\max} \in \mathbb{Z}^+$, $\psi \in \mathbb{R}$

$\psi_{\text{opt}}^{\text{pre}} \leftarrow 0$
$\psi_{\text{opt}} \leftarrow 1$
while $\left|\psi_{\text{opt}} - \psi_{\text{opt}}^{\text{pre}}\right| > \varepsilon$ **do**
    determine the normalized first eigenvector $|\Psi_0\rangle$ of the mean-field Hamiltonian matrix $\hat{\mathcal{H}}_{\text{MF},i}$ for ψ_{opt}
    $\psi_{\text{opt}}^{\text{pre}} \leftarrow \psi_{\text{opt}}$
    $\psi_{\text{opt}} \leftarrow \langle\Psi_0|\hat{a}_i|\Psi_0\rangle$
end while

2.1.5 Strong-coupling perturbation theory

Another approach to obtain the SF-MI phase diagram is a strong-coupling expansion with the perturbation theory [70, 71]. For that, the Hamiltonian is rewritten as

$$\hat{\mathcal{H}} \coloneqq \hat{\mathcal{H}}^{(0)} + \lambda\hat{\mathcal{V}}, \tag{2.80}$$

where

$$\hat{\mathcal{H}}^{(0)} \coloneqq \frac{U}{2}\sum_{i=1}^{N_{\text{s}}}\hat{n}_i(\hat{n}_i - 1) - \mu\sum_{i=1}^{N_{\text{s}}}\hat{n}_i \tag{2.81}$$

$$\hat{\mathcal{V}} \coloneqq -U\sum_{\langle i,j\rangle}\left(\hat{a}_i^\dagger\hat{a}_j + \text{h.c.}\right), \tag{2.82}$$

and $\lambda \coloneqq J/U$ is the small parameter. With such a choice, the MI ground state and its corresponding energy can be written as

$$|\Psi_{\text{MI}}\rangle_\lambda \coloneqq \left|n_0^{(0)}\right\rangle + \lambda\left|n_0^{(1)}\right\rangle + \lambda^2\left|n_0^{(2)}\right\rangle + \cdots \tag{2.83}$$

$$E_{n_0}(\lambda) \coloneqq E_{n_0}^{(0)} + \lambda\,\Delta E_{n_0}^{(1)} + \lambda^2\,\Delta E_{n_0}^{(2)} + \cdots, \tag{2.84}$$

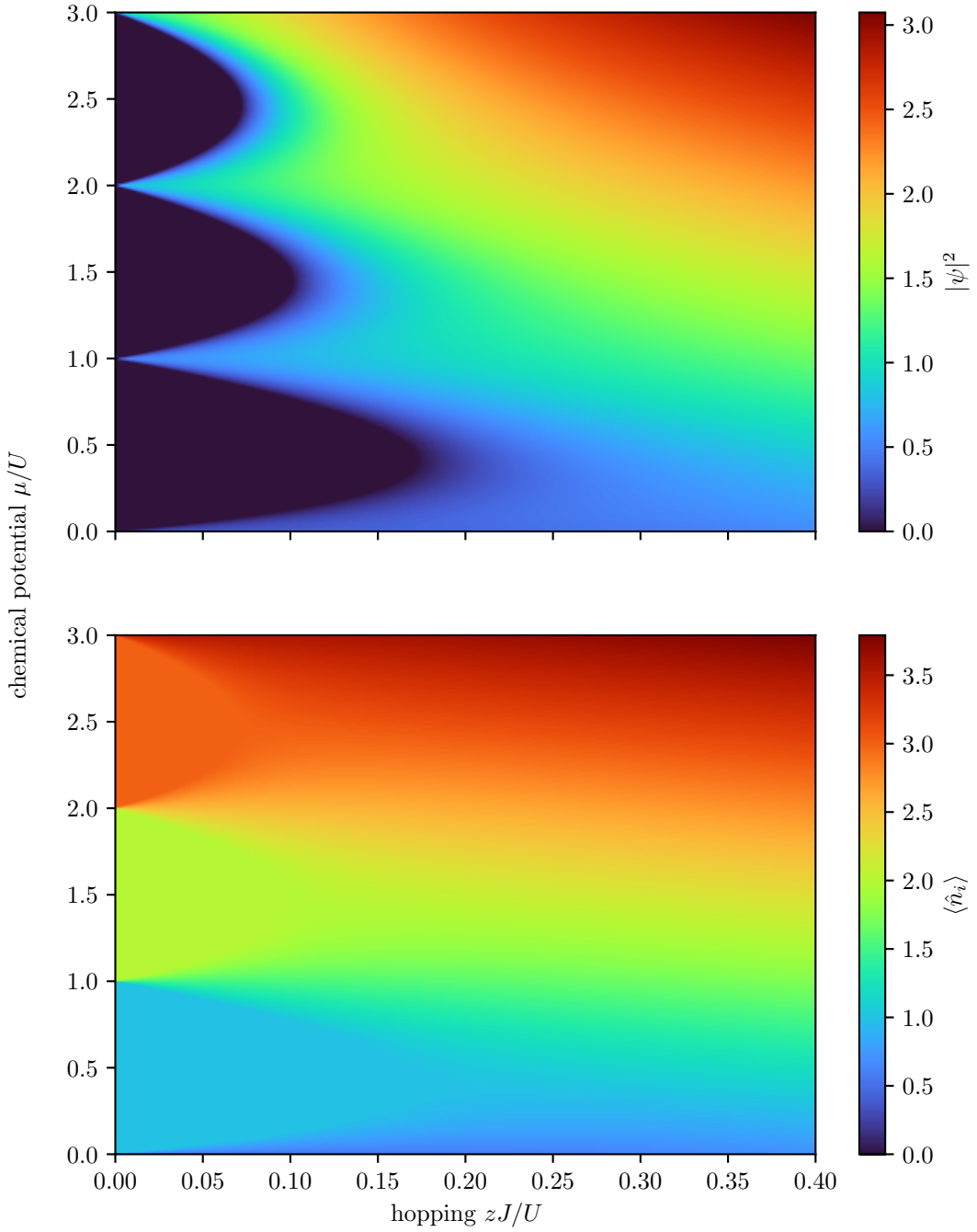

Figure 2.4: Mean-field phase diagrams of the Bose-Hubbard model in the zJ-μ plane with the self-consistency for $n_{\max} = 8$. The z-axis of the subplots denotes the modulus squared $|\psi|^2$ of the superfluid order parameter ψ and the expectation value $\langle \hat{n}_i \rangle$ of the local number operator $\hat{n}_i$, respectively.

where $\left|n_0^{(0)}\right\rangle$ indicates the perfect MI state $|\Psi_{\mathrm{MI}}\rangle$, so that

$$\hat{\mathcal{H}}^{(0)}\left|n_0^{(0)}\right\rangle = E_n^{(0)}\left|n_0^{(0)}\right\rangle, \tag{2.85}$$

where $E_{n_0}^{(0)} = N_{\mathrm{s}}\left[\frac{U}{2}n_0(n_0-1) - \mu n_0\right]$.

It is noted that the first-order energy correction vanishes, as follows:

$$\Delta E_{n_0}^{(1)} = \left\langle n_0^{(0)}\middle|\hat{\mathcal{V}}\middle|n_0^{(0)}\right\rangle = 0, \tag{2.86}$$

which also indicates that the odd-order energy corrections vanish, *i.e.* $\Delta E_{n_0}^{(p)} = 0$ where p is a positive odd integer.

Since the first-order state correction is defined as

$$\left|n_0^{(1)}\right\rangle = \sum_{m\neq n_0}\frac{\left\langle m^{(0)}\middle|\hat{\mathcal{V}}\middle|n_0^{(0)}\right\rangle}{E_{n_0}^{(0)} - E_m^{(0)}}\left|m^{(0)}\right\rangle, \tag{2.87}$$

the only valid ket state for that correction is

$$\begin{aligned}\left|m^{(0)}\right\rangle &:= \mathcal{N}\sum_{\langle i,j\rangle}\left(\hat{a}_i^\dagger\hat{a}_j + \mathrm{h.c.}\right)\left|n_0^{(0)}\right\rangle &(2.88)\\ &= -\frac{\mathcal{N}}{U}\hat{\mathcal{V}}\left|n_0^{(0)}\right\rangle, &(2.89)\end{aligned}$$

where $\mathcal{N}$ is a real-valued normalization constant, so this state correction becomes

$$\left|n_0^{(1)}\right\rangle = \frac{\mathcal{N}^2}{U^2}\frac{\left\langle n_0^{(0)}\middle|\hat{\mathcal{V}}^2\middle|n_0^{(0)}\right\rangle}{E_{n_0}^{(0)} - E_m^{(0)}}\hat{\mathcal{V}}\left|n_0^{(0)}\right\rangle \ . \tag{2.90}$$

Finding the result of the matrix element in the numerator in Eq. (2.90) is straightforward, as shown below:

$$\begin{aligned}\left\langle n_0^{(0)}\middle|\hat{\mathcal{V}}^2\middle|n_0^{(0)}\right\rangle &= U^2\sum_{\langle i,j\rangle}\left\langle n_0^{(0)}\middle|\left(\hat{a}_i^\dagger\hat{a}_j + \mathrm{h.c.}\right)^2\middle|n_0^{(0)}\right\rangle &(2.91)\\ &= U^2\sum_{\langle i,j\rangle}\left\langle n_0^{(0)}\middle|\hat{n}_i + \hat{n}_j + 2\hat{n}_i\hat{n}_j\middle|n_0^{(0)}\right\rangle &(2.92)\\ &= 2U^2N_{\mathrm{nn}}n_0(n_0+1), &(2.93)\end{aligned}$$

where N_{nn} is the number of the nearest-neighbour couples in the whole lattice, which is different from the coordination number z denoting the number of the nearest neighbours for a single lattice site. This matrix element also leads to finding the normalization constant $\mathcal{N}$ for $\left|m^{(0)}\right\rangle$ because

$$\begin{aligned} \left\langle m^{(0)}\middle|m^{(0)}\right\rangle = 1 &= \frac{\mathcal{N}^2}{U^2}\left\langle n_0^{(0)}\middle|\hat{\mathcal{V}}^2\middle|n_0^{(0)}\right\rangle && (2.94)\\ &= 2\mathcal{N}^2 N_{\mathrm{nn}} n_0(n_0+1)\ . && (2.95)\end{aligned}$$

Since $N_{\mathrm{nn}} \coloneqq zN_{\mathrm{s}}/2$ for a hypercubic lattice with the periodic boundary conditions, the constant turns out to be

$$\mathcal{N} = \frac{1}{\sqrt{zN_{\mathrm{s}}n_0(n_0+1)}}, \tag{2.96}$$

which also allows one to rewrite that matrix element as

$$\left\langle n_0^{(0)}\middle|\hat{\mathcal{V}}^2\middle|n_0^{(0)}\right\rangle = \frac{U^2}{\mathcal{N}^2}\ . \tag{2.97}$$

The energy for the ket state $\left|m^{(0)}\right\rangle$ is also required to determine the first-order state correction, which results as

$$\begin{aligned} E_m^{(0)} &= \left\langle m^{(0)}\middle|\hat{\mathcal{H}}^{(0)}\middle|m^{(0)}\right\rangle && (2.98)\\ &= \frac{\mathcal{N}^2}{2U}\left\langle n_0^{(0)}\middle|\hat{\mathcal{V}}\hat{\mathcal{H}}^{(0)}\hat{\mathcal{V}}\middle|n_0^{(0)}\right\rangle && (2.99)\\ &= \frac{\mathcal{N}^2}{U^2}\Big(E_{n_0}^{(0)}\left\langle n_0^{(0)}\middle|\hat{\mathcal{V}}^2\middle|n_0^{(0)}\right\rangle + \left\langle n_0^{(0)}\middle|\hat{\mathcal{V}}\left[\hat{\mathcal{H}}^{(0)},\hat{\mathcal{V}}\right]\middle|n_0^{(0)}\right\rangle\Big) && (2.100)\\ &= \frac{\mathcal{N}^2}{U^2}\big(E_{n_0}^{(0)}+U\big)\left\langle n_0^{(0)}\middle|\hat{\mathcal{V}}^2\middle|n_0^{(0)}\right\rangle && (2.101)\\ &= E_{n_0}^{(0)}+U, && (2.102)\end{aligned}$$

which leads to

$$\left|n_0^{(1)}\right\rangle = -\frac{\hat{\mathcal{V}}}{U}\left|n_0^{(0)}\right\rangle = \sum_{\langle i,j\rangle}\left(\hat{a}_i^\dagger\hat{a}_j + \hat{a}_j^\dagger\hat{a}_i\right)\left|n_0^{(0)}\right\rangle\ . \tag{2.103}$$

Using this first-order state correction, the second-order energy correction becomes

$$\Delta E^{(2)}_{n_0} = \left\langle n_0^{(0)} \middle| \hat{\mathcal{V}} \middle| n_0^{(1)} \right\rangle \tag{2.104}$$

$$= -\frac{1}{U} \left\langle n_0^{(0)} \middle| \hat{\mathcal{V}}^2 \middle| n_0^{(0)} \right\rangle \tag{2.105}$$

$$= -zUN_{\mathrm{s}} n_0(n_0+1), \tag{2.106}$$

so the perturbative energy of the state up to the econd order is obtained as

$$E^{(2)}_{n_0} = N_s \left[\frac{U}{2} n_0(n_0-1) - \mu n_0 - \frac{zJ^2}{U} n_0(n_0+1) \right]. \tag{2.107}$$

Determining the SF-MI phase boundaries also requires the energy of the defect states

$$\left| \Psi_{+,i}{}^{(0)} \right\rangle := \frac{1}{\sqrt{n_0+1}} \hat{a}_i^\dagger \left| n_0{}^{(0)} \right\rangle \tag{2.108}$$

$$\left| \Psi_{-,i}{}^{(0)} \right\rangle := \frac{1}{\sqrt{n_0}} \hat{a}_i \left| n_0{}^{(0)} \right\rangle, \tag{2.109}$$

which are simply the MI ground state with a particle and hole excitation in one lattice site. As can be seen in Eqs. (2.108) and (2.109), they are N_{s}-fold degenerate, so the degenerate perturbation theory is applied for these defect states. To utilize their correct form, their general states are defined as

$$\left| \Psi_{+}{}^{(0)} \right\rangle := \sum_{i=1}^{N_{\mathrm{s}}} f_i \left| \Psi_{+,i}{}^{(0)} \right\rangle \tag{2.110}$$

$$\left| \Psi_{-}{}^{(0)} \right\rangle := \sum_{i=1}^{N_{\mathrm{s}}} f_i \left| \Psi_{-,i}{}^{(0)} \right\rangle, \tag{2.111}$$

where f_i is the eigenvector of the single-particle matrix $S_{i,j} = -J_{i,j} = -J\delta[\mathbf{r}_i - (\mathbf{r}_j \pm \mathbf{e}_\mu)]$ where $\mu = x, y$ or z.

The calculation of the perturbative energy of the defect states is long but

explicit [70, 72], which gives the following perturbative energies:

$$E_+^{(2)} = E_{n_0}^{(2)} + Un_0 - \mu + \lambda_{\min}(n_0+1) + \frac{1}{2U}\sum_{i,j} J_{i,j}^2 f_j^2\, n_0(5n_0+4) - \frac{1}{U}\lambda_{\min}^2 n_0(n_0+1) \quad (2.112)$$

$$E_-^{(2)} = E_{n_0}^{(2)} - U(n_0-1) + \mu + \lambda_{\min} n_0 + \frac{1}{2U}\sum_{i,j} J_{i,j}^2 f_j^2\, (n_0+1)(5n_0+1) - \frac{1}{U}\lambda_{\min}^2 n_0(n_0+1)\ . \quad (2.113)$$

Since all possible hoppings are between the nearest neighbours in a hypercubic lattice, the lowest eigenvalue $\lambda_{\min}$ of the single-particle matrix $S_{i,j}$ is equal to $-zJ$ and the sum $\sum_{i,j} J_{i,j}^2 f_j^2$ thus becomes zJ^2, which makes Eqs. (2.112) and (2.113) to turn out to be

$$E_+^{(2)} = E_{n_0}^{(2)} + Un_0 - \mu - zJ(n_0+1) + \frac{zJ^2}{2U} n_0(5n_0+4) - \frac{z^2J^2}{U} n_0(n_0+1) \quad (2.114)$$

$$E_-^{(2)} = E_{n_0}^{(2)} - U(n_0-1) + \mu - zJn_0 + \frac{zJ^2}{2U}(n_0+1)(5n_0+1) - \frac{z^2J^2}{U} n_0(n_0+1)\ . \quad (2.115)$$

To find the boundaries of the Mott lobes, $E_+^{(2)} - E_{n_0}^{(2)} = 0$ and $E_{n_0}^{(2)} - E_-^{(2)} = 0$ are solved for $\mu_+^{(2)}$ and $\mu_-^{(2)}$, respectively:

$$\tilde{\mu}_+^{(2)} = n_0 - z\tilde{J}(n_0+1) + \tfrac{z\tilde{J}^2}{2} n_0(5n_0+4) - z^2\tilde{J}^2 n_0(n_0+1) \quad (2.116)$$

$$\tilde{\mu}_-^{(2)} = n_0 - 1 + z\tilde{J}n_0 - \tfrac{z\tilde{J}^2}{2}(n_0+1)(5n_0+1) + z^2\tilde{J}^2 n_0(n_0+1), \quad (2.117)$$

which are plotted for the first three Mott lobes in Fig. 2.5. Here, it is noted that the Mott lobes' shape is no longer concave as in the mean-field result but convex.

For example, for $n_0 = 1$ and an 1D hypercubic lattice, *i.e.* $z = 2$, $\tilde{\mu}_c$ occurs when $\tilde{\mu}_+ = \tilde{\mu}_-$ at $\tilde{J}_c$, which results as

$$5\tilde{J} - 6\tilde{J} + 1 = 0 \Rightarrow \tilde{J}_c = 0.2\ . \quad (2.118)$$

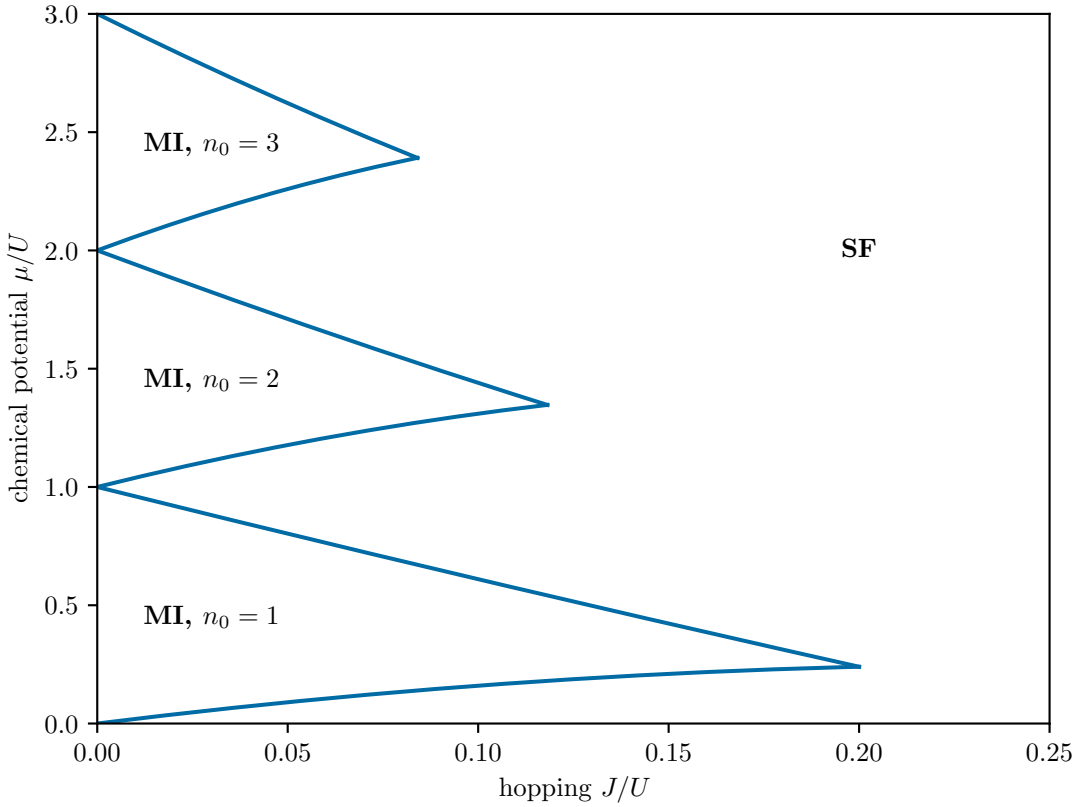

Figure 2.5: Phase diagram of a 1D Bose-Hubbard model with the strong-coupling perturbation theory.

2.2 Two-leg Bose-Hubbard ladder under an artificial magnetic flux

In this section, the Bose-Hubbard model is considered to be confined on an optical lattice in the form of ladder geometry with rungs under artificial magnetic flux, as shown in Fig. 2.6.

Its Hamiltonian can be defined as

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_{\mathrm{t}} + \hat{\mathcal{H}}_{\mathrm{int}}, \tag{2.119}$$

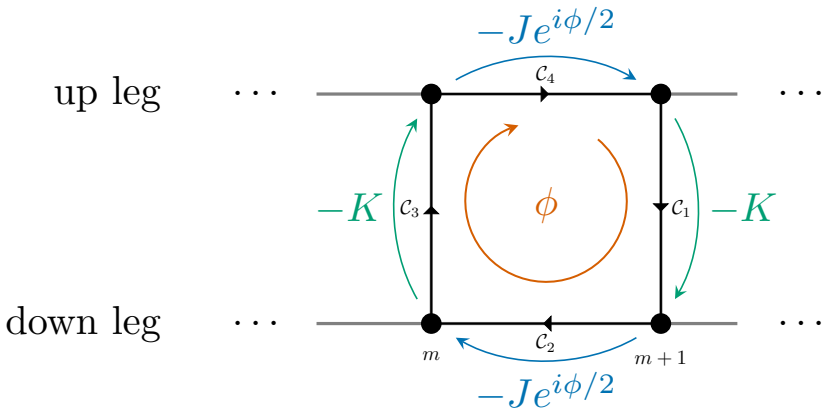

Figure 2.6: Schematic of the two-leg Bose-Hubbard ladder under magnetic field.

where

$$\hat{\mathcal{H}}_{\mathrm{t}} \coloneqq -J\sum_{m=1}^{L}\sum_{\ell\in\{\mathrm{u,d}\}}\left(e^{i\sigma_\ell\phi/2}\hat{a}^\dagger_{\ell,m+1}\hat{a}_{\ell,m}+\mathrm{h.c.}\right) \\ -K\sum_{m=1}^{L}\left(\hat{a}^\dagger_{\mathrm{u},m}\hat{a}_{\mathrm{d},m}+\mathrm{h.c.}\right), \tag{2.120}$$

and

$$\hat{\mathcal{H}}_{\mathrm{int}} \coloneqq \frac{U}{2}\sum_{m=1}^{L}\sum_{\ell\in\{\mathrm{u,d}\}}\hat{n}_{\ell,m}(\hat{n}_{\ell,m}-1)\ . \tag{2.121}$$

Here, $\hat{a}_{\ell,m}$ and $\hat{a}^\dagger_{\ell,m}$ are the bosonic field operators at the lattice site (ℓ,m), $\ell=\mathrm{u,d}$ indicates the upper and lower site of a rung, J and K is the intra- and inter-leg hopping amplitudes, U is the on-site interaction strength, $\sigma_\ell=+1$ (-1) for $\ell=\mathrm{u}$ (d), ϕ is the phase from the artificial magnetic field, and L is the number of rungs. Using the Peierls substitution [73], the phase can be related to a synthetic magnetic flux $\phi=\int_{\mathcal{A}}\mathbf{B}\cdot\mathrm{d}\mathbf{a}=\oint_{\mathcal{C}}\mathbf{A}\cdot\mathrm{d}\mathbf{l}$ passing through each plaquette, where $\mathbf{B}$ and $\mathbf{A}$ are the corresponding artificial magnetic field and vector potential, $\mathcal{C}$ is a closed path around a plaquette, $\mathcal{A}$ is the area of a plaquette, and $\mathbf{B}=\boldsymbol{\nabla}\times\mathbf{A}$. Since the Hamiltonian in Eq. (2.119) is invariant under the tranformation $(\mathrm{u,d},\phi)\to(\mathrm{d,u},-\phi)$, the domain of the flux is limited as $0<\phi\leq 2\pi$.

For the weakly interacting regime, the SF phases differentiate from each other [15, 58], which can be probed with the single-particle solution, as follows: Using

the translation invariance along legs, the bosonic field operators in the momentum space can be defined as

$$\hat{a}_k \coloneqq \frac{1}{\sqrt{L}} \sum_k e^{ikma} \hat{a}_{\mathrm{u},m} \tag{2.122}$$

$$\hat{b}_k \coloneqq \frac{1}{\sqrt{L}} \sum_k e^{ikma} \hat{a}_{\mathrm{d},m}, \tag{2.123}$$

so that the inter-leg coupling terms transform into

$$\sum_{m=1}^{L} \hat{a}^\dagger_{\mathrm{u},m} \hat{a}_{\mathrm{d},m} = \frac{1}{L} \sum_{k,k'} \sum_{m=1}^{L} e^{i(k-k')ma} \hat{a}^\dagger_k \hat{b}_{k'} \tag{2.124}$$

$$= \sum_k \hat{a}^\dagger_k \hat{b}_k, \tag{2.125}$$

and similarly for the intra-leg coupling terms:

$$\sum_{m=1}^{L} e^{i\phi/2} \hat{a}^\dagger_{\mathrm{u},m+1} \hat{a}_{\mathrm{u},m} = \frac{1}{L} \sum_{k,k'} \sum_{m=1}^{L} e^{i(k-k')ma} e^{i(ka+\phi/2)} \hat{a}^\dagger_k \hat{a}_{k'} \tag{2.126}$$

$$= \sum_k e^{i(ka+\phi/2)} \hat{a}^\dagger_k \hat{a}_k \tag{2.127}$$

$$\sum_{m=1}^{L} e^{-i\phi/2} \hat{a}^\dagger_{\mathrm{d},m+1} \hat{a}_{\mathrm{d},m} = \sum_k e^{i(ka-\phi/2)} \hat{b}^\dagger_k \hat{b}_k, \tag{2.128}$$

These definitions let one redefine the Hamiltonian in Eq. (2.119) for $U = 0$ as

$$\hat{\mathcal{H}}_{\mathrm{t}} = \sum_k \hat{\mathcal{H}}_k, \tag{2.129}$$

where

$$\hat{\mathcal{H}}_k \coloneqq -2J \left[\cos\left(ka + \frac{\phi}{2} \right) \hat{a}^\dagger_k \hat{a}_k + \cos\left(ka - \frac{\phi}{2} \right) \hat{b}^\dagger_k \hat{b}_k \right] - K \left(\hat{a}^\dagger_k \hat{b}_k + \hat{b}^\dagger_k \hat{a}_k \right), \tag{2.130}$$

or simply

$$\hat{\mathcal{H}}_k = \hat{\mathbf{c}}^\dagger_k \mathbf{h}_k \hat{\mathbf{c}}_k \tag{2.131}$$

with

$$\mathbf{h}_k \coloneqq -2J \left[\cos(ka) \cos\left(\frac{\phi}{2} \right) \boldsymbol{\sigma}_0 - \sin(ka) \sin\left(\frac{\phi}{2} \right) \boldsymbol{\sigma}_z \right] - K \boldsymbol{\sigma}_x, \tag{2.132}$$

where $\boldsymbol{\sigma}_i$ are the Pauli matrices.

It is noted that the single-particle Hamiltonian $\hat{\mathcal{H}}_k$ in Eq. (2.131) can be diagonalized via the Bogoliubov transformation, as follows:

$$\hat{\boldsymbol{\gamma}}_k \coloneqq \begin{pmatrix} \hat{\alpha}_k \\ \hat{\beta}_k \end{pmatrix} \coloneqq [\cos(\theta_k)\boldsymbol{\sigma}_0 + i\sin(\theta_k)\boldsymbol{\sigma}_y]\hat{\mathbf{c}}_k, \tag{2.133}$$

which transforms the local Hamiltonian into

$$\hat{\mathcal{H}}_k = \hat{\boldsymbol{\gamma}}_k^\dagger \begin{pmatrix} \epsilon_+(k) & 0 \\ 0 & \epsilon_-(k) \end{pmatrix} \hat{\boldsymbol{\gamma}}_k, \tag{2.134}$$

where

$$\theta_k \coloneqq \frac{1}{2}\arctan\left(\frac{-K}{2J\sin(ka)\sin(\phi/2)}\right), \tag{2.135}$$

and

$$\epsilon_\pm(k) = -2J\cos(ka)\cos\left(\frac{\phi}{2}\right) \pm \sqrt{4J^2\sin^2(ka)\sin^2\left(\frac{\phi}{2}\right) + K^2}\ . \tag{2.136}$$

As seen in Fig. 2.7, the minima of these two energy bands $\epsilon_\pm$ in Eq. (2.136) change depending on the choice of the hopping ratio K/J and the magnetic flux ϕ. The minimization of these bands $\epsilon_\pm$ can be done by calculating the first- and second-order derivative with respect to k:

$$\frac{\partial\epsilon_\pm}{\partial k} = 2Ja\sin(ka)\left[\cos\left(\frac{\phi}{2}\right) \pm \frac{2J\cos(ka)\sin^2\left(\frac{\phi}{2}\right)}{\sqrt{4J^2\sin^2(ka)\sin^2\left(\frac{\phi}{2}\right) + K^2}}\right] \tag{2.137}$$

$$\begin{aligned}\frac{\partial^2\epsilon_\pm}{\partial k^2} &= 2Ja^2\cos(ka)\cos\left(\frac{\phi}{2}\right) \pm \frac{4J^2a^2\cos(2ka)\sin^2\left(\frac{\phi}{2}\right)}{\sqrt{4J^2\sin^2(ka)\sin^2\left(\frac{\phi}{2}\right) + K^2}} \\ &\quad \mp \frac{4J^4a^2\sin^2(2ka)\sin^2\left(\frac{\phi}{2}\right)}{\left[4J^2\sin^2(ka)\sin^2\left(\frac{\phi}{2}\right) + K^2\right]^{3/2}},\end{aligned} \tag{2.138}$$

which suggests that the first obvious minimum of the lower band $\epsilon_-(k)$ occurs at $k = 0$ if $K > 2J\tan(\phi/2)\sin(\phi/2)$ since its second-order derivative in Eq. (2.138) must be positive. If $K < 2J\tan(\phi/2)\sin(\phi/2)$, then the following equality must be satisfied to minimize the lower band $\epsilon_-(k)$:

$$\cos\left(\frac{\phi}{2}\right) = \frac{2J\cos(ka)\sin^2\left(\frac{\phi}{2}\right)}{\sqrt{4J^2\sin^2(ka)\sin^2\left(\frac{\phi}{2}\right) + K^2}}\ . \tag{2.139}$$

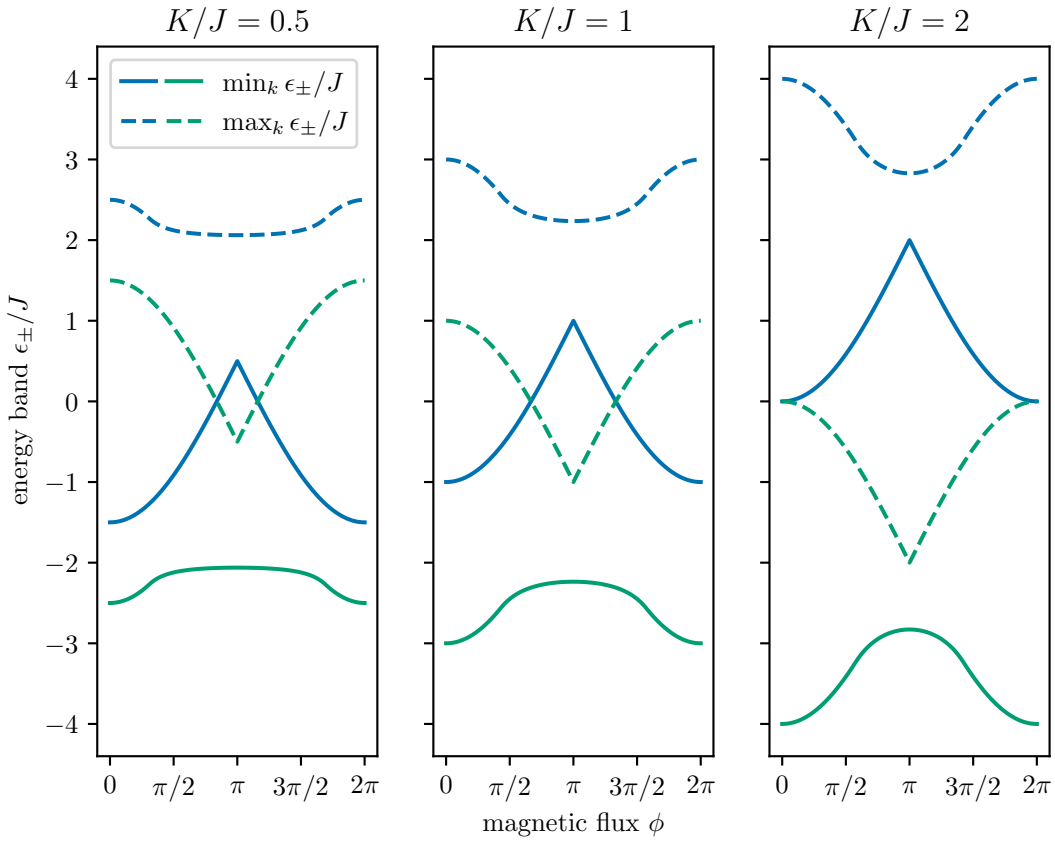

Figure 2.7: Normalized energy band minima (solid blue and green lines for the upper and lower bands, respectively) and maxima (dashed blue and green lines for the upper and lower bands, respectively) of the single-particle solution to a two-leg Bose-Hubbard ladder as a function of the magnetic flux ϕ for different hopping ratios $K/J = 0.5$ (left subplot), $K/J = 1$ (middle subplot) and $K/J = 2$ (right subplot). These plots can be considered as the *Hofstadter butterflies* of the two-leg flux ladder system for varying hopping ratio K/J.

This equality can be rewritten as

$$\sin^2(ka) = \sin^2\left(\frac{\phi}{2}\right) - \frac{K^2}{4J^2\tan^2\left(\frac{\phi}{2}\right)}, \tag{2.140}$$

which defines the other two arguments of the minimum as $k = \pm k_0$ where

$$k_0 \coloneqq \frac{1}{a}\arcsin\left(\sqrt{\sin^2\left(\frac{\phi}{2}\right) - \frac{K^2}{4J^2\tan^2\left(\frac{\phi}{2}\right)}}\right), \tag{2.141}$$

which is shown in Fig. 2.8. Additionally, the critical magnetic flux to divert the arguments of the minimum can be obtained by rewriting $K =$

$2J\tan(\phi_c/2)\sin(\phi_c/2)$ as

$$\phi_c = 2\arccos\left(-\frac{K}{4J} \pm \sqrt{\left(\frac{K}{4J}\right)^2 + 1}\right), \tag{2.142}$$

so that a single minimum occurs at $k = 0$ for $|\phi| < \phi_c$, and a degenerate minimum occurs at $k = \pm k_0$ for $|\phi| > \phi_c$, which is plotted in Fig. 2.9, as also indicated in [15].

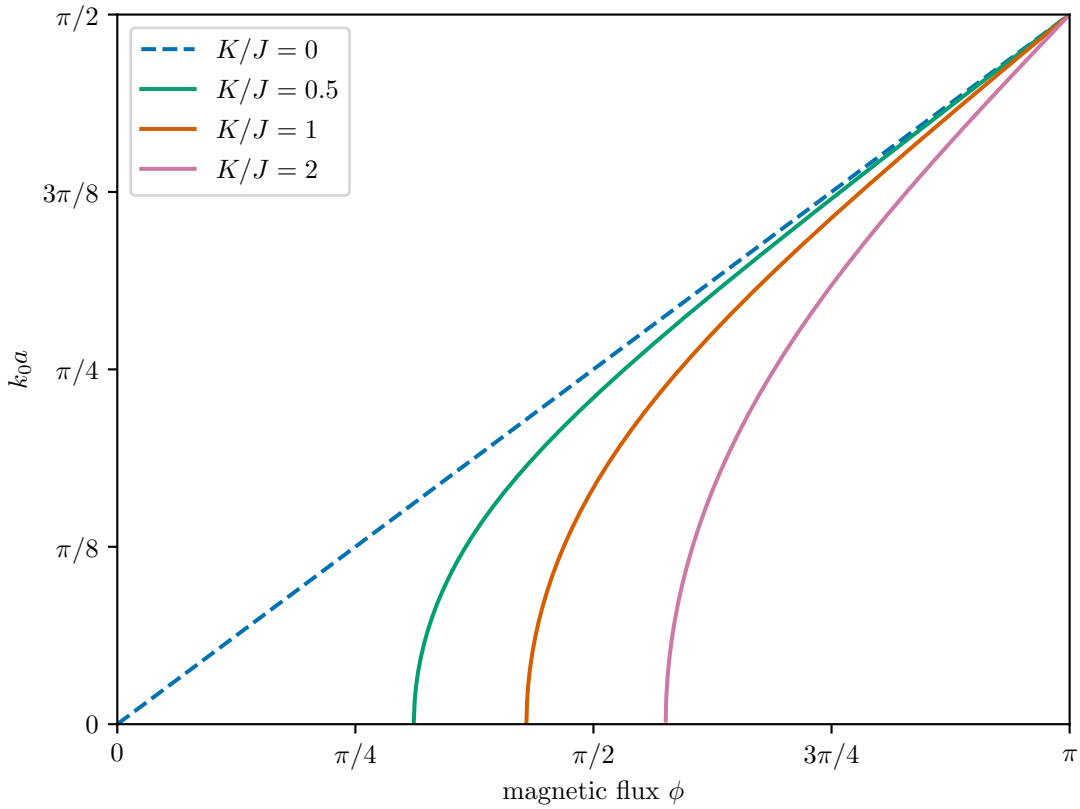

Figure 2.8: Wave number k_0, corresponding to the argument of degenerate energy band minimum, times the lattice constant a as a function of the magnetic flux ϕ at three hopping ratios $K/J = 0$ (blue dashed line), $K/J = 0.5$ (green line), $K/J = 1$ (orange line) and $K/J = 2$ (purple line) for the single-particle solution to a two-leg Bose-Hubbard ladder.

As [58] suggests, the following mean-field ansatz for the SF ground states is considered:

$$|G_{k_0}\rangle := \frac{1}{\sqrt{N!}}\left[\cos(\omega)\hat{\beta}^\dagger_{k_0} + \sin(\omega)\hat{\beta}^\dagger_{-k_0}\right]^N |0\rangle\,, \tag{2.143}$$

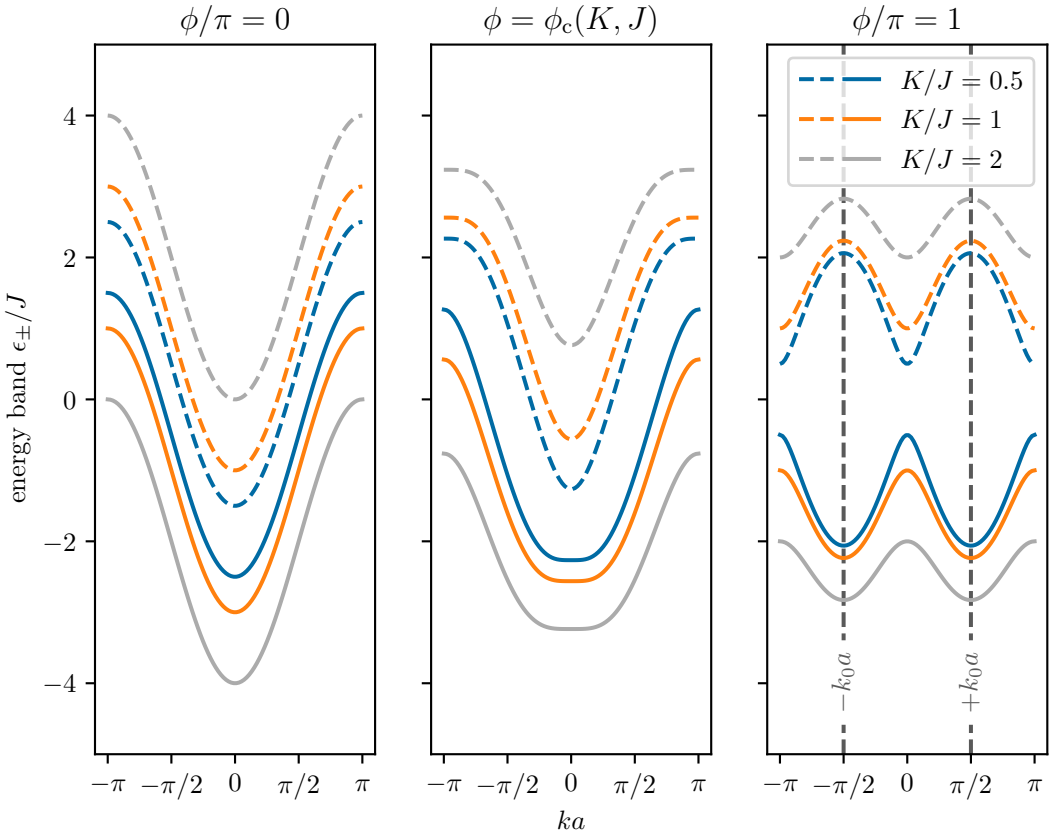

Figure 2.9: Normalized energy bands as a function of the wave number k times the lattice constant a at three magnetic fluxes $\phi = 0$ (left subplot), $\phi = \phi_c(K, J)$ (middle subplot) and $\phi = 1$ (right subplot) for different hopping ratios $K/J = 0.5$ (blue line), $K/J = 1$ (orange line) and $K/J = 2$ (grey line). The solid and dashed lines indicate the lower and upper bands, respectively.

where $0 < \omega < \phi/2$ for $k_0 > 0$ and $\omega = 0$ for $k_0 = 0$.

From this ansatz in Eq. (2.143), several calculations can be carried out, such as the net current j_n, which must be zero at equilibrium according to the continuity equation regardless of the value of the variable ω; thus it turns out to be

$$j_n \coloneqq \frac{1}{N}\sum_k \langle G_{k_0}|\hat{\mathbf{c}}_k^\dagger \frac{\partial \mathbf{h}_k}{\partial k}\hat{\mathbf{c}}_k|G_{k_0}\rangle\,, \tag{2.144}$$

where the operator $\hat{\mathbf{c}}_k^\dagger \frac{\partial \mathbf{h}_k}{\partial k}\hat{\mathbf{c}}_k$ results as

$$\hat{\mathbf{c}}_k^\dagger \frac{\partial \mathbf{h}_k}{\partial k}\hat{\mathbf{c}}_k = 2Ja\left[\sin\left(ka + \frac{\phi}{2}\right)\hat{n}_k^{(a)} + \sin\left(ka - \frac{\phi}{2}\right)\hat{n}_k^{(b)}\right], \tag{2.145}$$

and $\hat{n}_k^{(q)} \coloneqq \hat{q}_k^\dagger \hat{q}_k$ and $q \in \{a, b\}$.

To determine $\langle G_{k_0}|\hat{n}_k^{(q)}|G_{k_0}\rangle$, firstly, $\hat{q}_k\,|G_{k_0}\rangle$ must be obtained. After using some straightforward but lengthy commutation relations, it becomes

$$\hat{q}_k\,|G_{k_0}\rangle \;=\; \frac{N}{\sqrt{N!}}[\hat{q}_k, \hat{\nu}_{k_0}]\hat{\nu}_{k_0}^{N-1}\,|0\rangle\,, \tag{2.146}$$

where $\hat{\nu}_{k_0} \coloneqq \cos(\omega)\hat{\beta}_{k_0}^\dagger + \sin(\omega)\hat{\beta}_{-k_0}^\dagger$ and

$$\Big[\hat{a}_k, \hat{\nu}_{k_0}\Big] \;=\; -\sin(\theta_k)[\delta_{k,k_0}\cos(\omega) + \delta_{k,-k_0}\sin(\omega)] \tag{2.147}$$

$$\Big[\hat{b}_k, \hat{\nu}_{k_0}\Big] \;=\; \cos(\theta_k)[\delta_{k,k_0}\cos(\omega) + \delta_{k,-k_0}\sin(\omega)], \tag{2.148}$$

so that

$$\langle G_{k_0}|\hat{n}_k^{(q)}|G_{k_0}\rangle = N|[\hat{q}_k, \hat{\nu}_{k_0}]|^2\,. \tag{2.149}$$

Using Eqs. (2.147), (2.148) and (2.149), the net current is obtained as

$$j_{\rm n} = \cos^2(\omega)\Big(j_{k_0}^{(a)} + j_{k_0}^{(b)}\Big) + \sin^2(\omega)\Big(j_{-k_0}^{(a)} + j_{-k_0}^{(b)}\Big), \tag{2.150}$$

where the currents on each leg are

$$j_{k_0}^{(a)} \;\coloneqq\; 2Ja\sin\left(k_0 a + \frac{\phi}{2}\right)\sin^2(\theta_{k_0}) \tag{2.151}$$

$$j_{k_0}^{(b)} \;\coloneqq\; 2Ja\sin\left(k_0 a - \frac{\phi}{2}\right)\cos^2(\theta_{k_0})\,. \tag{2.152}$$

It is noted that Eq. (2.135) suggests the following relation:

$$j_{k_0}^{(a)} + j_{k_0}^{(b)} = \left.\frac{\partial\epsilon_-}{\partial k}\right|_{k_0} = 0\,. \tag{2.153}$$

Similarly, $j_{-k_0}^{(a)} + j_{-k_0}^{(b)} = 0$, so the net current vanishes, *i.e.*

$$j_{\rm n} = 0, \tag{2.154}$$

which is independent of the values of ω and k_0, as expected.

Rather than the net current $j_{\rm n}$, another measure can be the chiral current $j_{\rm c}$. It designates the chirality in the ladder system with the current difference

between the legs, which can be calculated, as follows:

$$j_{\mathrm{c}} \coloneqq \frac{1}{N}\sum_{k}\langle G_{k_0}|\hat{\mathbf{c}}_k^{\dagger}\boldsymbol{\sigma}_z\frac{\partial \mathbf{h}_k}{\partial k}\hat{\mathbf{c}}_k|G_{k_0}\rangle \tag{2.155}$$

$$= \cos^2(\omega)\left(j_{k_0}^{(a)}-j_{k_0}^{(b)}\right)+\sin^2(\omega)\left(j_{-k_0}^{(a)}-j_{-k_0}^{(b)}\right) \tag{2.156}$$

$$= 4Ja\sin\left(k_0a+\frac{\phi}{2}\right)\sin^2(\theta_{k_0})\ . \tag{2.157}$$

Addition to these currents, the local particle density $n_{\ell,m}$ for each site is defined using Eq. (2.149), as follows:

$$n_{\ell,m} = \langle G_{k_0}|\hat{n}_{\ell,m}|G_{k_0}\rangle = N|[\hat{a}_{\ell,m},\hat{\nu}_{k_0}]|^2\ . \tag{2.158}$$

By using the relation $\sin^2(\theta_{k_0}) = \cos^2(\theta_{-k_0})$ and the definition of mean density $\bar{n} \coloneqq N/L$, it turns out to be

$$n_{\ell,m} = \bar{n}_\ell + \delta n_{\ell,m}, \tag{2.159}$$

where

$$\begin{aligned}\bar{n}_\ell/\bar{n} \coloneqq\ & \cos^2(\omega)\left[\delta_{\ell,\mathrm{u}}\sin^2(\theta_{k_0})+\delta_{\ell,\mathrm{d}}\cos^2(\theta_{k_0})\right] \\ & +\sin^2(\omega)\left[\delta_{\ell,\mathrm{u}}\cos^2(\theta_{k_0})+\delta_{\ell,\mathrm{d}}\sin^2(\theta_{k_0})\right],\end{aligned} \tag{2.160}$$

and

$$\delta n_{\ell,m}/\bar{n} \coloneqq \frac{1}{2}\sin^2(\omega)\sin(2\theta_{k_0})\cos(2k_0ma)\ . \tag{2.161}$$

Combining the on-site interaction Hamiltonian $\hat{\mathcal{H}}_{\mathrm{int}}$ with the single-body Hamiltonian $\hat{\mathcal{H}}_{\mathrm{t}}$, the variational energy per boson for the ansatz in Eq. (2.143) becomes

$$E(\omega,k) \coloneqq \frac{1}{N}\langle G_k|\hat{\mathcal{H}}_{\mathrm{t}}+\hat{\mathcal{H}}_{\mathrm{int}}|G_k\rangle \tag{2.162}$$

$$= \epsilon_-(k)+\epsilon_{\mathrm{int}}(k,\omega), \tag{2.163}$$

where

$$\epsilon_{\mathrm{int}}(k,\omega) \coloneqq \frac{1}{N}\langle G_k|\hat{\mathcal{H}}_{\mathrm{int}}|G_k\rangle \tag{2.164}$$

$$= \frac{U}{2N}\sum_{m=1}^{L}\sum_{\ell\in\{\mathrm{u},\mathrm{d}\}}\left(\langle G_{k_0}|\hat{n}_{\ell,m}^2|G_{k_0}\rangle - n_{\ell,m}\right)\ . \tag{2.165}$$

$n_{\ell,m}$ is already known, and $\langle G_{k_0}|\hat{n}^2_{\ell,m}|G_{k_0}\rangle$ can be determined by using Eq. (2.146), as follows:

$$\begin{aligned}
\langle G_{k_0}|\hat{n}^2_{\ell,m}|G_{k_0}\rangle &= \frac{N^2}{N!}|[\hat{a}_{\ell,m},\hat{\nu}_k]|^2\left\langle 0\middle|\left(\hat{\nu}_k^\dagger\right)^{N-1}(\hat{n}_{\ell,m}+1)\,\hat{\nu}_k^N\middle|0\right\rangle && (2.166)\\
&= n_{\ell,m}+\frac{N-1}{N}n^2_{\ell,m} && (2.167)\\
&\approx n_{\ell,m}+n^2_{\ell,m}, && (2.168)
\end{aligned}$$

which leads to

$$\begin{aligned}
\epsilon_{\rm int}(k,\omega) &= \frac{U}{2N}\sum_{m=1}^{L}\sum_{\ell\in\{\rm u,d\}} n^2_{\ell,m} && (2.169)\\
&= \frac{U\bar{n}^2}{2N}\Bigg[L\big(\bar{n}_{\rm u}^2+\bar{n}_{\rm d}^2\big)+2(\bar{n}_{\rm u}+\bar{n}_{\rm d})\sum_m \delta n_{\ell,m} && (2.170)\\
&\qquad +2\sum_m(\delta n_{\ell,m})^2\Bigg]\,.
\end{aligned}$$

The first sum of the variational interaction energy in Eq. (2.170) vanishes because

$$\begin{aligned}
\sum_m \delta n_{\ell,m} &= \frac{1}{2}\sin(2\omega)\sin(2\theta_k)\sum_m\cos(2kma) && (2.171)\\
&= \delta_{k,0}\frac{L}{2}\sin(2\omega)\sin(2\theta_k) && (2.172)\\
&= 0\,. && (2.173)
\end{aligned}$$

The second sum also has a similar form as the previous one:

$$\sum_m(\delta n_{\ell,m})^2=\frac{1}{4}\sin^2(2\omega)\sin^2(2\theta_k)\,y(k), \tag{2.174}$$

where

$$y(k)\coloneqq\sum_m\cos^2(2kma)\,. \tag{2.175}$$

The function $y(k)$ can be determined by first differentiating it

$$\frac{\partial y}{\partial k}=-2a\sum_m m\sin(4kma), \tag{2.176}$$

and then integrating it

$$\int_{y(0)}^{y(k')} \mathrm{d}y = y(k') - L = -2a \sum_m \int_0^{k'} \sin(4kma)\,\mathrm{d}k \tag{2.177}$$

$$= \frac{1}{2}\sum_m [\cos(4kma) - 1] \tag{2.178}$$

$$= \frac{1}{2}L(\delta_{k,0} - 1), \tag{2.179}$$

which gives out

$$\sum_m (\delta n_{\ell,m})^2 = \frac{1}{8} L \sin^2(2\omega)\sin^2(2\theta_k)\ . \tag{2.180}$$

Using the result of these sums in Eqs. (2.173) and (2.180), the interaction energy ϵ_{int} results as

$$\epsilon_{\mathrm{int}}(k,\omega) = \frac{U\bar{n}}{2}\left\{1 - \frac{1}{2}\sin^2(2\theta_k) + \sin^2(2\omega)\left[\frac{3}{4}\sin^2(2\theta_k) - \frac{1}{2}\right]\right\}\ . \tag{2.181}$$

Since the derivative of the variational energy in Eq. (2.143) with respect to the variable ω only depends on the interaction energy

$$\frac{\partial E}{\partial \omega} = \frac{\partial \epsilon_{\mathrm{int}}}{\partial \omega} \tag{2.182}$$

$$= U\bar{n}\sin(4\omega)\left[\frac{3}{4}\sin^2(2\theta_k) - \frac{1}{2}\right], \tag{2.183}$$

it is then obvious that $E(k,\omega)$ is minimized at $\omega = 0$ if $\sin^2(2\theta_k) \geq 2/3$; otherwise, the energy minimum occurs at $\omega = \pi/4$.

A remarkable array of SF phases can be realized via this ansatz in this system, showcasing distinct vortex and local particle density profiles [27]. To simplify the discussion, the focus lies on three basic SF phases that can be tuned by altering the inter-leg hopping parameter K, while maintaining a fixed magnetic flux ϕ, which affects the choice of the variational parameter ω and the wave number k that minimizes the variational energy.

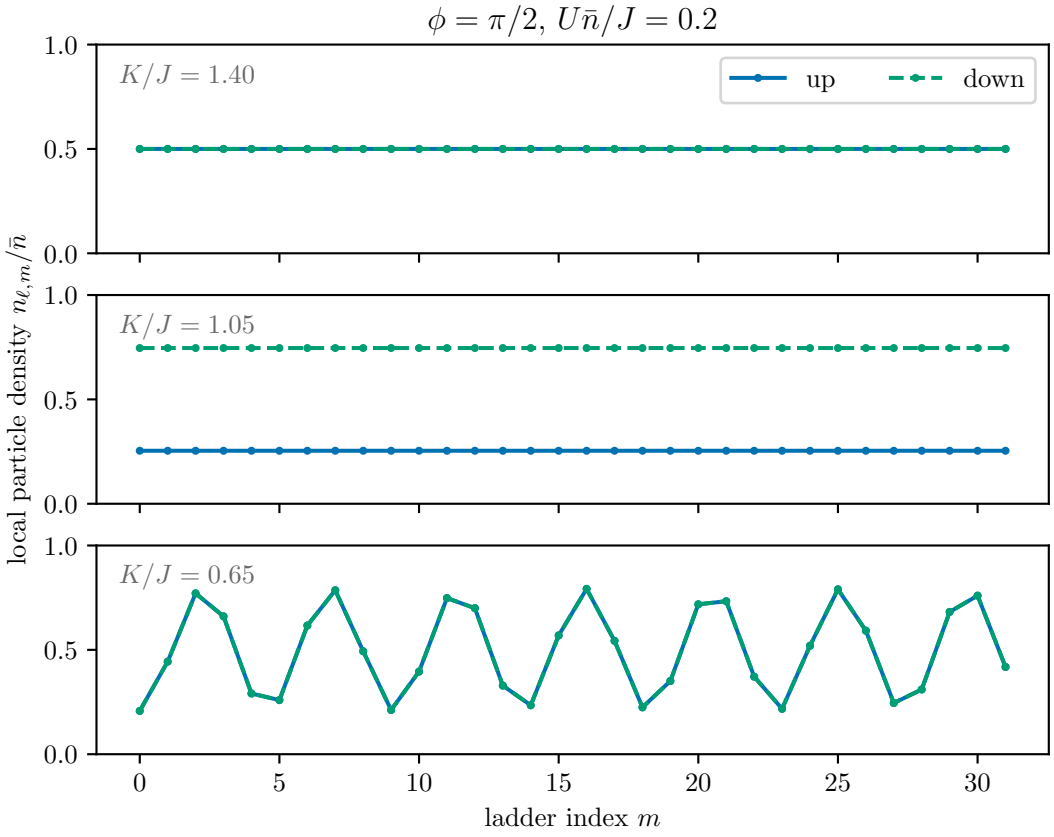

Figure 2.10: Local particle density as a function of the ladder index m for each leg (blue solid line for the up leg, green dashed line for the down leg) at a fixed magnetic flux $\phi = \pi/2$ and normalized on-site interaction strength $U\bar{n}/J = 0.2$ for different hopping ratios K/J via the mean-field approach. The particle density of the chiral, biased-ladder and vortex phases are drawn at the top ($K/J = 1.40$), middle ($K/J = 1.05$) and bottom ($K/J = 0.65$) of the figure, respectively.

2.2.1 Chiral phase

If $\omega = 0$ and $k = 0$ are the arguments that minimize the energy, the chiral current j_c becomes

$$j_c = j_{c,\max} \coloneqq 2Ja \sin\left(\frac{\phi}{2}\right), \tag{2.184}$$

and the local particle density for each site is invariant both along and across the legs, as shown in the top subplot of Figs. 2.10 and 2.11, so that

$$n_{\ell,m} = \frac{\bar{n}}{2}, \tag{2.185}$$

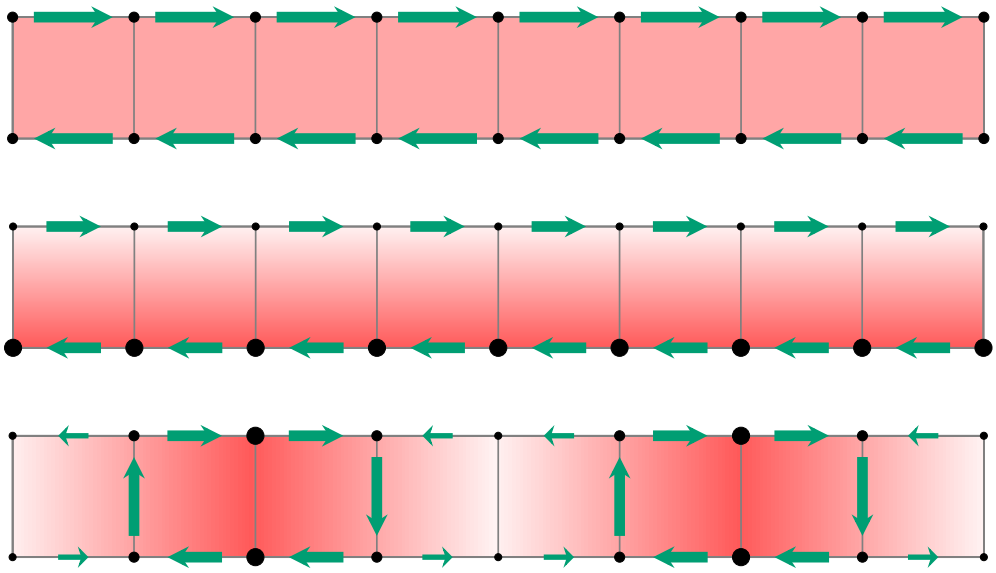

Figure 2.11: Representations of the superfluid phases in the weakly interacting regime on a two-leg ladder lattice. The direction and magnitude of the green arrows with their thickness and length represent the local currents between lattice sites. The black dots and the red background shadings indicate the local particle density on each site. The chiral, biased-ladder and vortex phases are figured at the top, middle and bottom of the figure, respectively.

where such a phase is called the *chiral phase* or *saturated-chiral-current phase*, which is also known as the *Meissner phase*. Here, the legs are strongly coupled, and the superfluid velocities across the legs are equal in magnitude but opposite in direction, without any rung current. Moreover, the only spontaneous broken symmetry is the $U(1)$ condensate phase symmetry.

2.2.2 Biased-ladder phase

If $\omega = 0$ and $k = \pm k_0$, the local particle density differentiates with a constant value in each leg, as shown in the middle subplot of Figs. 2.10 and 2.11, which is defined as

$$n_{\ell,m} = \bar{n}\left[\delta_{\ell,\mathrm{u}} \sin^2(\theta_k) + \delta_{\ell,\mathrm{d}} \cos^2(\theta_k)\right], \tag{2.186}$$

which indicates that the particle density is larger in one leg, and the superfluid velocity is larger in the other leg, which exhibits the *biased-ladder (BL) phase*. There are still no rung currents. Obviously, the $\mathbb{Z}_2$ reflection and $U(1)$ condensate phase symmetries are spontaneously broken.

2.2.3 Vortex phase

If $\omega = \pi/4$ and $k = \pm k_0$, a modulation appears in the local particle density, as shown in the bottom subplot of Figs. 2.10 and 2.11, so

$$n_{\ell,m} = \frac{\bar{n}}{2}[1 + \sin(2\theta_k)\cos(2kma)], \tag{2.187}$$

where its period increases with the hopping ratio K/J. The rung currents no longer vanish but are finite. This phase is called the *vortex phase* or *modulated-density phase*. Here, in addition to the spontaneous broken $U(1)$ condensate phase symmetry, another $U(1)$ symmetry is broken, where the variational energy is invariant under applying a phase to $\hat{\beta}^\dagger_{\pm k_0}$ of the ansatz in Eq. (2.143).

As seen in Fig. 2.12, the order of the phase transitions differentiates in this regime. The BL-chiral phase transition is second order, so the chiral current is continuous during the transition. For the vortex-BL phase transition, the order is first; therefore, a discontinuous jump in the chiral current occurs during the transition, where the jump size increases with the normalized on-site interaction strength $U\bar{n}/J$.

Rather than defined as in this mean-field approach, the local current operators along the legs and rungs can be defined more generally as

$$\hat{j}^{\parallel}_{\ell,m} := iJ\Big(e^{i\sigma_\ell\phi/2}\hat{a}^\dagger_{\ell,m+1}\hat{a}_{\ell,m} - \text{h.c.}\Big) \tag{2.188}$$

$$j^{\perp}_{m} := iK\big(\hat{a}^\dagger_{\text{u},m}\hat{a}_{\text{d},m} - \text{h.c.}\big), \tag{2.189}$$

which is utilized to write the chiral current operator as

$$\hat{j}_{\text{c}} = \frac{1}{L}\left|\sum_m \left(\hat{j}^{\parallel}_{\text{u},m} - \hat{j}^{\parallel}_{\text{d},m}\right)\right| . \tag{2.190}$$

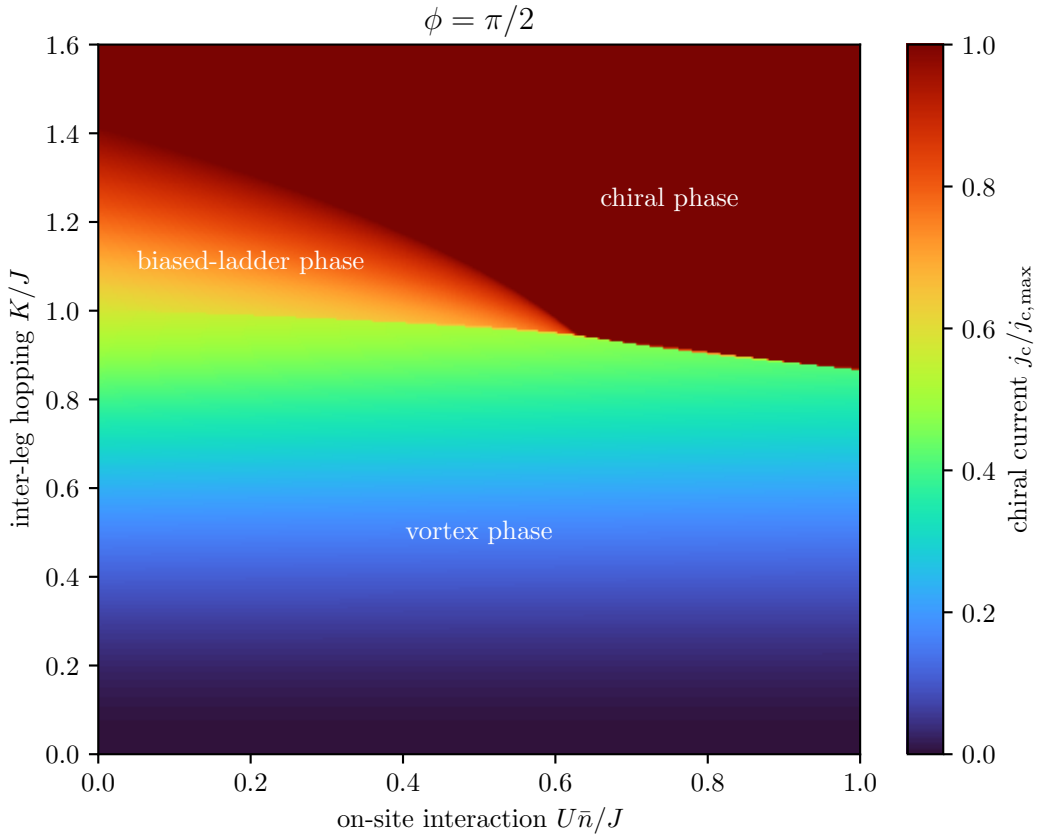

Figure 2.12: Mean-field phase diagram of a two-leg Bose-Hubbard ladder as a normalized chiral current function of the normalized on-site interaction strength $U\bar{n}/J$ and the hopping ratio K/J at a fixed magnetic flux $\phi = \pi/2$ via the mean-field approach.

These superfluid phases can also be distinguished with two order parameters. The first is the leg population imbalance ΔN — the boson population difference between the legs — defined as

$$\Delta N \coloneqq \frac{\left|\sum_{m=1}^{L} \left(n_{\mathrm{u},m} - n_{\mathrm{d},m}\right)\right|}{N}, \tag{2.191}$$

which has a non-zero value for the BL phase, but vanishes for the chiral and vortex phases. The second one is the absolute average value $j_{|\perp|}$ of the rung current, defined as

$$j_{|\perp|} \coloneqq \frac{1}{L} \sum_{m=1}^{L} \left|\left\langle \hat{j}_m^{\perp} \right\rangle\right|, \tag{2.192}$$

which becomes zero for the chiral and BL phases, and non-zero for the vortex

phase, where all these order parameters are summarized in Table 2.1.

superfluid phase	ΔN	$j_{\vert\perp\vert}$
chiral	0	0
biased-ladder	> 0	0
vortex	0	> 0

Table 2.1: Order parameters for the superfluid phases occurring in the weakly interacting regime for a two-leg Bose-Hubbard ladder under magnetic flux.

2.2.4 Gutzwiller variational approach

In the strongly interacting regime, the phases are well-known, which are nothing but SF and MI phases. As it is known, a perfect MI ground state can be defined as in Eq. (2.5) in the local Fock basis $|n_0\rangle_{\ell,m}$. Therefore, a SF ground state can be represented on the same basis with a particle-hole excitation around the perfect MI state as a small variation, which leads to the following site-dependent Gutzwiller variational ansatz for each lattice site:

$$|G_{\ell,m}\rangle \coloneqq \Delta_{\ell,m}\,|n_0-1\rangle_{\ell,m} + |n_0\rangle_{\ell,m} + \Delta'_{\ell,m}\,|n_0+1\rangle_{\ell,m}\,, \tag{2.193}$$

where $\mathbf{\Delta} \coloneqq \left\{\Delta_{\ell,m}, \Delta'_{\ell,m}\right\}$ are a set of small complex variational parameters. Thus, the unnormalized site-independent Gutzwiller variational ansatz can be written as

$$\left|\tilde{\Psi}_G\right\rangle \coloneqq \bigotimes_{m=1}^{L} \bigotimes_{\ell\in\{\mathrm{u,d}\}} |G_{\ell,m}\rangle\,, \tag{2.194}$$

With this definition, the variational energy up to the second order in $\mathbf{\Delta}$ results as

$$\begin{aligned}\tilde{E}^{(2)}(\mathbf{\Delta}) \;=\; & -\tilde{J}\sum_{m=1}^{L}\sum_{\ell\in\{\mathrm{u,d}\}} e^{-\sigma_\ell\phi/2} f(\ell,m;\ell,m+1) \\ & -\tilde{K}\sum_{m=1}^{L} f(\mathrm{u},m;\mathrm{d},m) \\ & +\sum_{m=1}^{L}\sum_{\ell\in\{\mathrm{u,d}\}} g(\ell,m;\tilde{\mu}),\end{aligned} \tag{2.195}$$

where the tilde sign corresponds to scaling with respect to U,

$$\begin{aligned} f(\ell,m;\ell',m') &\coloneqq n_0\Delta_{\ell,m}\Delta^*_{\ell',m'} + (n_0+1)(\Delta'_{\ell,m})^*\Delta'_{\ell',m'} \\ &\quad + \sqrt{n_0(n_0+1)}\left[\Delta_{\ell,m}\Delta'_{\ell',m'} + (\Delta'_{\ell,m})^*\Delta^*_{\ell',m'}\right], \end{aligned} \tag{2.196}$$

and

$$g(\ell,m;\tilde{\mu}) \coloneqq (1-n_0+\tilde{\mu})|\Delta_{\ell,m}|^2 + \frac{1}{2}n_0(n_0-1-2\tilde{\mu}) + (n_0-\tilde{\mu})\left|\Delta'_{\ell,m}\right|^2 . \tag{2.197}$$

It is obvious to see that the system is in the MI phase if the variational energy is minimized where all variational parameters $\boldsymbol{\Delta}$ become zero. This allows one to determine the SF-MI phase boundary where the variational energy turns out to be a local minimum in $\boldsymbol{\Delta}$. To do so, its Hessian matrix $\mathbf{H}_{\tilde{E}}$ with respect to the variational parameters should be positive definite so that its all eigenvalues are positive. This matrix can be written in the following compact way:

$$\mathbf{H}_{\tilde{E}} = -\tilde{J}\begin{pmatrix} n_0\boldsymbol{\mathcal{S}} & \sqrt{n_0(n_0+1)}\boldsymbol{\mathcal{S}} \\ \sqrt{n_0(n_0+1)}\boldsymbol{\mathcal{S}} & (n_0+1)\boldsymbol{\mathcal{S}} \end{pmatrix} \tag{2.198}$$

$$+\begin{pmatrix} (1-n_0+\tilde{\mu})\mathbb{I}_{N\times N} & \mathbf{0}_{N\times N} \\ \mathbf{0}_{N\times N} & (n_0-\tilde{\mu})\mathbb{I}_{N\times N} \end{pmatrix}, \tag{2.199}$$

where $\mathbb{I}$ and $\mathbf{0}$ are $N\times N$ identity and zero matrices, respectively,

$$\boldsymbol{\mathcal{S}} \coloneqq \begin{pmatrix} \mathbf{A} & \mathbf{B} & \mathbf{0}_{2\times2} & \cdots & \mathbf{0}_{2\times2} & \mathbf{B}^\dagger \\ \mathbf{B}^\dagger & \mathbf{A} & \mathbf{B} & \mathbf{0}_{2\times2} & \cdots & \mathbf{0}_{2\times2} \\ \mathbf{0}_{2\times2} & \mathbf{B}^\dagger & \mathbf{A} & \mathbf{B} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \mathbf{0}_{2\times2} \\ \mathbf{0}_{2\times2} & \cdots & \mathbf{0}_{2\times2} & \mathbf{B}^\dagger & \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{0}_{2\times2} & \cdots & \mathbf{0}_{2\times2} & \mathbf{B}^\dagger & \mathbf{A} \end{pmatrix}_{N\times N}, \tag{2.200}$$

which is nothing but a single-particle matrix, and the submatrices $\mathbf{A}$ and $\mathbf{B}$ are defined as

$$\mathbf{A} \coloneqq \frac{\tilde{K}}{\tilde{J}}\boldsymbol{\sigma}_x \tag{2.201}$$

$$\mathbf{B} \coloneqq \cos\left(\frac{\phi}{2}\right)\boldsymbol{\sigma}_0 + i\sin\left(\frac{\phi}{2}\right)\boldsymbol{\sigma}_z, \tag{2.202}$$

as discussed in [25, 74]. Here, the highest eigenvalue $\lambda_{\boldsymbol{\mathcal{S}}}$ of $\boldsymbol{\mathcal{S}}$ in the limit $L \to \infty$ corresponds to the maximized value of $\epsilon_+(k)/J$ in Eq. (2.136) with respect to k. The eigenvalues $\lambda_{\mathbf{H},\pm}$ of the Hessian matrix $\mathbf{H}_{\tilde{E}}$ can be found in terms of this highest eigenvalue if the eigenvectors of $\mathbf{H}_{\tilde{E}}$ are defined as

$$\mathbf{u} \coloneqq \begin{pmatrix} a\mathbf{v} \\ b\mathbf{v} \end{pmatrix}, \tag{2.203}$$

where $\mathbf{v}$ denotes the eigenvector of the single-particle matrix $\boldsymbol{\mathcal{S}}$ for the highest eigenvalue $\lambda_{\boldsymbol{\mathcal{S}}}$. Then, $\lambda_{\mathbf{H},\pm}$ is obtained as

$$\begin{aligned} \lambda_{\mathbf{H},\pm} \;=\; & \frac{1}{2}\Bigg[1-(2n_0+1)\tilde{J}\lambda_{\boldsymbol{\mathcal{S}}} \\ & \pm\sqrt{4(\tilde{\mu}-n_0)(1-n_0+\tilde{\mu})+4(\tilde{\mu}+1)\tilde{J}\lambda_{\boldsymbol{\mathcal{S}}}+\Big[1-(2n_0+1)\tilde{J}\lambda_{\boldsymbol{\mathcal{S}}}\Big]^2}\Bigg]. \end{aligned} \tag{2.204}$$

To determine the SF-MI phase boundary, the minimum eigenvalue $\lambda_{\mathbf{H},-}$ must be equal zero, which leads to find

$$\tilde{J}_{\mathrm{b}} = \left.\frac{(n_0-\tilde{\mu})(1-n_0+\tilde{\mu})}{(\tilde{\mu}+1)\lambda_{\boldsymbol{\mathcal{S}}}}\right|_{\tilde{K}_{\mathrm{b}},\tilde{J}_{\mathrm{b}}}, \tag{2.205}$$

which can be rewritten as

$$\tilde{\mu}_{\pm} = \left.\frac{1}{2}\left[2n_0-1-\tilde{J}_{\mathrm{b}}\lambda_{\boldsymbol{\mathcal{S}}} \pm \sqrt{\left(\tilde{J}_{\mathrm{b}}\lambda_{\boldsymbol{\mathcal{S}}}\right)^2-2\tilde{J}_{\mathrm{b}}\lambda_{\boldsymbol{\mathcal{S}}}(2n_0+1)+1}\right]\right|_{\tilde{K}_{\mathrm{b}},\tilde{J}_{\mathrm{b}}}, \tag{2.206}$$

where $n_0 - 1 \leq \tilde{\mu} \leq n_0$.

As can be seen in Fig. 2.13 where Eq. (2.206) for various hopping ratios K/J is plotted, a change in the magnetic flux ϕ affects the Mott lobes and thus their tips. The regions occupied by the Mott lobes grow with the modulus of the magnetic flux up to $|\phi| = \pi$ and shrink up to $|\phi| = 2\pi$.

It is also noted that the Gutzwiller approach is equivalent to the mean-field theory [75–77] in a two-leg flux ladder system [25, 74], which is not expected to have an accurate result because the effect of particle fluctuations in low-dimensional systems cannot be ignored. Nevertheless, this mean-field result gives a considerable impression to understanding such a two-leg ladder system under a magnetic field.

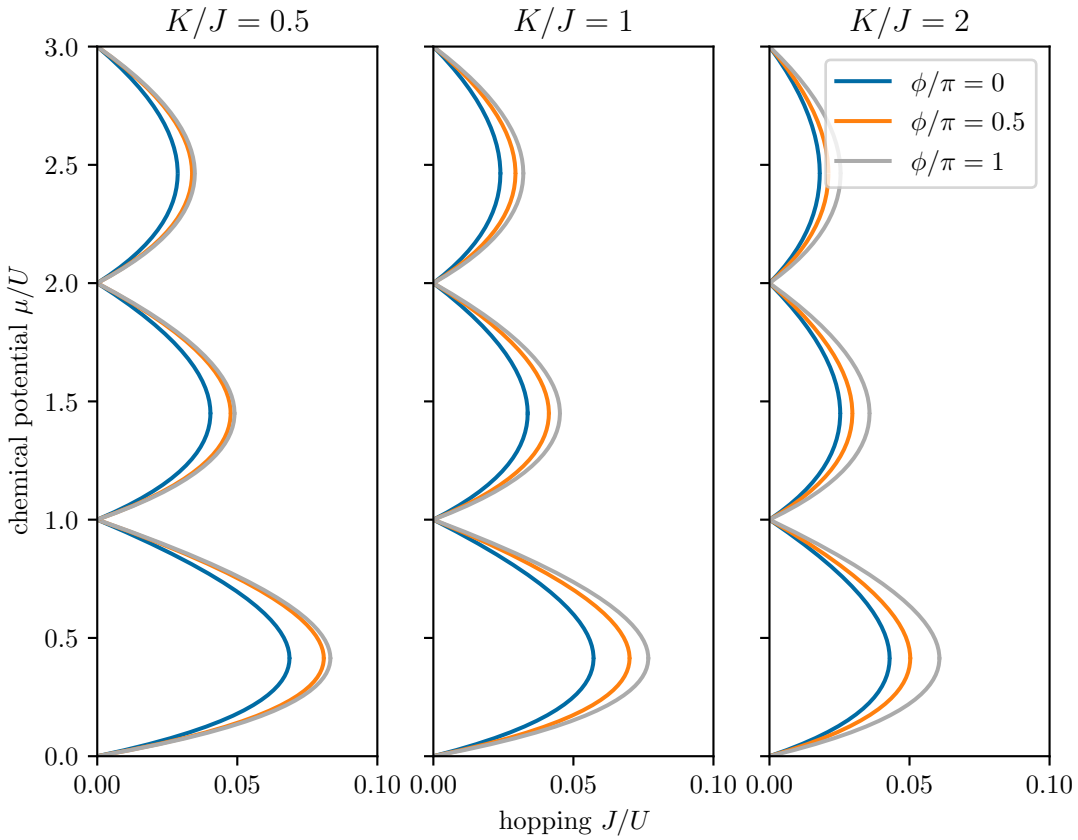

Figure 2.13: Mean-field phase diagrams of a two-leg Bose-Hubbard ladder at different hopping ratios $K/J = 0.5$ (left subplot), $K/J = 1$ (middle subplot) and $K/J = 2$ (right subplot) for the normalized magnetic fluxes $\phi/\pi = 0$ (blue line), $\phi/\pi = 0.5$ (orange line) and $\phi/\pi = 1$ (grey line), which is obtained with the Gutzwiller variational approach.

2.2.5 Strong-coupling perturbation theory

As discussed in the case of the vanilla Bose-Hubbard model, much more accurate boundaries can be obtained by using the strong-coupling expansion with the perturbation theory rather than the mean-field approach. To do that, the Hamiltonian is rewritten as

$$\hat{\mathcal{H}} = -J \sum_{i,j} \left(S_{i,j} \hat{a}_i^\dagger \hat{a}_j + \text{h.c.} \right) + \frac{U}{2} \sum_i \hat{n}_i (\hat{n}_i - 1) - \mu \sum_i \hat{n}_i, \tag{2.207}$$

by using the resulting normalized single-particle matrix $\boldsymbol{\mathcal{S}} = \{S_{i,j}\}$ defined in Eq. (2.200) in the previous section. After some lengthy but straightforward calculations, as done in [25], the normalized perturbative energies up to the second order result as

$$\tilde{E}^{(2)}_{n_0} = N_{\rm s}\left[\frac{1}{2}n_0(n_0-1) - \tilde{\mu}n_0 - \left(2\tilde{J}^2+\tilde{K}^2\right)n_0(n_0+1)\right] \tag{2.208}$$

$$\begin{aligned}\tilde{E}^{(2)}_{+} &= \tilde{E}^{(2)}_{n_0} + n_0 - \tilde{\mu} - \tilde{J}\lambda_{\boldsymbol{\mathcal{S}}}(n_0+1) - \left(\tilde{J}\lambda_{\boldsymbol{\mathcal{S}}}\right)^2 n_0(n_0+1) \\ &\quad + \frac{1}{2}\left(2\tilde{J}^2+\tilde{K}^2\right)n_0(5n_0+4)\end{aligned} \tag{2.209}$$

$$\begin{aligned}\tilde{E}^{(2)}_{-} &= \tilde{E}^{(2)}_{n_0} - (n_0-1) + \tilde{\mu} - \tilde{J}\lambda_{\boldsymbol{\mathcal{S}}}n_0 - \left(\tilde{J}\lambda_{\boldsymbol{\mathcal{S}}}\right)^2 n_0(n_0+1) \\ &\quad + \frac{1}{2}\left(2\tilde{J}^2+\tilde{K}^2\right)(n_0+1)(5n_0+1),\end{aligned} \tag{2.210}$$

which are similar to the energies without magnetic field in Eqs. (2.107), (2.114) and (2.115). Here, the effect of the magnetic field appears in the eigenvalue $\lambda_{\boldsymbol{\mathcal{S}}}$. Solving $\tilde{E}^{(2)}_{+} - \tilde{E}^{(2)}_{n_0} = 0$ and $\tilde{E}^{(2)}_{n_0} - \tilde{E}^{(2)}_{-} = 0$ gives the boundaries $\tilde{\mu}^{(2)}_{\pm}$ of the Mott lobes, as follows:

$$\begin{aligned}\tilde{\mu}^{(2)}_{+} &= n_0 - \tilde{J}\lambda_{\boldsymbol{\mathcal{S}}}(n_0+1) - \left(\tilde{J}\lambda_{\boldsymbol{\mathcal{S}}}\right)^2 n_0(n_0+1) \\ &\quad + \frac{1}{2}\left(2\tilde{J}^2+\tilde{K}^2\right)n_0(5n_0+4)\end{aligned} \tag{2.211}$$

$$\begin{aligned}\tilde{\mu}^{(2)}_{-} &= (n_0-1) + \tilde{J}\lambda_{\boldsymbol{\mathcal{S}}}n_0 + (\tilde{J}\lambda_{\boldsymbol{\mathcal{S}}})^2 n_0(n_0+1) \\ &\quad - \frac{1}{2}\left(2\tilde{J}^2+\tilde{K}^2\right)(n_0+1)(5n_0+1)\ .\end{aligned} \tag{2.212}$$

When the phase boundaries are plotted as in Fig. 2.14 for the first three Mott lobes at various hopping ratios K/J, the first noticeable thing is that the second-order perturbation theory fails beyond around $\phi \approx 0.6$ for $K/J \approx 2$, and the phase boundaries unexpectedly cross each other at a large intra-leg hopping amplitude $\tilde{J}$, as indicated in [25]. Moreover, a phase transition from the MI phase to the SF phase and then back to the MI phase starts to occur along the J/U axis after $\phi \approx 0.6$ at a fixed chemical potential, which is called the *re-entrant behaviour*. This behaviour is normally unexpected in the mean-field and second-order strong-coupling results without magnetic flux, as shown in Figs. 2.3, 2.5, 2.13 and 2.14.

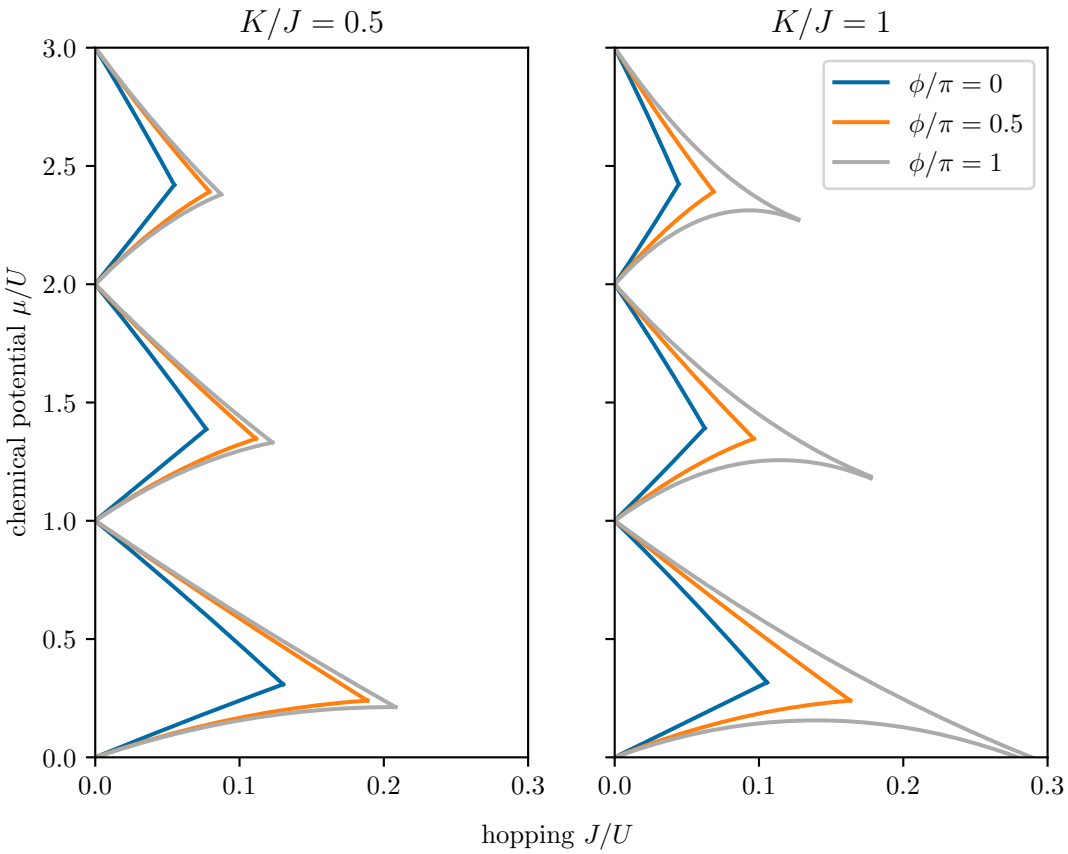

Figure 2.14: Phase diagrams of a two-leg Bose-Hubbard ladder for different hopping ratios $K/J = 0.5$ (left subplot) and $K/J = 1$ (right subplot) for the normalized magnetic fluxes $\phi/\pi = 0$ (blue line), $\phi/\pi = 0.5$ (orange line) and $\phi/\pi = 1$ (grey line), obtained with the strong-coupling perturbation theory.

Chapter 3

Neural-network quantum states

This chapter reviews neural-network quantum states (NQS) in the context of the Bose-Hubbard model and their stochastic optimization. Two different neural network architectures are considered for calculating the ground-state wave function of the Bose-Hubbard models. The first architecture considered is a restricted Boltzmann machine (RBM), which Carleo and Troyer [33] initially introduced for describing quantum spin models. It has also been applied to analyze the phase diagram of the one-dimensional Bose-Hubbard model [54]. The second architecture is a feedforward neural network (FNN), which is well-known in machine learning and has recently been applied to finding the ground state of the vanilla Bose-Hubbard model [55, 56]. Next, the measurement of expectation values, sampling a many-body configuration from the Fock space using a Markov chain Monte Carlo method, and the optimization techniques employed for these ansatzes, considering various approaches, are discussed.

3.1 Neural-network architectures

The choice of neural-network architecture makes wave function ansatzes different; however, any neural-network quantum state can be formulated, as follows:

$$|\Psi_{\boldsymbol{\mathcal{W}}}\rangle = \hat{\mathbb{I}}\,|\Psi_{\boldsymbol{\mathcal{W}}}\rangle = \sum_{\{\boldsymbol{\mathcal{S}}\}} |\boldsymbol{\mathcal{S}}\rangle\,\langle\boldsymbol{\mathcal{S}}|\Psi_{\boldsymbol{\mathcal{W}}}\rangle = \sum_{\{\boldsymbol{\mathcal{S}}\}} \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})\,|\boldsymbol{\mathcal{S}}\rangle\,, \tag{3.1}$$

where $\boldsymbol{\mathcal{W}}$ is the set of all network parameters of the chosen neural-network architecture and $\boldsymbol{\mathcal{S}}$ is the given many-body configuration.

3.1.1 Restricted Boltzmann machine

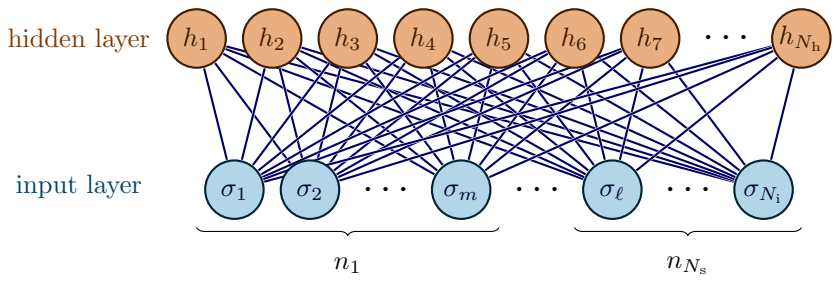

Figure 3.1: Restricted Boltzmann machine applied to the Bose-Hubbard model. The input layer $\boldsymbol{\mathcal{S}}$ corresponds to the physical space with the occupation number n_k, using a one-hot encoding with a maximum local occupation number $n_{\max}$. The number of input neurons is $N_\mathrm{i} = N_\mathrm{s} \times (n_{\max} + 1)$. The hidden layer consists of $N_\mathrm{h} = \alpha \times N_\mathrm{i}$ neurons, where $\alpha \in \mathbb{Z}^+$, with each neuron represented by a binary value $\{-1, 1\}$.

The RBM ansatz consists of input and hidden layers, similar to neural networks studied in machine learning, as illustrated in Fig. 3.1. The input layer represents the physical space, while the hidden layer represents an abstract space that determines the architecture of the variational parameters in the wave function. One-hot encoding is used for the bosonic occupation numbers in the physical space, denoted as $\boldsymbol{\mathcal{S}} = \{\,\sigma_j \mid \sigma_j \in \{\,0, 1\,\}\,\}_{j=1}^{N_\mathrm{i}}$, for a system of N_i sites in an arbitrary dimension. Each site in real space is represented by $n_k = \{\,0, 0, \ldots, 1, \ldots, 0\,\}$, where n_k has $n_{\max} + 1$ binary elements. Only the mth element is 1, while all

others are 0, indicating that there are $m-1$ bosons on the kth site. An upper bound is set by defining the maximum occupation number per site as $n_{\max}$.

The hidden layer consists of N_{h} neurons that take the values of -1 or 1. In the end, the variational wave function is written as a summation over the hidden layer neurons as follows:

$$\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) = \langle \boldsymbol{\mathcal{S}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle = \sum_{\{h_i\}} e^{\sum_i a_i h_i + \sum_j b_j \sigma_j + \sum_{ij} W_{ij} h_i \sigma_j}, \tag{3.2}$$

which is the probability amplitude of the given state $\boldsymbol{\mathcal{S}}$, and the dependence on the set of all variational parameters is shown with $\boldsymbol{\mathcal{W}} = \{ a_i, b_j, W_{ij} \}$ for $i = 1, \ldots, N_{\mathrm{h}}$ and $j = 1, \ldots, N_{\mathrm{i}}$, which are composed of biases and weights. Bias parameters for input and hidden layers are conventionally called a_i and b_j, respectively, and the weights of the links between the layers are called W_{ij}. In the RBM ansatz, any neuron in the visible layer is connected to all neurons in the hidden layer, but there is no connection between two neurons in the same layer; hence the name is *restricted*.

By performing the summation over $\{ h_i \}$ in Eq. (3.2) and assuming the restriction on connections, the hidden neurons can be traced out. For the variational ansatz, this results in the following compact expression

$$\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) = \langle \boldsymbol{\mathcal{S}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle = e^{\sum_j b_j \sigma_j} \prod_i \theta_i(\boldsymbol{\mathcal{S}}; \boldsymbol{\mathcal{W}}), \tag{3.3}$$

whose derivatives with respect to each variational network parameter can be calculated exactly without using automatic differentiation as follows:

$$\frac{1}{\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})} \frac{\partial}{\partial a_i} \left[\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \right] = \tanh \left[\theta_i(\boldsymbol{\mathcal{S}}; \boldsymbol{\mathcal{W}}) \right] \tag{3.4}$$

$$\frac{1}{\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})} \frac{\partial}{\partial b_j} \left[\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \right] = \sigma_j \tag{3.5}$$

$$\frac{1}{\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})} \frac{\partial}{\partial W_{ij}} \left[\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \right] = \tanh \left[\theta_i(\boldsymbol{\mathcal{S}}; \boldsymbol{\mathcal{W}}) \right] \sigma_j, \tag{3.6}$$

where

$$\theta_i(\boldsymbol{\mathcal{S}}; \boldsymbol{\mathcal{W}}) = 2 \cosh \left(a_i + \sum_j W_{ij} \sigma_j \right). \tag{3.7}$$

Given N_s lattice sites, N_h hidden neurons, and a given maximum occupation number n_max, the number of input neurons is $N_\mathrm{i} = (n_\mathrm{max} + 1) \times N_\mathrm{s}$. The density of hidden layer neurons is denoted by $\alpha = N_\mathrm{h}/N_\mathrm{i} \in \mathbb{Z}^+$, which can be adjusted to improve the accuracy of the ansatz. It is important to note that the network parameters can take real or complex values in the implementation, especially for systems with broken time-reversal symmetry.

3.1.2 Feedforward neural network

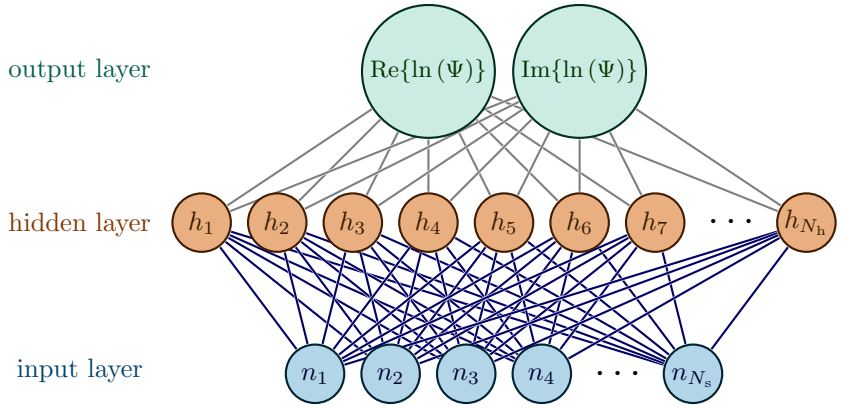

Figure 3.2: Feedforward neural network applied to the Bose-Hubbard model. The input layer consists of the site occupation numbers $\boldsymbol{\mathcal{S}} = \{\, n_k \mid n_k \in \mathbb{Z}_0^+ \,\}_{k=1}^{N_\mathrm{s}}$ without any cutoff on the maximum occupation number. The hidden layer consists of $N_\mathrm{h} = \alpha \times N_\mathrm{s}$ neurons, where $\alpha \in \mathbb{Z}^+$. The output layer comprises two neurons representing $\mathrm{Re}\{\ln(\Psi)\} = \ln|\Psi|$ and $\mathrm{Im}\{\ln(\Psi)\} = \Phi$, respectively, where $\Psi = |\Psi| e^{i\Phi}$ is the wave function.

The FNN ansatz, first introduced in [55] for the Bose-Hubbard model, consists of three layers, as depicted in Fig. 3.2. The input layer corresponds to the physical sites as in RBM, but each neuron takes integer values representing the number of bosons in that site, without imposing any cutoff on the maximum occupation. Specifically, the set is defined as $\boldsymbol{\mathcal{S}} = \{\, n_k \mid n_k \in \mathbb{Z}_0^+ \,\}_{k=1}^{N_\mathrm{s}}$. The hidden layer, which comprises N_h real-valued neurons, is obtained using the expression

$$h_j(\boldsymbol{\mathcal{S}}; \boldsymbol{\mathcal{V}}) = b_j + \sum_k V_{jk} n_k, \tag{3.8}$$

where $\boldsymbol{\mathcal{V}} = \{\, b_j, V_{jk} \,\}$ is the set of variational network parameters connecting the

input and hidden layers, b_j are the biases for the hidden layer, and V_{jk} are the weights of links between the layers.

The output layer consists of only two neurons, obtained from the hidden layer using the hyperbolic tangent activation function, given by

$$u_i(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}}) = a_i + \sum_j W_{ij} \tanh\left[h_j(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{V}})\right], \tag{3.9}$$

where $i = 1, 2$. Here $\boldsymbol{\mathcal{W}} = \{\, a_i, W_{ij}, b_j, V_{jk} \,\} \in \mathbb{R}^p$ is the combined set of all variational network parameters between the input and output layers.

The final variational wave function is expressed in terms of the neurons in the output layer as

$$\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) = \langle \boldsymbol{\mathcal{S}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle = e^{u_1(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}}) + i u_2(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}})}, \tag{3.10}$$

which is complex valued. All the parameters in the network are real valued, which provides a significant computational advantage compared to the RBM. The hidden layer neuron density is defined in a similar manner to the previous section, as $\alpha = N_{\mathrm{h}}/N_{\mathrm{s}}$.

3.2 Measurement of expectation values

Calculating the expectation value of an arbitrary operator is straightforward. Considering an operator $\hat{O}$, its expectation value in a many-body state with NQS ansatz results as

$$\left\langle \hat{O} \right\rangle = \frac{\langle \Psi_{\boldsymbol{\mathcal{W}}} | \hat{O} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle}{\langle \Psi_{\boldsymbol{\mathcal{W}}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle} \tag{3.11}$$

$$= \frac{\sum_{\{\boldsymbol{\mathcal{S}}\},\{\boldsymbol{\mathcal{S}}'\}} \langle \Psi_{\boldsymbol{\mathcal{W}}} | \boldsymbol{\mathcal{S}} \rangle \, \langle \boldsymbol{\mathcal{S}} | \hat{O} | \boldsymbol{\mathcal{S}}' \rangle \, \langle \boldsymbol{\mathcal{S}}' | \Psi_{\boldsymbol{\mathcal{W}}} \rangle}{\sum_{\{\boldsymbol{\mathcal{S}}'\}} \left| \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}') \right|^2} \tag{3.12}$$

$$= \frac{\sum_{\{\boldsymbol{\mathcal{S}}\}} \left| \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \right|^2 \sum_{\{\boldsymbol{\mathcal{S}}'\}} \langle \boldsymbol{\mathcal{S}} | \hat{O} | \boldsymbol{\mathcal{S}}' \rangle \, \frac{\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}')}{\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})}}{\sum_{\{\boldsymbol{\mathcal{S}}'\}} \left| \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}') \right|^2}, \tag{3.13}$$

where $\{\boldsymbol{\mathcal{S}}\}$ and $\{\boldsymbol{\mathcal{S}}'\}$ indicate a complete basis in the Fock space.

Eq. (3.13) can be simplified by dividing into the following two products:

$$O_{\mathrm{loc}}(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}}) = \sum_{\{\boldsymbol{\mathcal{S}}'\}} \langle \boldsymbol{\mathcal{S}}|\hat{O}|\boldsymbol{\mathcal{S}}'\rangle \frac{\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}')}{\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})} = \frac{\langle \boldsymbol{\mathcal{S}}|\hat{O}|\Psi_{\boldsymbol{\mathcal{W}}}\rangle}{\langle \boldsymbol{\mathcal{S}}|\Psi_{\boldsymbol{\mathcal{W}}}\rangle}, \tag{3.14}$$

which is the local estimator of an operator $\hat{O}$, and

$$p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) = \frac{|\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})|^2}{\sum_{\{\boldsymbol{\mathcal{S}}'\}} |\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}')|^2}, \tag{3.15}$$

which is the normalized probability distribution for the state $\boldsymbol{\mathcal{S}}$. With these choices, the expectation value in Eq. (3.13) can be reinterpreted as

$$\left\langle \hat{O} \right\rangle = \langle O_{\mathrm{loc}} \rangle = \sum_{\{\boldsymbol{\mathcal{S}}\}} p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})\, O_{\mathrm{loc}}(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}}) \ . \tag{3.16}$$

Since exponentially expanding Hilbert spaces cause the difficulty of calculating expectation values of operators, Markov chain Monte-Carlo (MCMC) sampling of a set $\left\{\boldsymbol{\mathcal{S}}^{(i)}\right\}_{i=1}^{M}$ of states from the probability distribution $p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})$ is an appropriate way to estimate expectation values of operators, as follows:

$$\left\langle \hat{O} \right\rangle \approx \left\langle \hat{O} \right\rangle_{\mathrm{MC}} = \frac{1}{M} \sum_{i=1}^{M} O_{\mathrm{loc}} \left(\boldsymbol{\mathcal{S}}^{(i)};\boldsymbol{\mathcal{W}} \right), \tag{3.17}$$

where M is the number of samples in the Markov chain. This estimate $\left\langle \hat{O} \right\rangle_{\mathrm{MC}}$ unbiasedly approaches the exact result of $\left\langle \hat{O} \right\rangle$ in Eq. (3.16) when M goes to infinity. However, it is not possible to choose infinity, so the choice of the number of samples must be sufficiently large but finite. Due to this estimation,

$$\mathrm{var}\left(\sum_{i=1}^{M} O_{\mathrm{loc}} \left(\boldsymbol{\mathcal{S}}^{(i)};\boldsymbol{\mathcal{W}} \right) \right) = M \left\langle O_{\mathrm{loc}}^2 \right\rangle, \tag{3.18}$$

and the resulting variance is

$$\mathrm{var}\left(\hat{O}\right) = \frac{1}{M^2} \mathrm{var} \left(\sum_{i=1}^{M} O_{\mathrm{loc}}\left(\boldsymbol{\mathcal{S}}^{(i)};\boldsymbol{\mathcal{W}} \right) \right) = \frac{1}{M^2} M \left\langle O_{\mathrm{loc}}^2 \right\rangle = \frac{\langle O_{\mathrm{loc}}^2 \rangle}{M} \ . \tag{3.19}$$

The standard error $\Delta \left\langle \hat{O} \right\rangle = \sqrt{\mathrm{var}\left(\hat{O}\right)} = \sqrt{\langle O_{\mathrm{loc}}^2 \rangle / M} \sim 1/\sqrt{M}$ in this estimate of expectation values declines as M increases.

3.3 Sampling

Among MCMC methods, the Metropolis-Hastings algorithm [78] is one of the appropriate and recognized MCMC methods for sampling configurations from the probability distribution for NQS. The algorithm aims to generate a randomized set of states, *i.e.* random samples based on a target probability distribution by proposing a Markov chain (process) with transition probabilities.

It can be derived by using the principle of detailed balance so that each configuration transition, for example, $\boldsymbol{\mathcal{S}} \to \boldsymbol{\mathcal{S}}'$, is reversible for any $\boldsymbol{\mathcal{S}}$ and $\boldsymbol{\mathcal{S}}'$, which proposes that the probability of being in the configuration $\boldsymbol{\mathcal{S}}$ and changing from $\boldsymbol{\mathcal{S}}$ to $\boldsymbol{\mathcal{S}}'$ must be equal to the probability of being in the configuration $\boldsymbol{\mathcal{S}}'$ and changing from $\boldsymbol{\mathcal{S}}'$ to $\boldsymbol{\mathcal{S}}$, and it can be denoted as follows:

$$p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})\, \mathcal{T}_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}} \to \boldsymbol{\mathcal{S}}') = p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}')\, \mathcal{T}_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}' \to \boldsymbol{\mathcal{S}}), \tag{3.20}$$

where $\mathcal{T}_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}} \to \boldsymbol{\mathcal{S}}')$ is the transition probability from $\boldsymbol{\mathcal{S}}$ to $\boldsymbol{\mathcal{S}}'$.

Eq. (3.20) can be rewritten as the following ratios

$$\frac{p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}')}{p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})} = \frac{\mathcal{T}_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}} \to \boldsymbol{\mathcal{S}}')}{\mathcal{T}_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}' \to \boldsymbol{\mathcal{S}})}\,. \tag{3.21}$$

Thanks to the utilization of this algorithm, the transition probability $\mathcal{T}_{\boldsymbol{\mathcal{W}}}$ can be redefined as a product of two distinct probabilities

$$\mathcal{T}_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}} \to \boldsymbol{\mathcal{S}}') = T(\boldsymbol{\mathcal{S}} \to \boldsymbol{\mathcal{S}}')\, A_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}} \to \boldsymbol{\mathcal{S}}'), \tag{3.22}$$

where the local transition kernel T is the conditional probability of proposing the configuration $\boldsymbol{\mathcal{S}}'$ given $\boldsymbol{\mathcal{S}}$, and the acceptance probability $A_{\boldsymbol{\mathcal{W}}}$ is the probability to accept the proposed configuration $\boldsymbol{\mathcal{S}}'$.

By inserting the redefined transition probability in Eq. (3.22) and after some reorganization, the detailed balance in Eq. (3.21) can written as

$$r_A(\boldsymbol{\mathcal{S}} \rightleftarrows \boldsymbol{\mathcal{S}}'; \boldsymbol{\mathcal{W}}) = r_p(\boldsymbol{\mathcal{S}}', \boldsymbol{\mathcal{S}}; \boldsymbol{\mathcal{W}})\, r_T(\boldsymbol{\mathcal{S}} \leftrightarrows \boldsymbol{\mathcal{S}}'), \tag{3.23}$$

where

$$r_p(\boldsymbol{\mathcal{S}}', \boldsymbol{\mathcal{S}}; \boldsymbol{\mathcal{W}}) \coloneqq \frac{p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}')}{p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})} = \left| \frac{\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}')}{\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})} \right|^2, \tag{3.24}$$

which is the ratio of the probability $p_{\boldsymbol{\mathcal{W}}}$,

$$r_A(\boldsymbol{\mathcal{S}} \rightleftarrows \boldsymbol{\mathcal{S}}'; \boldsymbol{\mathcal{W}}) \coloneqq \frac{A_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}} \to \boldsymbol{\mathcal{S}}')}{A_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}' \to \boldsymbol{\mathcal{S}})}, \tag{3.25}$$

which is the corrected ratio of the acceptance probability, and

$$r_T(\boldsymbol{\mathcal{S}} \leftrightarrows \boldsymbol{\mathcal{S}}') \coloneqq \frac{T(\boldsymbol{\mathcal{S}}' \to \boldsymbol{\mathcal{S}})}{T(\boldsymbol{\mathcal{S}} \to \boldsymbol{\mathcal{S}}')}, \tag{3.26}$$

which is the correcting ratio of the local transition kernel T in both transition directions ($\boldsymbol{\mathcal{S}} \leftrightarrows \boldsymbol{\mathcal{S}}'$).

One should notice that the computational cost of calculating r_T is generally high due to the following choice of the transition kernel

$$T(\boldsymbol{\mathcal{S}} \to \boldsymbol{\mathcal{S}}') = \left(1 - \delta_{\boldsymbol{\mathcal{S}},\boldsymbol{\mathcal{S}}'}\right) \frac{\left| \langle \boldsymbol{\mathcal{S}} | \hat{\mathcal{H}} | \boldsymbol{\mathcal{S}}' \rangle \right|}{\sum_{\{\boldsymbol{\mathcal{S}}'' \,|\, \boldsymbol{\mathcal{S}}'' \neq \boldsymbol{\mathcal{S}}\}} \left| \langle \boldsymbol{\mathcal{S}} | \hat{\mathcal{H}} | \boldsymbol{\mathcal{S}}'' \rangle \right|} \tag{3.27}$$

which is called Hamiltonian transition rule [33, 79, 80]. However, since boson tunneling in the canonical ensemble or boson annihilation/creation in the grand-canonical ensemble are mostly chosen as a transition rule to propose a new bosonic configuration, the ratio of the correcting ratio r_T of the local transition kernel T becomes one, so the corrected ratio r_A of the acceptance probability can be simplified, as follows:

$$r_A(\boldsymbol{\mathcal{S}} \rightleftarrows \boldsymbol{\mathcal{S}}'; \boldsymbol{\mathcal{W}}) = r_p(\boldsymbol{\mathcal{S}}', \boldsymbol{\mathcal{S}}; \boldsymbol{\mathcal{W}}) \ . \tag{3.28}$$

After finalizing the definition of the ratio r_A depending on a chosen transition rule, in the ith step of the sampling process, a candidate configuration $\boldsymbol{\mathcal{S}}'$ is proposed according to the local transition kernel $T(\boldsymbol{\mathcal{S}} \to \boldsymbol{\mathcal{S}}')$ where $\boldsymbol{\mathcal{S}}^{(i)} = \boldsymbol{\mathcal{S}}$.

If $r_A(\boldsymbol{\mathcal{S}} \rightleftarrows \boldsymbol{\mathcal{S}}'; \boldsymbol{\mathcal{W}}) \geq 1$, the next configuration $\boldsymbol{\mathcal{S}}^{(i+1)}$ in the Markov chain is the candidate configuration $\boldsymbol{\mathcal{S}}'$. If r_A is in the interval $[0, 1)$, $\boldsymbol{\mathcal{S}}'$ is accepted as the next configuration $\boldsymbol{\mathcal{S}}^{(i+1)}$ with a probability r_A or rejected with a probability $1 - r_A$.

Obviously, the probability for the candidate configuration $\boldsymbol{\mathcal{S}}'$ to be accepted can defined as

$$\alpha_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}} \to \boldsymbol{\mathcal{S}}') \coloneqq \min[1, r_A(\boldsymbol{\mathcal{S}} \rightleftarrows \boldsymbol{\mathcal{S}}'; \boldsymbol{\mathcal{W}})] \ . \tag{3.29}$$

At each optimization iteration, a thermalization (or warm-up or burn-in) run is mostly necessary before generating the actual samples and carrying out measurements with them because the probability distribution of the variational wave function does not represent the actual stationary distribution at the first iterations of the sampling process.

The algorithm can thus be summarized as in Algorithm 2.

Algorithm 2 Metropolis-Hastings algorithm

Require: $M > 0$
1: randomly initialize a configuration $\boldsymbol{\mathcal{S}}^{(1)}$
2: **for** $i = 1$ **to** M **do**
3: propose a candidate configuration $\boldsymbol{\mathcal{S}}'$ according to $T\left(\boldsymbol{\mathcal{S}}^{(i)} \to \boldsymbol{\mathcal{S}}'\right)$ under a chosen transition rule
4: $\alpha_{\boldsymbol{\mathcal{W}}}\left(\boldsymbol{\mathcal{S}}^{(i)} \to \boldsymbol{\mathcal{S}}'\right) \leftarrow \min\left[1, r_A\left(\boldsymbol{\mathcal{S}}^{(i)} \rightleftarrows \boldsymbol{\mathcal{S}}'; \boldsymbol{\mathcal{W}}\right)\right]$
5: $r \sim U[0, 1)$
6: **if** $r \leq \alpha_{\boldsymbol{\mathcal{W}}}\left(\boldsymbol{\mathcal{S}}^{(i)} \to \boldsymbol{\mathcal{S}}'\right)$ **then**
7: $\boldsymbol{\mathcal{S}}^{(i+1)} \leftarrow \boldsymbol{\mathcal{S}}'$
8: **else**
9: $\boldsymbol{\mathcal{S}}^{(i+1)} \leftarrow \boldsymbol{\mathcal{S}}$
10: **end if**
11: **end for**

3.4 Optimization

To find out the ground-state wave functions, the variational network parameters require an optimization process. For this, a gradient descent algorithm is utilized usually, as follows:

$$\mathcal{W}_k \to \mathcal{W}_k - \eta \, C_k(\boldsymbol{\mathcal{W}}), \tag{3.30}$$

where the learning rate η is a configurable hyperparameter used to determine the step size at each optimization step of the gradient descent algorithm, and $C_k(\boldsymbol{\mathcal{W}})$ is the derivative of the cost function with respect to kth variational network parameter. Here, the cost function is the variational energy defined below

$$\left\langle \hat{\mathcal{H}} \right\rangle = \langle E_{\text{loc}} \rangle = \frac{\langle \Psi_{\boldsymbol{\mathcal{W}}} | \hat{\mathcal{H}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle}{\langle \Psi_{\boldsymbol{\mathcal{W}}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle} \geq E_{\text{g}}, \tag{3.31}$$

where E_{loc} is the local energy, *i.e.* the local estimator of the Hamiltonian, as follows:

$$E_{\text{loc}}(\boldsymbol{\mathcal{S}}; \boldsymbol{\mathcal{W}}) \coloneqq \sum_{\{\boldsymbol{\mathcal{S}}'\}} \langle \boldsymbol{\mathcal{S}} | \hat{\mathcal{H}} | \boldsymbol{\mathcal{S}}' \rangle \, \frac{\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}')}{\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})} = \frac{\langle \boldsymbol{\mathcal{S}} | \hat{\mathcal{H}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle}{\langle \boldsymbol{\mathcal{S}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle} \, . \tag{3.32}$$

With MCMC, the variational energy is approximated as shown below

$$E(\boldsymbol{\mathcal{W}}) \coloneqq \left\langle \hat{\mathcal{H}} \right\rangle_{\text{MC}} = \frac{1}{M} \sum_{i=1}^{M} E_{\text{loc}}\left(\boldsymbol{\mathcal{S}}^{(i)}; \boldsymbol{\mathcal{W}}\right) \, . \tag{3.33}$$

Each optimization step needs the differentiation of the variational energy with respect to each variational network parameter, *i.e.* the energy gradients g_k. However, calculating the energy gradients varies depending on the parametrization of the variational wave function. For convenience, the subscript MC is dropped in the rest of this thesis even if MCMC is used to calculate an expectation value.

3.4.1 Energy gradients

If the real-valued variational wave function, *i.e.* $\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \in \mathbb{R}$, is parameterized by real-valued network parameters, *i.e.* $\mathcal{W}_k \in \mathbb{R}$, the energy gradient is

$$\begin{aligned} g_k(\boldsymbol{\mathcal{W}}) \coloneqq \frac{\partial}{\partial \mathcal{W}_k} \left\langle \hat{\mathcal{H}} \right\rangle &= \frac{\partial}{\partial \mathcal{W}_k} \left(\frac{\langle \Psi_{\boldsymbol{\mathcal{W}}} | \hat{\mathcal{H}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle}{\langle \Psi_{\boldsymbol{\mathcal{W}}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle} \right) & (3.34) \\ &= 2 \frac{\sum_{\{\boldsymbol{\mathcal{S}}\},\{\boldsymbol{\mathcal{S}}'\}} \left[\langle \boldsymbol{\mathcal{S}} | \hat{\mathcal{H}} | \boldsymbol{\mathcal{S}}' \rangle \, \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}') \, \partial_{\mathcal{W}_k} \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \right]}{\sum_{\{\boldsymbol{\mathcal{S}}''\}} \Psi^2_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}'')} & (3.35) \\ &\quad - 2 \, \langle \Psi_{\boldsymbol{\mathcal{W}}} | \hat{\mathcal{H}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle \, \frac{\sum_{\{\boldsymbol{\mathcal{S}}\}} \left[\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \, \partial_{\mathcal{W}_k} \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \right]}{\left[\sum_{\{\boldsymbol{\mathcal{S}}''\}}^{M} \Psi^2_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}'') \right]^2} \, . \end{aligned}$$

By introducing the derivative of the logarithm of the wave function with respect to the network parameters

$$O_k(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}}) \coloneqq \frac{\partial}{\partial \mathcal{W}_k} \ln |\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})| = \frac{1}{\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})} \frac{\partial \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})}{\partial \mathcal{W}_k}, \tag{3.36}$$

which can be rewritten in terms of the an operator as follows:

$$O_k(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}}) = \langle \boldsymbol{\mathcal{S}} | \hat{\mathcal{O}}_k | \boldsymbol{\mathcal{S}} \rangle \,, \tag{3.37}$$

where $\hat{\mathcal{O}}_k \coloneqq \partial_{\mathcal{W}_k}$ because

$$\begin{aligned}
\hat{\mathcal{O}}_k \, |\Psi_{\boldsymbol{\mathcal{W}}}\rangle &= \frac{\partial}{\partial \mathcal{W}_k} |\Psi_{\boldsymbol{\mathcal{W}}}\rangle && (3.38)\\
&= \sum_{\{\boldsymbol{\mathcal{S}}\}} |\boldsymbol{\mathcal{S}}\rangle \frac{\partial}{\partial \mathcal{W}_k} \langle \boldsymbol{\mathcal{S}} | \Psi_{\boldsymbol{\mathcal{W}}}\rangle && (3.39)\\
&= \sum_{\{\boldsymbol{\mathcal{S}}\}} |\boldsymbol{\mathcal{S}}\rangle \, \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \, O_k(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}}) && (3.40)\\
&= \sum_{\{\boldsymbol{\mathcal{S}}\}} |\boldsymbol{\mathcal{S}}\rangle \, \langle \boldsymbol{\mathcal{S}} | \Psi_{\boldsymbol{\mathcal{W}}}\rangle \, O_k(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}}) && (3.41)\\
&= \left[\sum_{\{\boldsymbol{\mathcal{S}}\}} |\boldsymbol{\mathcal{S}}\rangle \, O_k(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}}) \, \langle \boldsymbol{\mathcal{S}}| \right] |\Psi_{\boldsymbol{\mathcal{W}}}\rangle \,, && (3.42)
\end{aligned}$$

and by substituting the local energy defined in Eq. (3.32), the energy gradient becomes

$$\begin{aligned}
g_k(\boldsymbol{\mathcal{W}}) &= 2 \sum_{\{\boldsymbol{\mathcal{S}}\}} p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \, E_{\text{loc}}(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}}) \, O_k(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}}) && (3.43)\\
&\quad -2 \langle E_{\text{loc}} \rangle \sum_{\{\boldsymbol{\mathcal{S}}\}} p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \, O_k(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}})\\
&= 2 \langle E_{\text{loc}}; O_k \rangle \,, && (3.44)
\end{aligned}$$

where $\langle A; B \rangle = \langle AB \rangle - \langle A \rangle \langle B \rangle$ is the correlation function for A and B.

If the complex-valued variational wave function, *i.e.* $\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \in \mathbb{C}$, is parameterized by real-valued network parameters, *i.e.* $\mathcal{W}_k \in \mathbb{R}$, the energy gradient is similar to the result in Eq. (3.44), but now $O_k \neq O_k^*$ in general, so the energy

gradient is defined as

$$
\begin{aligned}
g_k(\boldsymbol{\mathcal{W}}) \coloneqq \frac{\partial}{\partial \mathcal{W}_k} \left\langle \hat{\mathcal{H}} \right\rangle &= \frac{\partial}{\partial \mathcal{W}_k} \left(\frac{\langle \Psi_{\boldsymbol{\mathcal{W}}} | \hat{\mathcal{H}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle}{\langle \Psi_{\boldsymbol{\mathcal{W}}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle} \right) && (3.45) \\
&= \frac{\sum_{\{\boldsymbol{\mathcal{S}}\},\{\boldsymbol{\mathcal{S}}'\}} \left[\langle \boldsymbol{\mathcal{S}} | \hat{\mathcal{H}} | \boldsymbol{\mathcal{S}}' \rangle \, \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}') \, \partial_{\mathcal{W}_k} \Psi^*_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \right]}{\sum_{\{\boldsymbol{\mathcal{S}}''\}} |\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}'')|^2} && (3.46) \\
&\quad + \frac{\sum_{\{\boldsymbol{\mathcal{S}}\},\{\boldsymbol{\mathcal{S}}'\}} \left[\Psi^*_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}') \, \langle \boldsymbol{\mathcal{S}}' | \hat{\mathcal{H}} | \boldsymbol{\mathcal{S}} \rangle \, \partial_{\mathcal{W}_k} \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \right]}{\sum_{\{\boldsymbol{\mathcal{S}}''\}} |\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}'')|^2} \\
&\quad - \langle \Psi_{\boldsymbol{\mathcal{W}}} | \hat{\mathcal{H}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle \, \frac{\sum_{\{\boldsymbol{\mathcal{S}}\}} \left[\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \, \partial_{\mathcal{W}_k} \Psi^*_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \right]}{\left[\sum_{\{\boldsymbol{\mathcal{S}}''\}}^{M} |\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}'')|^2 \right]^2} \\
&\quad - \langle \Psi_{\boldsymbol{\mathcal{W}}} | \hat{\mathcal{H}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle \, \frac{\sum_{\{\boldsymbol{\mathcal{S}}\}} \left[\Psi^*_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \, \partial_{\mathcal{W}_k} \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \right]}{\left[\sum_{\{\boldsymbol{\mathcal{S}}''\}} |\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}'')|^2 \right]^2} \\
&= \sum_{\{\boldsymbol{\mathcal{S}}\}} p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \, E_{\mathrm{loc}}(\boldsymbol{\mathcal{S}}; \boldsymbol{\mathcal{W}}) \, O^*_k(\boldsymbol{\mathcal{S}}; \boldsymbol{\mathcal{W}}) && (3.47) \\
&\quad - \langle E_{\mathrm{loc}} \rangle \sum_{\{\boldsymbol{\mathcal{S}}\}} p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \, O^*_k(\boldsymbol{\mathcal{S}}; \boldsymbol{\mathcal{W}}) + \text{c.c.} \\
&= \langle E_{\mathrm{loc}}; O^*_k \rangle + \text{c.c.} && (3.48) \\
&= 2 \, \mathrm{Re}\{ \langle E_{\mathrm{loc}}; O^*_k \rangle \}, && (3.49)
\end{aligned}
$$

which can be rewritten as

$$
g_k(\boldsymbol{\mathcal{W}}) = 2 \, \mathrm{Re}\{ f_k(\boldsymbol{\mathcal{W}}) \}, \tag{3.50}
$$

by defining so-called energy forces as

$$
f_k(\boldsymbol{\mathcal{W}}) \coloneqq \langle E_{\mathrm{loc}}; O^*_k \rangle \ . \tag{3.51}
$$

If the complex-valued variational wave function, *i.e.* $\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \in \mathbb{C}$, is parameterized by complex-valued network parameters, *i.e.* $\mathcal{W}_k = \mathcal{R}_k + i\mathcal{I}_k \in \mathbb{C}$ where $\mathcal{R}_k, \mathcal{I}_k \in \mathbb{R}$, the calculation of the gradients differs by the holomorphicity of the wave function.

The holomorphicity of a function $f(\mathcal{W}) = u(\mathcal{W}) + i \, v(\mathcal{W}) \in \mathbb{C}$ where $\mathcal{W} = \mathcal{R} + i\mathcal{I} \in \mathbb{C}$ and $\mathcal{R}, \mathcal{I}, u, v \in \mathbb{R}$ at a point w_0 can be determined by checking

whether it is complex differentiable at $w_0 \in \mathbb{C}$ and satisfies the Cauchy-Riemann equations at w_0, as defined below

$$\left.\frac{\partial u}{\partial \mathcal{R}}\right|_{w_0} = \left.\frac{\partial v}{\partial \mathcal{I}}\right|_{w_0} \tag{3.52}$$

$$\left.\frac{\partial u}{\partial \mathcal{I}}\right|_{w_0} = -\left.\frac{\partial v}{\partial \mathcal{R}}\right|_{w_0}, \tag{3.53}$$

or, equivalently, by checking whether its derivative with respect to the complex-conjugated parameter $\mathcal{W}^*$ at w_0 turns out to be zero:

$$\left.\frac{\partial f}{\partial \mathcal{W}^*}\right|_{w_0} = 0, \tag{3.54}$$

with the following definition of Wirtinger derivatives [81], shown below

$$\frac{\partial}{\partial \mathcal{W}} = \frac{1}{2}\left(\frac{\partial}{\partial \mathcal{R}} - i\frac{\partial}{\partial \mathcal{I}}\right) \tag{3.55}$$

$$\frac{\partial}{\partial \mathcal{W}^*} = \frac{1}{2}\left(\frac{\partial}{\partial \mathcal{R}} + i\frac{\partial}{\partial \mathcal{I}}\right). \tag{3.56}$$

For holomorphic wave functions, the energy gradient can be determined by relating the complex-valued parameters and their real and imaginary parts with a transformation matrix $\boldsymbol{J}$ as follows:

$$\boldsymbol{c}_k \coloneqq \begin{pmatrix} \mathcal{W}_k \\ \mathcal{W}_k^* \end{pmatrix} = \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} \mathcal{R}_k \\ \mathcal{I}_k \end{pmatrix} \coloneqq \boldsymbol{J}\,\boldsymbol{r}_k, \tag{3.57}$$

where the differentiation operators with respect to the vectors $\boldsymbol{c}_k$ and $\boldsymbol{r}_k$ can be defined as

$$\frac{\partial}{\partial \boldsymbol{c}_k} \coloneqq \begin{pmatrix} \partial_{\mathcal{W}_k} \\ \partial_{\mathcal{W}_k^*} \end{pmatrix} \tag{3.58}$$

$$\frac{\partial}{\partial \boldsymbol{r}_k} \coloneqq \begin{pmatrix} \partial_{\mathcal{R}_k} \\ \partial_{\mathcal{I}_k} \end{pmatrix}. \tag{3.59}$$

These differentiation operators can also be related to each other by the following equation

$$\frac{\partial}{\partial \boldsymbol{r}_k} = \frac{\partial \boldsymbol{c}_k^\top}{\partial \boldsymbol{r}_k}\frac{\partial}{\partial \boldsymbol{c}_k} = \boldsymbol{J}^\top \frac{\partial}{\partial \boldsymbol{c}_k}, \tag{3.60}$$

which also implies

$$\left(\frac{\partial}{\partial \boldsymbol{r}_k}\right)^* = \frac{\partial}{\partial \boldsymbol{r}_k} = \left(\boldsymbol{J}^\top \frac{\partial}{\partial \boldsymbol{c}_k}\right)^* = \boldsymbol{J}^\dagger \frac{\partial}{\partial \boldsymbol{c}_k^*} \,. \tag{3.61}$$

The minimization of the variational energy can done by applying $\delta \boldsymbol{r}_k \coloneqq \partial_{\boldsymbol{r}_k} \left\langle \hat{\mathcal{H}} \right\rangle$ at each optimization step in the parameter space $\mathbb{R}^{2p}$. To carry out the same update in the parameter space $\mathbb{C}^p$, the transformation in Eq. (3.57) can be used, as follows:

$$\delta \boldsymbol{c}_k \coloneqq \boldsymbol{J}\, \delta \boldsymbol{r}_k = \boldsymbol{J}\boldsymbol{J}^\dagger \frac{\partial}{\partial \boldsymbol{c}^*} \left\langle \hat{\mathcal{H}} \right\rangle = 2 \frac{\partial}{\partial \boldsymbol{c}^*} \left\langle \hat{\mathcal{H}} \right\rangle , \tag{3.62}$$

which leads to

$$g_k(\boldsymbol{\mathcal{W}}) \coloneqq 2 \frac{\partial}{\partial \mathcal{W}_k^*} \left\langle \hat{\mathcal{H}} \right\rangle , \tag{3.63}$$

where the factor 2 is usually ignored since it can be compensated by tuning the learning rate η.

In regard of Eq. (3.63), the energy gradient is

$$g_k(\boldsymbol{\mathcal{W}}) \coloneqq \frac{\partial}{\partial \mathcal{W}_k^*} \left\langle \hat{\mathcal{H}} \right\rangle = \frac{\partial}{\partial \mathcal{W}_k^*} \left(\frac{\langle \Psi_{\boldsymbol{\mathcal{W}}} | \hat{\mathcal{H}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle}{\langle \Psi_{\boldsymbol{\mathcal{W}}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle} \right) \tag{3.64}$$

$$= \frac{\sum_{\{\boldsymbol{\mathcal{S}}\},\{\boldsymbol{\mathcal{S}}'\}} \left[\langle \boldsymbol{\mathcal{S}} | \hat{\mathcal{H}} | \boldsymbol{\mathcal{S}}' \rangle \, \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}') \, \partial_{\mathcal{W}_k^*} \Psi_{\boldsymbol{\mathcal{W}}}^*(\boldsymbol{\mathcal{S}}) \right]}{\sum_{\{\boldsymbol{\mathcal{S}}''\}} |\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}'')|^2} \tag{3.65}$$

$$+ \frac{\sum_{\{\boldsymbol{\mathcal{S}}\},\{\boldsymbol{\mathcal{S}}'\}} \left[\Psi_{\boldsymbol{\mathcal{W}}}^*(\boldsymbol{\mathcal{S}}') \, \langle \boldsymbol{\mathcal{S}}' | \hat{\mathcal{H}} | \boldsymbol{\mathcal{S}} \rangle \, \partial_{\mathcal{W}_k^*} \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \right]}{\sum_{\{\boldsymbol{\mathcal{S}}''\}} |\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}'')|^2}$$

$$- \langle \Psi_{\boldsymbol{\mathcal{W}}} | \hat{\mathcal{H}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle \frac{\sum_{\{\boldsymbol{\mathcal{S}}\}} \left[\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \, \partial_{\mathcal{W}_k^*} \Psi_{\boldsymbol{\mathcal{W}}}^*(\boldsymbol{\mathcal{S}}) \right]}{\left[\sum_{\{\boldsymbol{\mathcal{S}}''\}}^{M} |\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}'')|^2 \right]^2}$$

$$- \langle \Psi_{\boldsymbol{\mathcal{W}}} | \hat{\mathcal{H}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle \frac{\sum_{\{\boldsymbol{\mathcal{S}}\}} \left[\Psi_{\boldsymbol{\mathcal{W}}}^*(\boldsymbol{\mathcal{S}}) \, \partial_{\mathcal{W}_k^*} \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) \right]}{\left[\sum_{\{\boldsymbol{\mathcal{S}}''\}} |\Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}'')|^2 \right]^2} \,. \tag{3.66}$$

Since $\partial_{\mathcal{W}_k^*} \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}}) = \partial_{\mathcal{W}_k} \Psi_{\boldsymbol{\mathcal{W}}}^*(\boldsymbol{\mathcal{S}}) = 0$ for holomorphic wave functions as the

Cauchy-Riemann equations suggest in Eq. (3.54), the energy gradient becomes

$$
\begin{aligned}
g_k(\boldsymbol{\mathcal{W}}) &= \sum_{\{\boldsymbol{\mathcal{S}}\}} p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})\, E_{\mathrm{loc}}(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}})\, O_k^*(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}}) && (3.67)\\
&\quad - \langle E_{\mathrm{loc}}\rangle \sum_{\{\boldsymbol{\mathcal{S}}\}} p_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})\, O_k^*(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}})\\
&= \langle E_{\mathrm{loc}}; O_k^*\rangle && (3.68)\\
&= f_k(\boldsymbol{\mathcal{W}})\ . && (3.69)
\end{aligned}
$$

For a non-holomorphic wave function, another approach for differentiation is used since the Cauchy-Riemann equations do not hold. The real and imaginary parts of the complex-valued parameters are treated separately, which leads to $\boldsymbol{\mathcal{W}}_{\cup} \coloneqq \boldsymbol{\mathcal{R}} \cup \boldsymbol{\mathcal{I}} \in \mathbb{R}^{2p}$ rather than $\boldsymbol{\mathcal{W}} = \boldsymbol{\mathcal{R}} + i\boldsymbol{\mathcal{I}} \in \mathbb{C}^p$, and rewrite the function in terms of these combined real-valued network parameters $\boldsymbol{\mathcal{W}}_{\cup}$ so that $\Psi_{\boldsymbol{\mathcal{W}}_{\cup}}(\boldsymbol{\mathcal{S}}) = \Psi_{\boldsymbol{\mathcal{W}}}(\boldsymbol{\mathcal{S}})$ [32, 79]. With this choice, the energy derivates result as in Eq. (3.50).

3.4.2 Zero-variance property

Throughout the optimization, the variational energy is expected to converge a minimum value, *i.e.* the ground-state energy that can be unknown for some many-body systems. To overcome this indefiniteness, the variational energy variance can be chosen as an auxiliary convergence measure.

Ideally, if sampling is carried out from the probability distribution $p_{\boldsymbol{\mathcal{W}}_0}(\boldsymbol{\mathcal{S}})$ for the ground state $|\Psi_{\boldsymbol{\mathcal{W}}_0}\rangle$ where $\boldsymbol{\mathcal{W}}_0$ is the exact variational network parameters for the ground-state, this state is then an eigenstate of the Hamiltonian $\hat{\mathcal{H}}$ with its eigenvalue $E(\boldsymbol{\mathcal{W}}_0) = E_{\mathrm{g}}$, so the local energy becomes

$$
\begin{aligned}
E_{\mathrm{loc}}(\boldsymbol{\mathcal{S}};\boldsymbol{\mathcal{W}}_0) &= \frac{\langle \boldsymbol{\mathcal{S}}|\hat{\mathcal{H}}|\Psi_{\boldsymbol{\mathcal{W}}_0}\rangle}{\langle \boldsymbol{\mathcal{S}}|\Psi_{\boldsymbol{\mathcal{W}}_0}\rangle} && (3.70)\\
&= \frac{E(\boldsymbol{\mathcal{W}}_0)\,\langle \boldsymbol{\mathcal{S}}|\Psi_{\boldsymbol{\mathcal{W}}_0}\rangle}{\langle \boldsymbol{\mathcal{S}}|\Psi_{\boldsymbol{\mathcal{W}}_0}\rangle} && (3.71)\\
&= E_{\mathrm{g}}, && (3.72)
\end{aligned}
$$

which exactly equals the ground-state energy. Thus, the variational energy variance for the ground state turns out to be

$$\begin{aligned} \mathrm{var}(\hat{\mathcal{H}}) &= \left\langle \hat{\mathcal{H}}^2 \right\rangle - \left\langle \hat{\mathcal{H}} \right\rangle^2 && (3.73) \\ &= \frac{\langle \Psi_{\boldsymbol{\mathcal{W}}_0} | \hat{\mathcal{H}}^2 | \Psi_{\boldsymbol{\mathcal{W}}_0} \rangle}{\langle \Psi_{\boldsymbol{\mathcal{W}}_0} | \Psi_{\boldsymbol{\mathcal{W}}_0} \rangle} - \left(\frac{\langle \Psi_{\boldsymbol{\mathcal{W}}_0} | \hat{\mathcal{H}} | \Psi_{\boldsymbol{\mathcal{W}}_0} \rangle}{\langle \Psi_{\boldsymbol{\mathcal{W}}_0} | \Psi_{\boldsymbol{\mathcal{W}}_0} \rangle} \right)^2 && (3.74) \\ &= E_g^2 - E_g^2 = 0, && (3.75) \end{aligned}$$

which indicates so-called *zero-variance property*.

Whereas, for non-eigenstates, this variance is not zero, but a positive number because the expectation value $\left\langle \hat{\mathcal{H}}^2 \right\rangle$ of the Hamiltonian squared $\hat{\mathcal{H}}^2$ results as

$$\left\langle \hat{\mathcal{H}}^2 \right\rangle = \frac{\langle \Psi_{\boldsymbol{\mathcal{W}}} | \hat{\mathcal{H}}^2 | \Psi_{\boldsymbol{\mathcal{W}}} \rangle}{\langle \Psi_{\boldsymbol{\mathcal{W}}} | \Psi_{\boldsymbol{\mathcal{W}}} \rangle} = \left\langle E_{\mathrm{loc}}^2 \right\rangle , \tag{3.76}$$

which is definitely not equal to $\langle E_{\mathrm{loc}} \rangle^2 = [E(\boldsymbol{\mathcal{W}})]^2$. Thus, in general, the variational energy variance can be written as follows:

$$\mathrm{var}(\hat{\mathcal{H}}) = \left\langle E_{\mathrm{loc}}^2 \right\rangle - \langle E_{\mathrm{loc}} \rangle^2 \geq 0 \ . \tag{3.77}$$

The convergence to the ground state $|\Psi_{\boldsymbol{\mathcal{W}}_0}\rangle$ can also be done by differentiating the variance with respect to each network parameter $\mathcal{W}_k$ instead of calculating the energy gradient g_k to obtain the ground-state energy E_{g}. However, it is not guaranteed to converge to the ground state since the only state that satisfies the zero-variance property is not the ground state, but also all excited states. That is why utilizing the variational energy variance as a subsidiary measure is much more favored.

3.4.3 Stochastic gradient descent

Due to the exponentially increasing size of the Hilbert space, it is preferable to calculate and apply the energy gradient from stochastically sampled many-body configurations. For such a case, a straightforward way to minimize the variational energy is to utilize stochastic gradient descent (SGD) defined in Algorithm 3.

Algorithm 3 Stochastic gradient descent (SGD)

Require: $\eta > 0$
1: **repeat**
2: get M sampled many-body configurations $\{\boldsymbol{\mathcal{S}}_j\}_{j=1}^{M}$ with MCMC
3: **for** each $\mathcal{W}_k$ **do**
4: calculate $g_k(\boldsymbol{\mathcal{W}})$ with these configurations
5: $\mathcal{W}_k \leftarrow \mathcal{W}_k - \eta\, g_k(\boldsymbol{\mathcal{W}})$
6: **end for**
7: **until** $E(\boldsymbol{\mathcal{W}})$ is converged to a minimum energy value within an error range

This method does not guarantee to converge to the ground-state energy. It generally requires much more iterations than other optimization algorithms, even if it converges.

3.4.4 ADAM: Adaptive moment estimation

Without introducing further measurements, the adaptive moment estimation (ADAM) method makes it possible to develop and accelerate the energy optimization by computing adaptive learning rates instead of fixed rates for each variational network parameter as in SGD [82]. To do so, ADAM estimates the first two moments of gradients with exponentially decaying gradient averages.

At each iteration, the biased first-moment estimate m_k stores past energy derivates, as follows:

$$m_k^{(i)} \coloneqq \beta_1\, m_k^{(i-1)} + (1-\beta_1)\, g_k\Big(\boldsymbol{\mathcal{W}}^{(i-1)}\Big), \tag{3.78}$$

where β_1 is the exponential decay rate of the first-moment estimate. To make m_k biased corrected, it is divided by an exponentially increasing factor determined by both the number of iterations and β_1

$$\hat{m}_k^{(i)} \coloneqq \frac{m_k^{(i)}}{1-\beta_1^i}, \tag{3.79}$$

which corrects the discrepancy between $\langle m_k \rangle$ and its true moment $\langle g_k \rangle$.

Similarly, the biased second-moment estimate v_k is made up of past squared

energy gradients, like RMSProp [83], as shown below

$$v_k^{(i)} \coloneqq \beta_2\, v_k^{(i-1)} + (1-\beta_2)\left[g_k\Big(\boldsymbol{\mathcal{W}}^{(i-1)}\Big)\right]^2, \tag{3.80}$$

where β_2 is the exponential decay rate for the second-moment estimate, so its biased-corrected value is

$$\hat{v}_k^{(i)} \coloneqq \frac{v_k^{(i)}}{1-\beta_2^i}\ . \tag{3.81}$$

With these definitions, the optimization update can be written as

$$\mathcal{W}_k^{(i+1)} \coloneqq \mathcal{W}_k^{(i)} - \eta \frac{\hat{m}_k^{(i)}}{\sqrt{\hat{v}_k^{(i)} + \overline{\varepsilon}} + \varepsilon} g_k\Big(\boldsymbol{\mathcal{W}}^{(i)}\Big), \tag{3.82}$$

where $\overline{\varepsilon}$ and ε are the small constants in the denominator of the adaptive learning rate to avoid dividing by zero. For the initialization of ADAM, $m_k^{(0)}$ and $v_k^{(0)}$ is set to zero. All method is summarized in Algorithm 4.

Algorithm 4 Adaptive moment estimation (ADAM)

Require: $\eta, \varepsilon, \beta_1, \beta_2 > 0$, $\overline{\varepsilon} \geq 0$
$i \leftarrow 0$
$(m_k, v_k) \leftarrow (0, 0)$
repeat
  get M sampled many-body configurations $\{\boldsymbol{\mathcal{S}}_j\}_{j=1}^{M}$ with MCMC
  for each $\mathcal{W}_k$ **do**
    calculate $g_k(\boldsymbol{\mathcal{W}})$ with these configurations
    $m_k \leftarrow \beta_1\, m_k + (1-\beta_1)\, g_k(\boldsymbol{\mathcal{W}})$
    $v_k \leftarrow \beta_2\, v_k + (1-\beta_2)\, [g_k(\boldsymbol{\mathcal{W}})]^2$
    $\hat{m}_k \leftarrow m_k / (1-\beta_1^i)$
    $\hat{v}_k \leftarrow v_k / (1-\beta_2^i)$
    $i \leftarrow i+1$
    $\mathcal{W}_k \leftarrow \mathcal{W}_k - \eta\, g_k(\boldsymbol{\mathcal{W}})\, \hat{m}_k / \left(\sqrt{\hat{v}_k + \overline{\varepsilon}} + \varepsilon\right)$
  end for
until $E(\boldsymbol{\mathcal{W}})$ is converged to a minimum energy value within an error range

3.4.5 Stochastic reconfiguration

Canonical gradient descents as first-order optimization methods such as SGD and ADAM optimize variational parameters in the steepest-descent direction in

a Euclidean space, so the optimization update is carried out with the following distance (for convenience, the real and imaginary parts of the variational parameters are accepted as they are separated, so it is assumed that all the variational parameters are real valued)

$$\|\delta \boldsymbol{\mathcal{W}}\|^2 = \sum_k (\delta \mathcal{W}_k)^2, \tag{3.83}$$

which is the same for each parameter. This approach has some problems. It can require excessively many iterations for convergence to the ground-state energy that is only generally guaranteed. The reason is that such gradient descents are unaware of the non-Euclidean geometry of the variational network parameter space, so even an infinitesimally small change in these parameters could cause an utterly divergent variational energy.

A solution to this problem is to move on the parameter space under consideration of its geometry to find optimal variational parameters, which can be done by using the natural gradient descent [84]. It is a second-order optimization method that updates the variational parameters in the steepest-descent direction in a Riemannian space. For such a case, the distance turns out to be different from Eq. (3.83), as follows:

$$\|\boldsymbol{\delta \mathcal{W}}\|^2 = \sum_{k,k'} s_{kk'}(\boldsymbol{\mathcal{W}})\, \delta \mathcal{W}_{k'}\, \delta \mathcal{W}_k, \tag{3.84}$$

where $s_{kk'}$ is the quantum Fisher information metric. The distance of the canonical steepest descent in Eq. (3.83) can be recovered by setting $s_{kk'}(\boldsymbol{\mathcal{W}}) = \delta_{kk'}$.

With this metric, the natural gradient descent for the minimization of variational energy can be shown as

$$\mathcal{W}_k^{(i+1)} = \mathcal{W}_k^{(i)} - \sum_{k'} s_{kk'}^{-1}\left(\boldsymbol{\mathcal{W}}^{(i)}\right) f_{k'}\left(\boldsymbol{\mathcal{W}}^{(i)}\right), \tag{3.85}$$

which is also called the stochastic reconfiguration (SR) [30, 32, 33]. It can be derived by solving the imaginary time evolution of the variational wave function in the variational network parameter space, which gives $\boldsymbol{\delta \mathcal{W}}$ as a function of $\delta\tau$.

The imaginary time evolution of the variational wave function at a time τ by a time step $\delta\tau$ is carried out by

$$|\Psi_{\boldsymbol{\mathcal{W}}}(\tau+\delta\tau)\rangle = e^{-\delta\tau\hat{\mathcal{H}}}\,|\Psi_{\boldsymbol{\mathcal{W}}}(\tau)\rangle \approx \left(\hat{\mathbb{I}} - \delta\tau\hat{\mathcal{H}}\right)|\Psi_{\boldsymbol{\mathcal{W}}}(\tau)\rangle\,, \tag{3.86}$$

which can be written as the first-order Taylor expansion of the wave function since the time step $\delta\tau$ is infinitesimal. Here, the aim is to determine $\boldsymbol{\delta\mathcal{W}}$, so the wave function needs to be expanded around each $\delta\mathcal{W}_k$ in the first order, as follows (from now on, it is assumed that each network parameter is complex valued and the wave function is holomorphic):

$$\begin{aligned} |\Psi_{\boldsymbol{\mathcal{W}}+\delta\boldsymbol{\mathcal{W}}}(\tau)\rangle &= |\Psi_{\boldsymbol{\mathcal{W}}}(\tau)\rangle + \sum_k \left(\delta\mathcal{W}_k\frac{\partial}{\partial\mathcal{W}_k} + \delta\mathcal{W}_k^*\frac{\partial}{\partial\mathcal{W}_k^*}\right)|\Psi_{\boldsymbol{\mathcal{W}}}(\tau)\rangle \quad (3.87)\\ &= |\Psi_{\boldsymbol{\mathcal{W}}}(\tau)\rangle + \sum_k \delta\mathcal{W}_k\,\hat{\mathcal{O}}_k\,|\Psi_{\boldsymbol{\mathcal{W}}}(\tau)\rangle\,, \quad (3.88)\end{aligned}$$

which is only valid for a normalized wave function, whereas no NQS ansatz is normalized. Thus, it is redefined by introducing a zeroth-order factor $\delta\mathcal{W}_0$ to the Taylor series in Eq. (3.88) for normalization, which becomes

$$|\Psi_{\boldsymbol{\mathcal{W}}+\delta\boldsymbol{\mathcal{W}}}(\tau)\rangle = \delta\mathcal{W}_0\,|\Psi_{\boldsymbol{\mathcal{W}}}(\tau)\rangle + \sum_k \delta\mathcal{W}_k\,\hat{\mathcal{O}}_k\,|\Psi_{\boldsymbol{\mathcal{W}}}(\tau)\rangle\ . \tag{3.89}$$

By matching these two expansions in Eqs. (3.86) and (3.89), a new approximated equality arises

$$\delta\mathcal{W}_0\,|\Psi_{\boldsymbol{\mathcal{W}}}\rangle + \sum_{k\neq 0}\delta\mathcal{W}_k\,\hat{\mathcal{O}}_k\,|\Psi_{\boldsymbol{\mathcal{W}}}\rangle \approx \left(\hat{\mathbb{I}} - \delta\tau\hat{\mathcal{H}}\right)|\Psi_{\boldsymbol{\mathcal{W}}}\rangle\ . \tag{3.90}$$

To find the made-up normalization factor $\delta\mathcal{W}_0$, the state $\frac{\langle\Psi_{\boldsymbol{\mathcal{W}}}(\tau)|}{\langle\Psi_{\boldsymbol{\mathcal{W}}}(\tau)|\Psi_{\boldsymbol{\mathcal{W}}}(\tau)\rangle}$ and Eq. (3.90) are multiplied, which results as

$$\delta\mathcal{W}_0 + \sum_{k\neq 0}\delta\mathcal{W}_k\,\langle O_k\rangle = 1 - \delta\tau\,\langle E_{\mathrm{loc}}\rangle\,, \tag{3.91}$$

where $\delta\mathcal{W}_0$ can be substituted as follows

$$\delta\mathcal{W}_0 = 1 - \delta\tau\,\langle E_{\mathrm{loc}}\rangle - \sum_{k\neq 0}\delta\mathcal{W}_k\,\langle O_k\rangle\ . \tag{3.92}$$

By inserting Eq. (3.92), Eq. (3.90) becomes

$$\sum_{k\neq 0} \delta\mathcal{W}_k \left(\hat{\mathcal{O}}_k - \langle O_k \rangle\right) |\Psi_{\boldsymbol{\mathcal{W}}}\rangle = -\delta\tau \left(\hat{\mathcal{H}} - \langle E_{\text{loc}} \rangle\right) |\Psi_{\boldsymbol{\mathcal{W}}}\rangle \,, \tag{3.93}$$

which does not depend on $\delta\mathcal{W}_0$ anymore.

To obtain $\{\delta\mathcal{W}_k\}_{k\neq 0}$, the state $\frac{\langle\Psi_{\boldsymbol{\mathcal{W}}}(\tau)|\hat{\mathcal{O}}^*_{k'}}{\langle\Psi_{\boldsymbol{\mathcal{W}}}(\tau)|\Psi_{\boldsymbol{\mathcal{W}}}(\tau)\rangle}$ and Eq. (3.90) are multiplied, which gives

$$\sum_{k\neq 0} \delta\mathcal{W}_k \langle O^*_{k'}; O_k \rangle = -\delta\tau \langle E_{\text{loc}}; O^*_{k'} \rangle \ . \tag{3.94}$$

By setting the metric $s_{k'k} \coloneqq \langle O^*_{k'}; O_k \rangle$, Eq. (3.94) turns out to be

$$\sum_{k\neq 0} s_{k'k}(\boldsymbol{\mathcal{W}})\, \delta\mathcal{W}_k = -\delta\tau f_{k'}(\boldsymbol{\mathcal{W}}), \tag{3.95}$$

which can be rewritten as a matrix-vector multiplication

$$\boldsymbol{S}\, \delta\boldsymbol{\mathcal{W}} = -\delta\tau \boldsymbol{F}, \tag{3.96}$$

where $\boldsymbol{S} = \{s_{k'k}\}$ is the quantum Fisher matrix, $\boldsymbol{\delta\mathcal{W}}$ is the vector $\begin{pmatrix} \delta\mathcal{W}_1 & \cdots & \delta\mathcal{W}_p \end{pmatrix}^\top$ and $\boldsymbol{F} = \begin{pmatrix} f_1 & \cdots & f_p \end{pmatrix}^\top$ is the energy force vector. Thus, $\boldsymbol{\delta\mathcal{W}}$ as a function of the time step $\delta\tau$ can be substituted by multiplying Eq. (3.96) by the inverse of $\boldsymbol{S}$

$$\boldsymbol{\delta\mathcal{W}} = -\delta\tau \boldsymbol{S}^{-1} \boldsymbol{F} \ . \tag{3.97}$$

From now on, the infinitesimal time step $\delta\tau$ corresponds to the learning rate η in MCMC, so

$$\boldsymbol{\delta\mathcal{W}} = -\eta\, \boldsymbol{S}^{-1} \boldsymbol{F}, \tag{3.98}$$

or it can be rewritten explicitly

$$\delta\mathcal{W}_k = -\eta \sum_{k'\neq 0} s^{-1}_{kk'}(\boldsymbol{\mathcal{W}})\, f_{k'}(\boldsymbol{\mathcal{W}}) \ . \tag{3.99}$$

SR algorithm is summarized in Algorithm 5.

To look into how the variational energy is affected with such an update, its first-order Taylor expansion can be calculated around $\boldsymbol{\delta\mathcal{W}}$

$$E(\boldsymbol{\mathcal{W}} + \boldsymbol{\delta\mathcal{W}}) \approx E(\boldsymbol{\mathcal{W}}) + \sum_{k\neq 0} \left(\delta\mathcal{W}_k \frac{\partial E}{\partial \mathcal{W}_k} + \delta\mathcal{W}^*_k \frac{\partial E}{\partial \mathcal{W}^*_k} \right), \tag{3.100}$$

Algorithm 5 Stochastic reconfiguration (SR)

Require: $\eta > 0$

1: **repeat**
2: get M sampled many-body configurations $\{\boldsymbol{\mathcal{S}}_j\}_{j=1}^{M}$ with MCMC
3: calculate $\boldsymbol{F} = (\, f_1 \,\cdots\, f_P \,)^\top$ and $\boldsymbol{S} = \{\langle O_{k'}^*; O_k \rangle\}$ with these configurations
4: determine $\boldsymbol{S}^{-1}$ by solving $-\eta\, \boldsymbol{F} = \boldsymbol{S}\, \boldsymbol{\delta\mathcal{W}}$ for $\boldsymbol{\delta\mathcal{W}}$
5: $\boldsymbol{\mathcal{W}} \leftarrow \boldsymbol{\mathcal{W}} - \eta\, \boldsymbol{S}^{-1}\boldsymbol{F}$
6: **until** $E(\boldsymbol{\mathcal{W}})$ is converged to a minimum energy value within an error range

so the energy difference due to the update $\boldsymbol{\delta\mathcal{W}}$ is

$$\Delta E(\boldsymbol{\mathcal{W}}) \;\coloneqq\; E(\boldsymbol{\mathcal{W}} + \delta\boldsymbol{\mathcal{W}}) - E(\boldsymbol{\mathcal{W}}) \tag{3.101}$$

$$\Delta E(\boldsymbol{\mathcal{W}}) \;=\; -\eta \sum_{k,k' \neq 0} \left[s_{kk'}^{-1}(\boldsymbol{\mathcal{W}})\, f_{k'}(\boldsymbol{\mathcal{W}})\, f_k^*(\boldsymbol{\mathcal{W}}) + \text{c.c.} \right] \tag{3.102}$$

$$\;=\; -2\eta \sum_{k,k' \neq 0} \mathrm{Re}\big\{ s_{kk'}^{-1}(\boldsymbol{\mathcal{W}})\, f_{k'}(\boldsymbol{\mathcal{W}})\, f_k^*(\boldsymbol{\mathcal{W}}) \big\}, \tag{3.103}$$

which indicates that the variational energy decreases monotonically without considering stochastic noises [32] since $\boldsymbol{S}$ is a positive-definite matrix when the number of samples is finite.

For wave functions of real-valued parameters and non-holomorphic wave functions, the quantum Fisher information metric $s_{kk'}$ and energy forces f_k become

$$s_{kk'}(\boldsymbol{\mathcal{W}}) = \mathrm{Re}\{\langle O_k^*; O_{k'} \rangle\}, \tag{3.104}$$

$$f_k(\boldsymbol{\mathcal{W}}) = \mathrm{Re}\{\langle E_{\mathrm{loc}}; O_k^* \rangle\}, \tag{3.105}$$

which leads to

$$\Delta E(\boldsymbol{\mathcal{W}}) = -\eta \sum_{k,k' \neq 0} s_{kk'}^{-1}(\boldsymbol{\mathcal{W}})\, f_{k'}(\boldsymbol{\mathcal{W}})\, f_k(\boldsymbol{\mathcal{W}}) < 0\ . \tag{3.106}$$

Under consideration of stochastic noises, some eigenvalues of the Fisher matrix $\boldsymbol{S}$ that are close or equal to zero could cause such noises to increase when inverting the matrix. Thus, the invertibility of $\boldsymbol{S}$ can be ensured by adding a damping factor λ to $\boldsymbol{S}$. With this regularization, the quantum Fisher information matrix at the ith optimization step is written as

$$\boldsymbol{S}^{(i)} \to \boldsymbol{S}^{(i)} + \lambda(i)\, \mathbb{I}, \tag{3.107}$$

where $\lambda(i)$ is defined as $\max(\lambda_0 b^i, \lambda_{\min})$, and $(\lambda_{\min}, \lambda_0, b)$ is chosen as, for example, $(10^{-4}, 100, b = 0.9)$ in [33].

Chapter 4

Results

4.1 1D Bose-Hubbard model

Here, NQS are utilized for the 1D Bose-Hubbard model, which is previously analyzed with such analytical approximations as the mean-field approach and strong-coupling expansion in Sec. 2.1. For a total number of lattice sites N_s, the variational energies for N_s and $N_s \pm 1$ bosons by optimizing the network parameters of a chosen neural-network architecture. Then, one can use the following expressions for the phase boundaries $\mu_\pm$, which are also known as the charge gaps:

$$\pm\mu_\pm(N_s) = E(N_s \pm 1, N_s) - E(N_s, N_s), \tag{4.1}$$

where $E(N, N_s)$ denotes the ground state energy as a function of the total number of particles N and the number of lattice sites N_s.

For a start, a real-valued RBM is chosen as a neural-network architecture to plot the phase diagram of the 1D Bose-Hubbard model for $N_s = 6, 7, 8, 9$. As shown in Fig. 4.1, the results produced by the RBM ansatz with ADAM begin to correlate with the strong-coupling expansion as the number of lattice sites N_s increases. These results also capture the re-entrant behaviour. To get the results in the thermodynamic limit, the variational energies must be calculated for large

N_s. However, due to the definition of the RBM ansatz, an increase in N_s leads to an exponential increase in the number of neurons in the network layers when the maximum occupation number n_{max} is kept as to be equivalent to the number of bosons N, which is the case in Fig. 4.1 for more accurate results. This causes the computational cost to hike up vastly. To prevent this, an alternative but efficient neural-network architecture should be chosen.

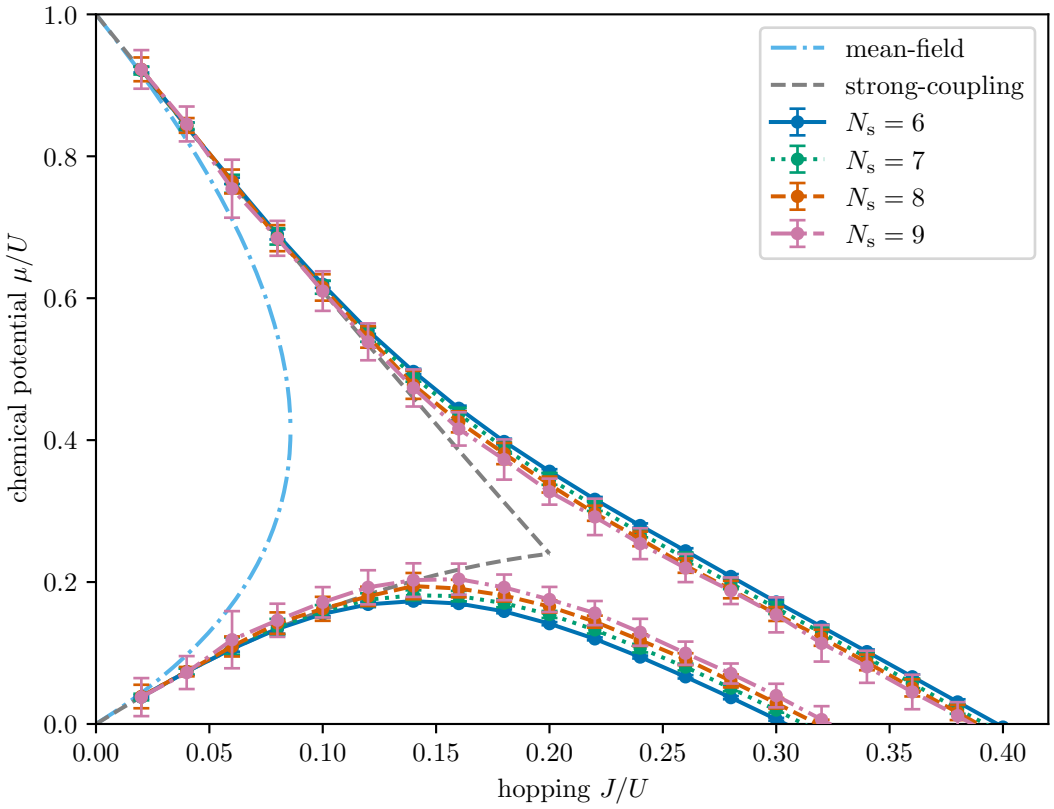

Figure 4.1: Phase diagram of the 1D Bose-Hubbard model using the restricted Boltzmann machine ansatz for different numbers of lattice sites N_s, compared to the result of the strong-coupling perturbation theory (grey dashed line) and mean-field theory (soft blue dashdot line). The error bars represent the standard deviations of the last 200 values through iterations.

Such a candidate architecture is thus a FNN, as discussed in Sec. 3.1.2. With the FNN ansatz, the phase diagram is plotted in Fig. 4.2 for $N_s = 6, 7, 8, 9$. Its resulting phase boundaries are nearly exact with the boundaries obtained via the RBM ansatz for the same N_s. To check what happens to these ground-state energies for the MI state and defect states used in determining the phase

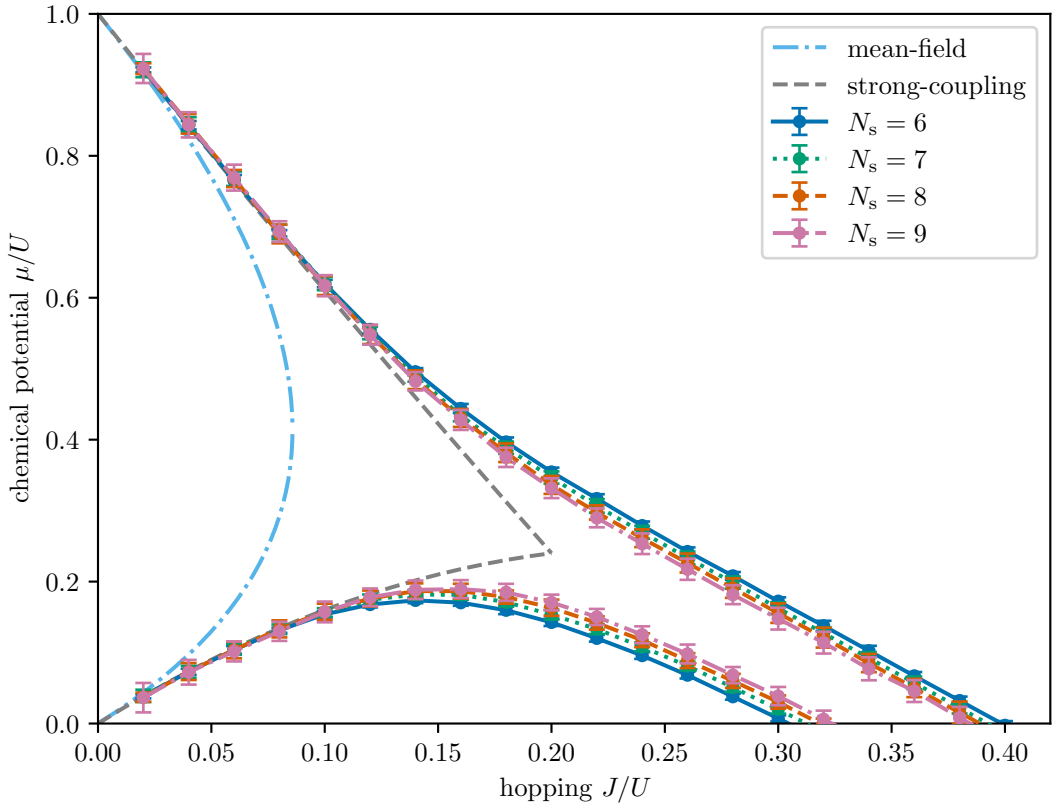

Figure 4.2: Phase diagram of the 1D Bose-Hubbard model using the feedforward neural network ansatz for different numbers of lattice sites N_s, compared to the result of the strong-coupling perturbation theory (grey dashed line) and mean-field theory (soft blue dashdot line). The error bars represent the standard deviations of the last 200 values through iterations.

boundaries in the thermodynamic limit, one can analyze the ground-state energy per boson as a function of the inverse number of lattice sites. In Fig. 4.3, this function is plotted for two different hopping amplitudes $J/U = 0.02, 0.40$ with its linear regression. For such a small hopping amplitude as $J/U = 0.02$, the energy gaps with particle-hole excitations are finite as expected since the system is more likely in the MI phase in this limit. Similarly, this finite gap vanishes at such a large hopping amplitude as $J/U = 0.40$ as the system size increases because the phase of the system becomes the SF phase. These observations correlate with the related phase analysis done in Secs. 2.1.1 and 2.1.2.

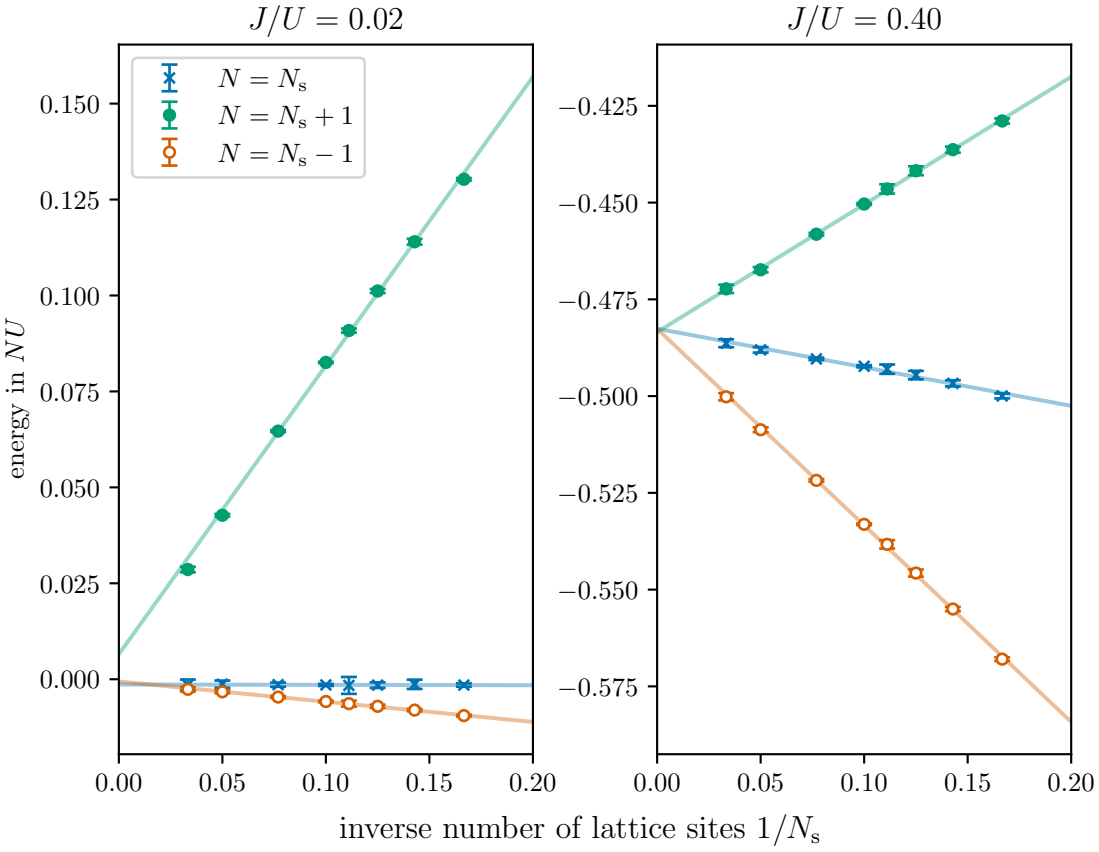

Figure 4.3: Normalized variational ground-state energy per boson as a function of the inverse number of lattice sites for two different hopping amplitudes $J/U = 0.02$ (left subplot) and $J/U = 0.40$ (right subplot), which is obtained with the feedforward neural network ansatz. The linear lines indicate the linear regressions on these variational ground-state energies. The error bars represent the standard deviations of the last 200 values through iterations.

4.2 Two-leg Bose-Hubbard ladder under synthetic magnetic flux

4.2.1 Strongly interacting regime

An initial benchmark for the RBM and FNN ansatzes in a two-leg Bose-Hubbard ladder is established by focusing on the strongly interacting regime where $J/U \ll 1$ and aiming to determine the SF-MI phase boundaries under

magnetic flux. As discussed earlier in Sec. 2.1.3.4, the Hamiltonian in Eq. (2.119) breaks the time-reversal symmetry, necessitating the use of a complex-valued restricted Boltzmann machine (ℂ-RBM).

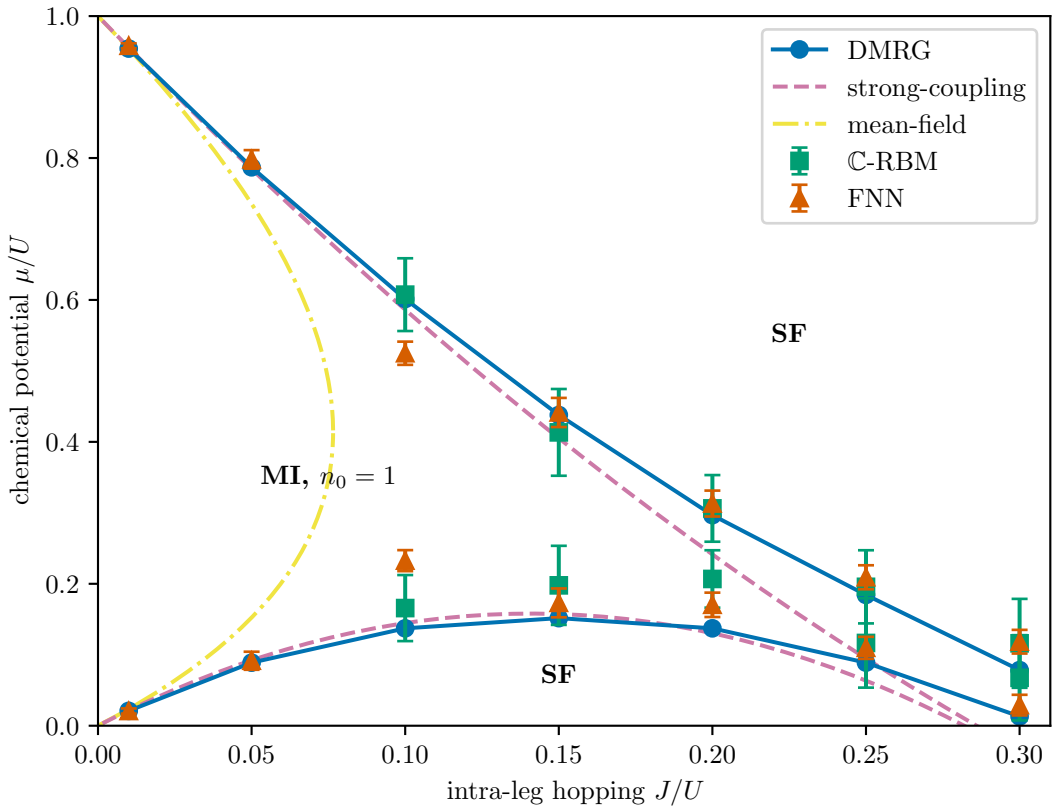

Figure 4.4: Phase diagram of a two-leg Bose-Hubbard ladder with $K/J = 1.00$ and $\phi/\pi = 0.90$ from the complex-valued RBM (green square) and FNN (orange up-triangle) ansatzes for a system of $L = 12$ rungs compared with DMRG (blue circle with line) [Keleş *et al.*, Phys. Rev. A **91**, 013629 (2015)], the strong-coupling expansion (purple dashed line) and the mean-field Gutzwiller ansatz (yellow dashdot line). The error bars represent the standard deviations of the last 200 values through iterations.

Similar to the previous section, the stochastic calculation is done to determine the phase boundaries as defined in Eq. (4.1). The resulting phase diagram in the $\mu - J$ plane, along with the results obtained with DMRG [25], the strong-coupling perturbation theory and the mean-field Gutzwiller ansatz, is illustrated in Fig. 4.4. Remarkably, when J/U is small, where the mean-field Gutzwiller ansatz offers more reliability, both ℂ-RBM and FNN exhibit an agreement with

DMRG, showcasing minimal statistical fluctuations. It is also noted that the second-order strong-coupling expansion fails to correlate with other results after $J/U \approx 0.20$. As the value of J/U increases, deviations from the DMRG data become more evident with amplified statistical fluctuations. Nevertheless, the obtained results are still superior to the mean-field approach, demonstrating the capacity of neural networks to capture basic correlations. Notably, the proximity of the MI phase is expected to facilitate the emergence of more complicated correlated phases deep within the SF phase, due to the competing effects of the magnetic field, kinetic energy, and strong interactions [74], such as fractional quantum Hall phases [5] and charge-density waves [85]. Therefore, this particular region presents a promising avenue for designing innovative neural networks in future studies, which may also find direct experimental relevance in the context of new-generation cold-atom setups. It is important to mention that in the vicinity of the tip of the Mott lobe, the transition follows a Berezinskii-Kosterlitz-Thouless type driven by phase fluctuations rather than number fluctuations. Analyzing this transition requires a different approach from Eq. (4.1), falling beyond the scope of the present thesis [86].

4.2.2 Weakly interacting regime

The regime of weaker interactions $J/U \gg 1$ outside the MI is now examined to investigate the SF phases exclusively using the FNN ansatz, demonstrating greater accuracy in [57]. The identification of these phases involves the calculations of site densities and currents along and across the legs. The overall normalization for currents is based on the maximum value $j_{\mathrm{c,max}}$ of the chiral current j_{c}. In the chiral phase, characterized by $j_{\mathrm{c}} \approx j_{\mathrm{c,max}}$, and other phases where $j_{\mathrm{c}} < j_{\mathrm{c,max}}$, this current is smaller.

Fig. 4.5 presents a cut along the phase diagram, specifically focusing on K/J for fixed $U/J = 0.20$, and $\phi/\pi = 0.50$. The expectation value of the chiral current is shown for $N = L = 16$, and the local particle densities are displayed for $N = L = 32$ in Fig. 4.6. The vortex phase (green region) exhibits gradual

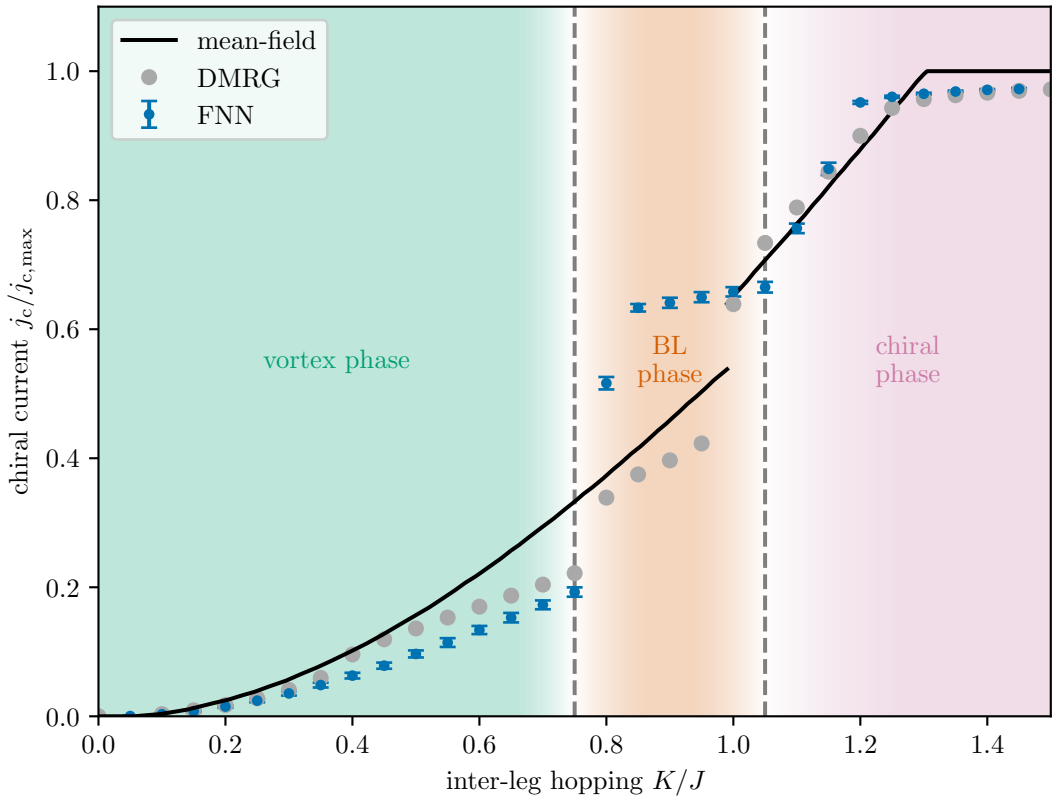

Figure 4.5: Chiral current as a function of K/J for fixed $N = L = 16$, $U/J = 0.20$, and $\phi/\pi = 0.50$ from the FNN ansatz (blue dot) normalized by the maximum current $j_{c,max} = 2\overline{n}J\sin(\phi/2) = J\sqrt{2}$ compared with the mean-field result for $L = 32$ (black line) and DMRG for $L = 64$ (dark-gray circle). The phase diagram has vortex, biased-ladder, and chiral phases. The dashed lines show the approximate phase boundaries. The error bars represent the standard deviations of the last 200 values through iterations.

growth of the chiral current, reaching a finite value significantly smaller than $j_{c,max}$. The local particle densities oscillate along the legs and are equal in the upper and lower legs. Beyond the critical point at $K/J \approx 0.75$, the chiral current has a rapid increase to a larger value, still below $j_{c,max}$. Within the intermediate BL phase (orange region), the current continues to grow slowly, and the local particle densities differ between the upper and lower legs but become uniform along the legs. Another critical point occurs at $K/J \approx 1.05$, indicating the beginning of the expected chiral phase (purple region). In this phase, the chiral current saturates close to $j_{c,max}$, and the local particle densities are still uniform along the legs, with equality between the upper and lower sites.

The convergence of the variational energy within the BL phase turns out to be notably challenging due to the delicate competition between the vortex and chiral phases. In contrast, the regions around the MI, vortex and chiral phases converged relatively quickly (within a few thousand iterations) (see Fig. 4.7), the energy in the BL phase exhibited a slow drift, taking around 60-70 thousand iterations for convergence. Additionally, the local particle density profiles underwent substantial changes before reaching the converged BL phase. The RBM ansatz, although not presented here, also performed poorly in this region. Future work could focus on exploring alternative network architectures that can accurately capture this region. Nevertheless, the ability of the FNN ansatz to reveal this phase without bias underscores the remarkable potential of neural-network quantum states in discovering novel quantum many-body phases.

To establish a benchmark, the chiral current was also calculated using the state-of-the-art DMRG simulation for $L = 64$ [87]. To obtain these results, MPS bond dimensions were increased up to 160, and 15 sweeps were performed for each data point. In order to decrease the impact of the initial random MPS configuration and avoid DMRG getting trapped in local minima, the value of K was gradually increased from zero. The optimized MPS obtained for a smaller K was used as the initial ansatz for the subsequent value $K + \delta K$. In Fig. 4.5, it can be observed that the FNN results align with the DMRG simulation in the vortex phase region, while the mean-field results exhibit significantly higher chiral currents compared to these two simulations. In the chiral phase region, there is a strong agreement among all the results. However, in the BL phase region, the chiral currents begin to deviate more noticeably from the DMRG simulation. This discrepancy is expected since DMRG employs open boundary conditions, which significantly impact the chiral current due to boundary effects.

4.3 Conclusions

The application of neural-network quantum states in this thesis focused on identifying many-body phases in the 1D Bose-Hubbard model and a two-leg Bose-Hubbard ladder under an artificial magnetic field. For the one-dimensional case, the resulting phase diagram correlates with the results obtained via strong-coupling perturbation theory, validating the theoretical predictions. Additionally, the analysis reveals energy gaps associated with particle-hole excitations, which align with the analytical predictions. This agreement between the computational findings and the theoretical expectations highlights the robustness and accuracy of the proposed variational method. Moreover, the two-leg flux ladder model's characteristics, including strong magnetic flux, inter- and intra-leg hopping, and on-site interactions, allow for much more competing quantum phases. This model serves as an ideal testing ground for evaluating the capabilities of neural networks. To address the broken time-reversal symmetry, two different neural-network architectures are employed: the RBM ansatz with complex network parameters and the FNN with real parameters but separated output neurons for the real and imaginary parts of the wave function. In the regime where on-site interactions dominate, both RBM and FNN ansatzes successfully describe the SF-MI phase transition with precision comparable to the results obtained from DMRG. In the weakly interacting regime, which features competing SF phases, the focus was primarily on the FNN ansatz, which predicts three distinct phases: a vortex phase characterized by modulations in local particle density and homogeneous superflow under weak leg coupling; a chiral phase exhibiting uniform density and leg currents in opposite directions; and a BL phase where local particle density and SF velocities differ between legs, while the total current remains zero for all of them. Importantly, it should be noted that these phases are obtained through the minimization of variational energies without any bias.

This thesis showcases the suitability of a two-leg Bose-Hubbard ladder with magnetic flux as an ideal test ground for future advancements in neural-network quantum states. Other promising neural network architectures can be explored to investigate several vital questions. For instance, the proximity of the MI

phase, where the phase boundary demonstrates enhanced fluctuations, presents an opportunity to explore strongly correlated phases [25, 88]. More sophisticated networks, such as recently developed convolutional networks with improved accuracy [49, 89], can potentially unveil these phases. Additionally, the regime near the MI phase offers excellent prospects for experimental investigations in next-generation quantum gas setups. Another important observation is the considerable slowdown in the convergence of the neural network, particularly in the weak interaction regime characterized by competing SF phases. This issue can be addressed by employing alternative ansatzes. However, it is worth noting that the FNN ansatz adequately captures the BL phase as expected. The ability to benchmark this toy model against other powerful numerical techniques like DMRG or quantum Monte Carlo using worm sampling [90] further enhances the value of future studies on this system. Furthermore, the implementation of this model in cold atom experiments, which offer a high degree of control, holds great promise for practical realizations.

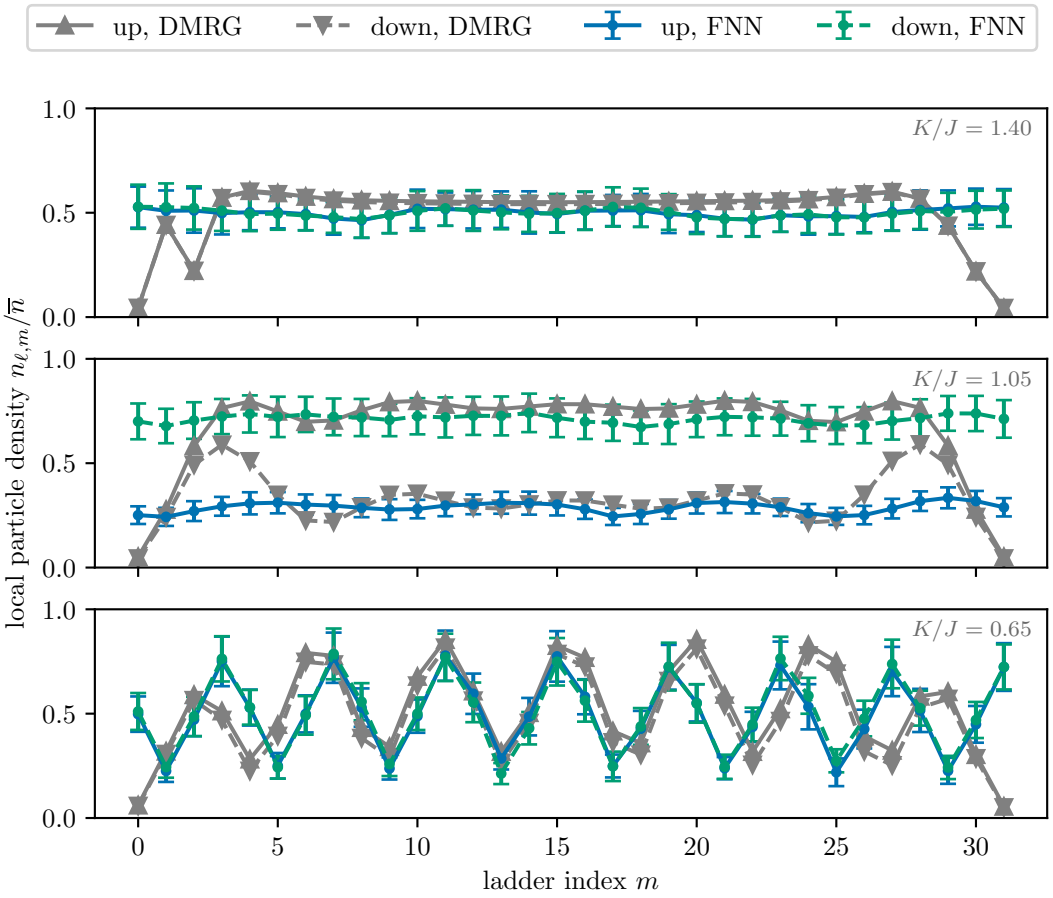

Figure 4.6: Local particle density normalized by the mean density $\overline{n} = N/L$ for $L = 32$, $U/J = 0.20$ and $\phi/\pi = 0.50$ as a function site index m in upper (blue dot with line) and lower (green dot with dashed line) legs for selected points in the phase diagram $K/J = 1.40$ (top), $K/J = 1.05$ (middle), and $K/J = 0.65$ (bottom) from the FNN ansatz, compared with DMRG (grey up- and down-triangles). The error bars represent the standard deviations of the last 200 values through iterations.

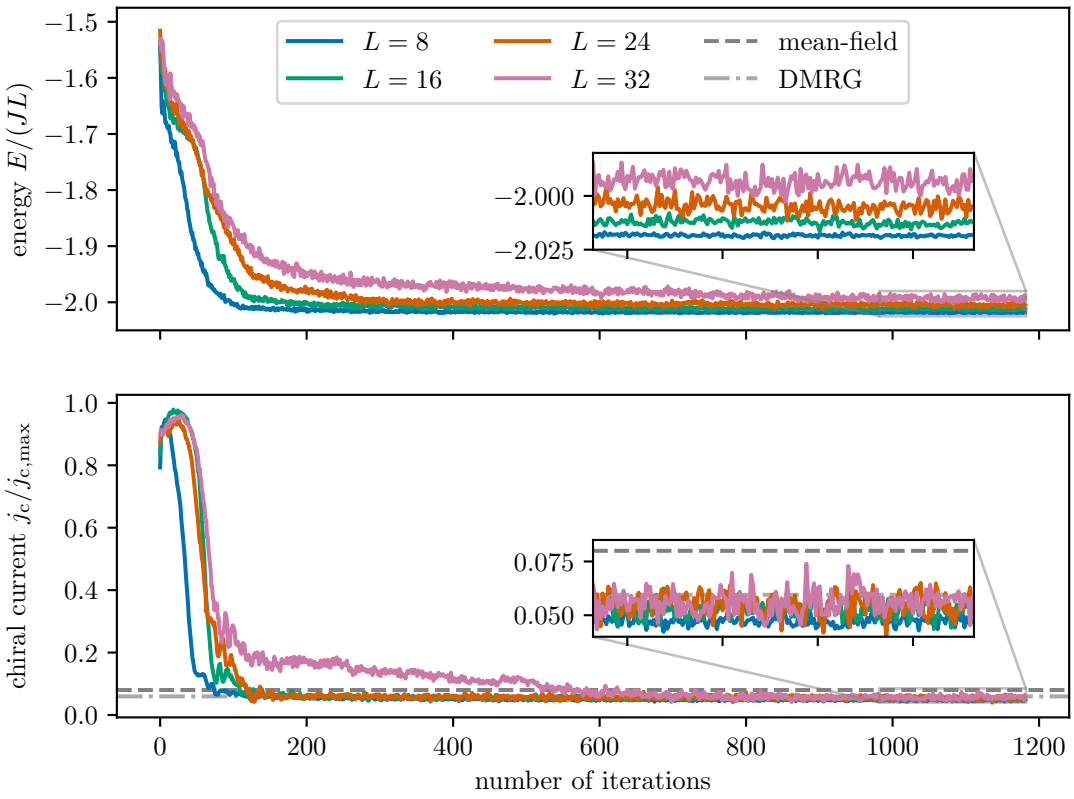

Figure 4.7: Variational energy normalized by the intra-leg hopping amplitude J times the number of rungs L and the normalized chiral current as a function of the number of ADAM iterations for fixed $K/J = 0.35$, $U/J = 0.20$, and $\phi/\pi = 0.50$ from the FNN ansatz in $L = 8$ (blue line), 16 (green line), 24 (orange line), 32 (pink line) compared with the mean-field result for $L = 32$ (dashed line) and DMRG for $L = 64$ (dash-dot line). The error bars represent the standard deviations of the last 200 values through iterations.

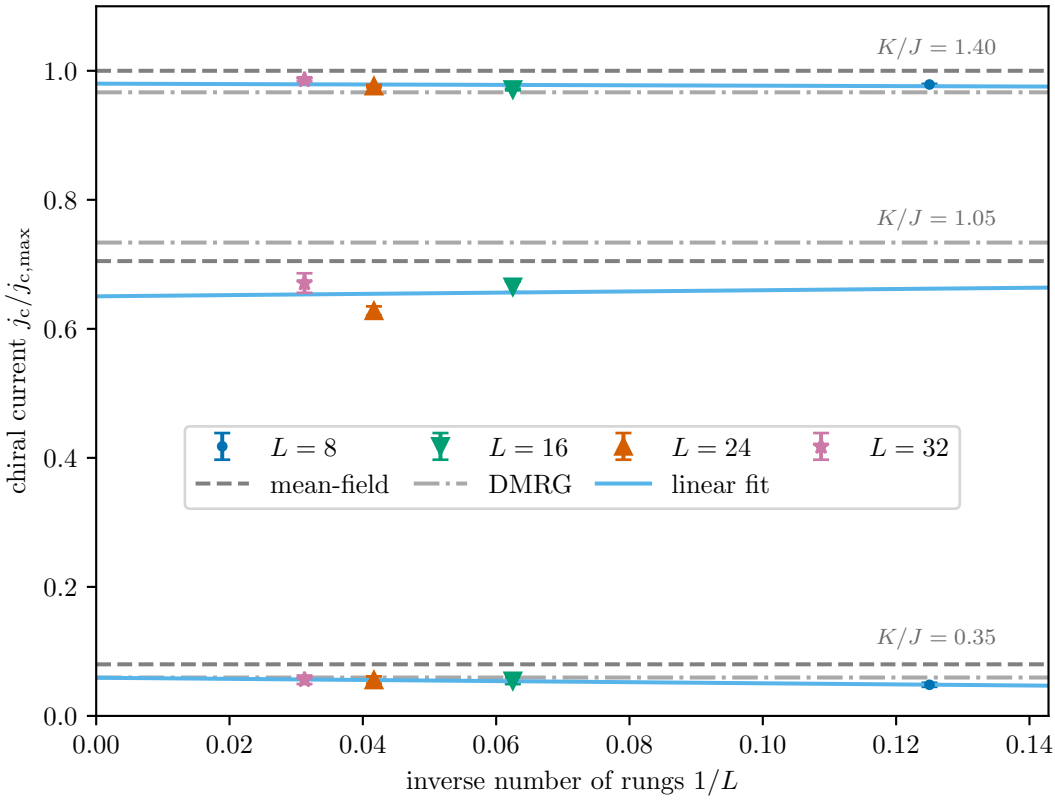

Figure 4.8: Normalized chiral current as a function of the inverse number of rungs $1/L$ compared with the mean-field result and DMRG. The light blue lines indicate the linear fitting of chiral currents over $1/L = 8^{-1}$ (blue dot), 16^{-1} (green down-triangle), 24^{-1} (orange up-triangle), 32^{-1} (pink star) for the selected points in the phase diagram $K/J = 1.40$ (top), $K/J = 1.05$ (middle), and $K/J = 0.35$ (bottom). The error bars represent the standard deviations of the last 200 values through iterations.

Bibliography

[1] C. Gross and I. Bloch, Quantum simulations with ultracold atoms in optical lattices, Science **357**, 995–1001 (2017).

[2] A. J. Daley, I. Bloch, C. Kokail, S. Flannigan, N. Pearson, M. Troyer, and P. Zoller, Practical quantum advantage in quantum simulation, Nature **607**, 667–676 (2022).

[3] I. Bloch, J. Dalibard, and W. Zwerger, Many-body physics with ultracold gases, Rev. Mod. Phys. **80**, 885–964 (2008).

[4] J. Quintanilla and C. Hooley, The strong-correlations puzzle, Phys. World **22**, 32–37 (2009).

[5] R. O. Umucalılar and E. J. Mueller, Fractional quantum Hall states in the vicinity of Mott plateaus, Phys. Rev. A **81**, 053628 (2010).

[6] F. Schäfer, T. Fukuhara, S. Sugawa, Y. Takasu, and Y. Takahashi, Tools for quantum simulation with ultracold atoms in optical lattices, Nature Reviews Physics **2**, 411–425 (2020).

[7] R. A. Hart, P. M. Duarte, T.-L. Yang, X. Liu, T. Paiva, E. Khatami, R. T. Scalettar, N. Trivedi, D. A. Huse, and R. G. Hulet, Observation of antiferromagnetic correlations in the Hubbard model with ultracold atoms, Nature **519**, 211–214 (2015).

[8] M. P. A. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher, Boson localization and the superfluid-insulator transition, Phys. Rev. B **40**, 546–570 (1989).

[9] M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch, and I. Bloch, Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms, Nature **415**, 39–44 (2002).

[10] Y.-J. Lin, R. L. Compton, K. Jiménez-García, J. V. Porto, and I. B. Spielman, Synthetic magnetic fields for ultracold neutral atoms, Nature **462**, 628–632 (2009).

[11] M. Aidelsburger, M. Atala, S. Nascimbène, S. Trotzky, Y.-A. Chen, and I. Bloch, Experimental Realization of Strong Effective Magnetic Fields in an Optical Lattice, Phys. Rev. Lett. **107**, 255301 (2011).

[12] M. Aidelsburger, M. Atala, M. Lohse, J. T. Barreiro, B. Paredes, and I. Bloch, Realization of the Hofstadter Hamiltonian with Ultracold Atoms in Optical Lattices, Phys. Rev. Lett. **111**, 185301 (2013).

[13] H. Miyake, G. A. Siviloglou, C. J. Kennedy, W. C. Burton, and W. Ketterle, Realizing the Harper Hamiltonian with Laser-Assisted Tunneling in Optical Lattices, Phys. Rev. Lett. **111**, 185302 (2013).

[14] A. Celi, P. Massignan, J. Ruseckas, N. Goldman, I. B. Spielman, G. Juzeliūnas, and M. Lewenstein, Synthetic Gauge Fields in Synthetic Dimensions, Phys. Rev. Lett. **112**, 043001 (2014).

[15] M. Atala, M. Aidelsburger, M. Lohse, J. T. Barreiro, B. Paredes, and I. Bloch, Observation of chiral currents with ultracold atoms in bosonic ladders, Nat. Phys. **10**, 588–593 (2014).

[16] B. K. Stuhl, H.-I. Lu, L. M. Aycock, D. Genkina, and I. B. Spielman, Visualizing edge states with an atomic Bose gas in the quantum Hall regime, Science **349**, 1514–1518 (2015).

[17] M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, T. Menke, D. Borgnia, P. M. Preiss, F. Grusdt, A. M. Kaufman, and M. Greiner, Microscopy of the interacting Harper–Hofstadter model in the two-body limit, Nature **546**, 519–523 (2017).

[18] M. Mancini, G. Pagano, G. Cappellini, L. Livi, M. Rider, J. Catani, C. Sias, P. Zoller, M. Inguscio, M. Dalmonte, and L. Fallani, Observation of chiral edge states with neutral fermions in synthetic Hall ribbons, Science **349**, 1510–1513 (2015).

[19] E. Anisimovas, M. Račiūnas, C. Sträter, A. Eckardt, I. B. Spielman, and G. Juzeliūnas, Semisynthetic zigzag optical lattice for ultracold bosons, Phys. Rev. A **94**, 063632 (2016).

[20] J. H. Han, D. Bae, and Y. Shin, Synthetic Hall ladder with tunable magnetic flux, Phys. Rev. A **105**, 043306 (2022).

[21] E. Orignac and T. Giamarchi, Meissner effect in a bosonic ladder, Phys. Rev. B **64**, 144515 (2001).

[22] A. Dhar, M. Maji, T. Mishra, R. V. Pai, S. Mukerjee, and A. Paramekanti, Bose-Hubbard model in a strong effective magnetic field: Emergence of a chiral Mott insulator ground state, Phys. Rev. A **85**, 041602(R) (2012).

[23] A. Petrescu and K. Le Hur, Bosonic Mott Insulator with Meissner Currents, Phys. Rev. Lett. **111**, 150601 (2013).

[24] A. Tokuno and A. Georges, Ground states of a Bose–Hubbard ladder in an artificial magnetic field: Field-theoretical approach, New J. Phys. **16**, 073005 (2014).

[25] A. Keleş and M. Ö. Oktel, Mott transition in a two-leg Bose-Hubbard ladder under an artificial magnetic field, Phys. Rev. A **91**, 013629 (2015).

[26] M. Piraud, F. Heidrich-Meisner, I. P. McCulloch, S. Greschner, T. Vekua, and U. Schollwöck, Vortex and Meissner phases of strongly interacting bosons on a two-leg ladder, Phys. Rev. B **91**, 140406(R) (2015).

[27] S. Greschner, M. Piraud, F. Heidrich-Meisner, I. P. McCulloch, U. Schollwöck, and T. Vekua, Symmetry-broken states in a system of interacting bosons on a two-leg ladder with a uniform Abelian gauge field, Phys. Rev. A **94**, 063628 (2016).

[28] M. Buser, U. Schollwöck, and F. Grusdt, Snapshot-based characterization of particle currents and the Hall response in synthetic flux lattices, Phys. Rev. A **105**, 033303 (2022).

[29] W. L. McMillan, Ground State of Liquid He^4, Phys. Rev. **138**, A442–A451 (1965).

[30] S. Sorella, Green Function Monte Carlo with Stochastic Reconfiguration, Phys. Rev. Lett. **80**, 4558–4561 (1998).

[31] J. P. F. LeBlanc, A. E. Antipov, F. Becca, I. W. Bulik, G.-L. Chan, C.-M. Chung, Y. Deng, M. Ferrero, T. M. Henderson, C. A. Jimenez-Hoyos, E. Kozik, X.-W. Liu, A. J. Millis, N. V. Prokof'ev, M. Qin, G. E. Scuseria, H. Shi, B. V. Svistunov, L. F. Tocchio, I. S. Tupitsyn, S. R. White, S. Zhang, B.-X. Zheng, Z. Zhu, and E. Gull (Simons Collaboration on the Many-Electron Problem), Solutions of the Two-Dimensional Hubbard Model: Benchmarks and Results from a Wide Range of Numerical Algorithms, Phys. Rev. X **5**, 041041 (2015).

[32] F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, Cambridge, 2017).

[33] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science **355**, 602–606 (2017).

[34] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, Rev. Mod. Phys. **91**, 045002 (2019).

[35] Yusuke Nomura, Andrew S. Darmawan, Youhei Yamaji, and Masatoshi Imada, Restricted Boltzmann machine learning for solving strongly correlated quantum systems, Phys. Rev. B **96**, 205152 (2017).

[36] D.-L. Deng, X. Li, and S. Das Sarma, Quantum Entanglement in Neural Network States, Phys. Rev. X **7**, 021021 (2017).

[37] X. Gao and L.-M. Duan, Efficient representation of quantum many-body states with deep neural networks, Nat. Commun. **8**, 1–6 (2017).

[38] Z. Cai and J. Liu, Approximating quantum many-body wave functions using artificial neural networks, Phys. Rev. B **97**, 035116 (2018).

[39] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, Equivalence of restricted Boltzmann machines and tensor network states, Phys. Rev. B **97**, 085104 (2018).

[40] S. Lu, X. Gao, and L.-M. Duan, Efficient representation of topologically ordered states with restricted Boltzmann machines, Phys. Rev. B **99**, 155136 (2019).

[41] Y. Levine, O. Sharir, N. Cohen, and A. Shashua, Quantum Entanglement in Deep Learning Architectures, Phys. Rev. Lett. **122**, 065301 (2019).

[42] Y. Huang and J. E. Moore, Neural Network Representation of Tensor Network and Chiral States, Phys. Rev. Lett. **127**, 170601 (2021).

[43] O. Sharir, A. Shashua, and G. Carleo, Neural tensor contractions and the expressive power of deep neural quantum states, Phys. Rev. B **106**, 205136 (2022).

[44] M. Kliesch, D. Gross, and J. Eisert, Matrix-Product Operators and States: NP-Hardness and Undecidability, Phys. Rev. Lett. **113**, 160503 (2014).

[45] S. Yan, D. A. Huse, and S. R. White, Spin-Liquid Ground State of the $S = 1/2$ Kagome Heisenberg Antiferromagnet, Science **332**, 1173–1176 (2011).

[46] R. G. Melko, G. Carleo, J. Carrasquilla, and J. I. Cirac, Restricted Boltzmann machines in quantum physics, Nat. Phys. **15**, 887–892 (2019).

[47] G. Lami, G. Carleo, and M. Collura, Matrix product states with backflow correlations, Phys. Rev. B **106**, L081111 (2022).

[48] K. Choo, G. Carleo, N. Regnault, and T. Neupert, Symmetries and Many-Body Excitations with Neural-Network Quantum States, Phys. Rev. Lett. **121**, 167204 (2018).

[49] K. Choo, T. Neupert, and G. Carleo, Two-dimensional frustrated J_1-J_2 model studied with neural network quantum states, Phys. Rev. B **100**, 125124 (2019).

[50] F. Ferrari, F. Becca, and J. Carrasquilla, Neural Gutzwiller-projected variational wave functions, Phys. Rev. B **100**, 125131 (2019).

[51] A. Szabó and C. Castelnovo, Neural network wave functions and the sign problem, Phys. Rev. Res. **2**, 033075 (2020).

[52] Y. Nomura and M. Imada, Dirac-Type Nodal Spin Liquid Revealed by Refined Quantum Many-Body Solver Using Neural-Network Wave Function, Correlation Ratio, and Level Spectroscopy, Phys. Rev. X **11**, 031034 (2021).

[53] N. Astrakhantsev, T. Westerhout, A. Tiwari, K. Choo, A. Chen, M. H. Fischer, G. Carleo, and T. Neupert, Broken-Symmetry Ground States of the Heisenberg Model on the Pyrochlore Lattice, Phys. Rev. X **11**, 041021 (2021).

[54] K. McBrian, G. Carleo, and E. Khatami, Ground state phase diagram of the one-dimensional Bose-Hubbard model from restricted Boltzmann machines, J. Phys. Conf. Ser. **1290**, 012005 (2019).

[55] H. Saito, Solving the Bose–Hubbard Model with Machine Learning, J. Phys. Soc. Jpn. **86**, 093001 (2017).

[56] H. Saito and M. Kato, Machine Learning Technique to Find Quantum Many-Body Ground States of Bosons on a Lattice, J. Phys. Soc. Jpn. **87**, 014001 (2018).

[57] K. Çeven, M. Ö. Oktel, and A. Keleş, Neural-network quantum states for a two-leg Bose-Hubbard ladder under magnetic flux, Phys. Rev. A **106**, 063320 (2022).

[58] R. Wei and E. J. Mueller, Theory of bosons in two-leg ladders with large magnetic fields, Phys. Rev. A **89**, 063617 (2014).

[59] H. A. Gersch and G. C. Knollman, Quantum Cell Model for Bosons, Phys. Rev. **129**, 959–967 (1963).

[60] E. Noether, Invariante Variationsprobleme, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse , 235–257 (1918).

[61] Lih-King Lim, C. Morais Smith, and Andreas Hemmerich, Staggered-Vortex Superfluid of Ultracold Bosons in an Optical Lattice, Phys. Rev. Lett. **100**, 130402 (2008).

[62] Fabrice Gerbier and Jean Dalibard, Gauge fields for ultracold atoms in optical superlattices, New Journal of Physics **12**, 033007 (2010).

[63] G. Möller and N. R. Cooper, Condensed ground states of frustrated Bose-Hubbard models, Phys. Rev. A **82**, 063625 (2010).

[64] Arya Dhar, Tapan Mishra, Maheswar Maji, R. V. Pai, Subroto Mukerjee, and Arun Paramekanti, Chiral Mott insulator with staggered loop currents in the fully frustrated Bose-Hubbard model, Phys. Rev. B **87**, 174501 (2013).

[65] S. Sachdev, Boson Hubbard Model, in *Quantum Phase Transitions* (Cambridge University Press, 2000) p. 193–202.

[66] J. J. Sakurai and J. Napolitano, Time-Reversal Operator, in *Modern Quantum Mechanics* (Cambridge University Press, 2017) pp. 275–279, 2nd ed.

[67] V. L. Berezinskii, Destruction of long range order in one-dimensional and two-dimensional systems having a continuous symmetry group. I. Classical systems, J. Exp. Theor. Phys. **32**, 493–500 (1971).

[68] J. M. Kosterlitz and D. J. Thouless, Ordering, metastability and phase transitions in two-dimensional systems, Journal of Physics C: Solid State Physics **6**, 1181 (1973).

[69] L. D. Landau, On the theory of phase transitions, J. Exp. Theor. Phys. **7**, 19–32 (1937).

[70] J. K. Freericks and H. Monien, Phase diagram of the Bose-Hubbard Model, Europhysics Letters **26**, 545 (1994).

[71] D. van Oosten, P. van der Straten, and H. T. C. Stoof, Quantum phases in an optical lattice, Phys. Rev. A **63**, 053601 (2001).

[72] J. K. Freericks and H. Monien, Strong-coupling expansions for the pure and disordered Bose-Hubbard model, Phys. Rev. B **53**, 2691–2700 (1996).

[73] D. R. Hofstadter, Energy levels and wave functions of Bloch electrons in rational and irrational magnetic fields, Phys. Rev. B **14**, 2239–2249 (1976).

[74] R. O. Umucalılar and M. Ö. Oktel, Phase boundary of the boson Mott insulator in a rotating optical lattice, Phys. Rev. A **76**, 055601 (2007).

[75] D. S. Rokhsar and B. G. Kotliar, Gutzwiller projection for bosons, Phys. Rev. B **44**, 10328–10332 (1991).

[76] W. Krauth, M. Caffarel, and J.-P. Bouchaud, Gutzwiller wave function for a model of strongly interacting bosons, Phys. Rev. B **45**, 3137–3140 (1992).

[77] K. Sheshadri, H. R. Krishnamurthy, R. Pandit, and T. V. Ramakrishnan, Superfluid and Insulating phases in an Interacting-Boson Model: Mean-Field Theory and the RPA, Europhysics Letters **22**, 257 (1993).

[78] W. K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika **57**, 97–109 (1970).

[79] F. Vicentini, D. Hofmann, A. Szabó, D. Wu, C. Roth, C. Giuliani, G. Pescia, J. Nys, V. Vargas-Calderón, N. Astrakhantsev, and G. Carleo, NetKet 3: Machine Learning Toolbox for Many-Body Quantum Systems, SciPost Phys. Codebases , 7 (2022).

[80] A. Dawid, J. Arnold, B. Requena, A. Gresch, M. Płodzień, K. Donatella, K. A. Nicoli, P. Stornati, R. Koch, M. Büttner, R. Okuła, G. Muñoz-Gil, R. A. Vargas-Hernández, A. Cervera-Lierta, J. Carrasquilla, V. Dunjko, M. Gabrié, P. Huembeli, E. van Nieuwenburg, F. Vicentini, L. Wang, S. J. Wetzel, G. Carleo, E. Greplová, R. Krems, F. Marquardt, M. Tomza, M. Lewenstein, and A. Dauphin, Modern applications of machine learning in quantum sciences (2022), arXiv:2204.04198 [quant-ph] .

[81] W. Wirtinger, Zur formalen Theorie der Funktionen von mehr komplexen Veränderlichen, Mathematische Annalen **97**, 357–376 (1927).

[82] D. P. Kingma and J. Ba, Adam: A Method for Stochastic Optimization (2014), arXiv:1412.6980 [cs.LG] .

[83] A. Graves, Generating Sequences With Recurrent Neural Networks (2014), arXiv:1308.0850 [cs.NE] .

[84] S. Amari, Natural Gradient Works Efficiently in Learning, Neural Computation **10**, 251–276 (1998).

[85] F. Grusdt and M. Höning, Realization of fractional Chern insulators in the thin-torus limit with ultracold bosons, Phys. Rev. A **90**, 053623 (2014).

[86] T. D. Kühner and H. Monien, Phases of the one-dimensional Bose-Hubbard model, Phys. Rev. B **58**, R14741–R14744 (1998).

[87] M. Fishman, S. White, and E. Stoudenmire, The ITensor Software Library for Tensor Network calculations, SciPost Physics Codebases , 4 (2022).

[88] S. Greschner, M. Piraud, F. Heidrich-Meisner, I. P. McCulloch, U. Schollwöck, and T. Vekua, Symmetry-broken states in a system of interacting bosons on a two-leg ladder with a uniform Abelian gauge field, Phys. Rev. A **94**, 063628 (2016).

[89] C. Roth and A. H. MacDonald, Group convolutional neural networks improve quantum state accuracy (2021), arXiv:2104.05085 [quant-ph] .

[90] B. Capogrosso-Sansone, Ş. G. Söyler, N. Prokof'ev, and B. Svistunov, Monte Carlo study of the two-dimensional Bose-Hubbard model, Phys. Rev. A **77**, 015602 (2008).