Analytic and machine learning approaches to composite quantum impurities

by

Wojciech Rządkowski

February, 2022

A thesis submitted to the Graduate School of the
Institute of Science and Technology Austria (IST Austria)
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Committee in charge:

Prof. Dr. Julian Fischer, Chair
Prof. Dr. Mikhail Lemeshko
Prof. Dr. Onur Hosten
Prof. Dr. Johan Mentink

Institute of Science and Technology

The thesis of Wojciech Rządkowski, titled *Analytic and machine learning approaches to composite quantum impurities*, is approved by:

Supervisor: Prof. Dr. Mikhail Lemeshko, IST Austria, Klosterneuburg, Austria

Signature: _____

Committee Member: Prof. Dr. Onur Hosten, IST Austria, Klosterneuburg, Austria

Signature: _____

Committee Member: Prof. Dr. Johan Mentink, Radboud University Nijmegen, The Netherlands

Signature: _____

Defense Chair: Prof. Dr. Julian Fischer, IST Austria, Klosterneuburg, Austria

Signature: _____

Signed page is on file

IST Austria Thesis, ISSN: 2663-337X

I hereby declare that this thesis is my own work and that it does not contain other people's work without this being so stated; this thesis does not contain my previous work without this being stated, and the bibliography contains all the literature that I used in writing the dissertation.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee, and that this thesis has not been submitted for a higher degree to any other university or institution.

I certify that any republication of materials presented in this thesis has been approved by the relevant publishers and co-authors.

Signature: _____

Wojciech Rządkowski
February, 2022

Signed page is on file

Abstract

In this Thesis, I study composite quantum impurities with variational techniques, both inspired by machine learning as well as fully analytic. I supplement this with exploration of other applications of machine learning, in particular artificial neural networks, in many-body physics.

In Chapters 3 and 4, I study quasiparticle systems with variational approach. I derive a Hamiltonian describing the angulon quasiparticle in the presence of a magnetic field. I apply analytic variational treatment to this Hamiltonian. Then, I introduce a variational approach for non-additive systems, based on artificial neural networks. I exemplify this approach on the example of the polaron quasiparticle (Fröhlich Hamiltonian).

In Chapter 5, I continue using artificial neural networks, albeit in a different setting. I apply artificial neural networks to detect phases from snapshots of two types physical systems. Namely, I study Monte Carlo snapshots of multilayer classical spin models as well as molecular dynamics maps of colloidal systems. The main type of networks that I use here are convolutional neural networks, known for their applicability to image data.

Acknowledgements

First, I thank my supervisor, Prof. Mikhail Lemeshko. I am grateful to him for finding the perfect balance between continuous scientific support and allowing me a degree of independence in pursuing my scientific interests.

I thank the members of my Thesis Committee, Prof. Johan Mentink and Prof. Onur Hosten for their collaboration and guidance. I also appreciate their help through periodic reviews of my progress on this Thesis. I am grateful to Prof. Christoph Lampert and Prof. Maksym Serbyn for hosting me during my rotation projects – skills I learned in their groups were beneficial throughout my thesis research.

I also had the pleasure to collaborate with Giacomo Bighin, Silvia Chiacchiera, Nicolò Defenu, and Andrea Trombettoni on spin models and with Athanassios Panagiotopoulos and Jakub Pękalski on colloidal systems. Working with each of them was a remarkable and enlightening experience.

Special thanks goes to the entire Lemeshko group for being great peers for all kinds of discussions. In particular, I thank Igor Cherepanov for sharing an office with me over the years.

I am grateful to my high and middle school physics teachers, Elżbieta Zawistowska and Janina Kaczmarczyk-Multarzyńska. It was them who kindled my enthusiasm for physics and taught me the foundation of the scientific method.

This Thesis would not have been possible without the unconditional support that I enjoyed in my private life. I am grateful to my family for creating an environment which empowered me to always be curious about the surrounding world. Last but definitely not least, I thank my partner Barbara. She was the greatest companion I could have ever dreamed of, both in navigating the challenges and enjoying the bright moments of those PhD years. Dziękuję.

I acknowledge financial support from the Austrian Academy of Sciences through a DOC fellowship.

I acknowledge the European Union for funding through Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 665385. I note that the EU's principle of free movement of persons greatly simplified my Austrian life.

The numerical calculations presented in this Thesis were partially powered by the IST Austria's in-house high performance computing cluster.

About the Author

I come from Warsaw, Poland.

There, in 2015, I obtained my Bachelor's degrees – in Physics at the University of Warsaw and in Electrical and Computer Engineering at the Warsaw University of Technology. In 2017, I graduated with my Master's degree in Physics, again from the University of Warsaw.

During my studies, I did several research internships in Gwangju Institute of Science and Technology (Republic of Korea), Weizmann Institute (Israel), CERN (France/Switzerland), University of Innsbruck (Austria), and IST Austria.

In 2017, I started my PhD studies at IST Austria. I affiliated with the group of Prof. Dr. Mikhail Lemeshko in 2018. In early 2020, I received a DOC Fellowship from the Austrian Academy of Sciences.

During my PhD, I did two software engineering internships at Google, in 2020 and 2021, focusing on machine learning.

List of Publications

This Thesis is based on the following publications and a preprint.

Chapter 3:
Wojciech Rzadkowski and Mikhail Lemeshko. Effect of a magnetic field on molecule–solvent angular momentum transfer. *The Journal of chemical physics*, 148(10):104307, 2018

Chapter 4:
Wojciech Rzadkowski, Mikhail Lemeshko, and Johan H Mentink. Artificial neural network states for non-additive systems. *arXiv preprint arXiv:2105.15193*, 2021

Chapter 5:
Wojciech Rządkowski, Nicolo Defenu, Silvia Chiacchiera, Andrea Trombettoni, and Giacomo Bighin. Detecting composite orders in layered models via machine learning. *New Journal of Physics*, 22(9):093026, 2020

Jakub Pękalski, Wojciech Rządkowski, and Athanassios Z Panagiotopoulos. Shear-induced ordering in systems with competing interactions: A machine learning study. *The Journal of chemical physics*, 152(20):204905, 2020

Table of Contents

List of Figures

CHAPTER 1

Introduction

1.1 Many-body physics

In the search for fundamental laws of nature, one usually starts from studying the properties of its single constituents. However, very often, interesting phenomena will only be observed in a large ensemble of particles. Let us take water as an example. By looking at just one H_2O molecule, we cannot predict that it will be vapour, liquid, or solid, depending on temperature and pressure. Even more so, by studying single molecules we cannot predict the various snowflake shapes. However, studying many-body ensembles is usually extremely difficult.

To emphasize how hard many-body physics is, let us provide two arguments, one from classical and one from quantum physics. The classical example is N bodies (e.g. celestial objects in space) with masses m_n, $n \in 1, 2, \ldots, N$. They are initially set at positions $\{\vec{x}_n(t=0)\}$ with momenta $\{\vec{p}_n(t=0)\}$, interacting with gravitational force:

$$\vec{F}_{nm} = G\frac{m_n m_m(\vec{x}_n - \vec{x}_m)}{|\vec{x}_n - \vec{x}_m|^3}. \tag{1.1}$$

The task to find their evolution $\{\vec{x}_n(t)\}$ is analytically feasible only for $N = 1$ and $N = 2$ [Heg05]. Already the task for $N = 3$ does not have a general analytic solution, despite the example being classical and only one type of interaction involved.

The quantum argument is the size of the Hilbert space, i.e. the vector space containing all possible states of the system. For a system of N quantum particles, each of them being in d possible states, the size of Hilbert space is d^N. This exponential growth with N quickly prohibits any exact calculations, even if they are numerical.

Due to the notorious challenges that many-body studies pose, some of the most important methods, concepts, and approximations in physics have been developed to tackle many-body systems. A surely non-exhaustive list of only most notable examples would include mean field theory [OS01], path integrals [GS98], density functional theory [SS11], Monte Carlo methods [NB99], quasiparticle approach [VVG+16], matrix product states [Orú14], and artificial neural networks [CCC+19]. We will use some of these techniques throughout this Thesis.

We will deal with many-body problems from the area of condensed matter physics, both solid-state and soft-matter ones. We will introduce these systems in the next sections

of this Chapter. In Chapter 2, we introduce methods that will be used to study these systems. The details and results of these studies will be presented in Chapters 3, 4, and 5.

1.2 Impurity problems

1.2.1 Quasiparticle approach

Sometimes many-body objects can be treated in a simpler way as one conceptual object capturing the most important properties of all physically existing objects. A very simple example could be a bubble of air in water, moving upwards to surface driven by buoyancy. Studying the behavior of each molecule of nitrogen or oxygen inside the bubble would be very involved. However, if we treat the bubble as one object, we can accurately compute the distance to surface as a function of time. Such considerations give rise to the concept of quasiparticles.

The *quasiparticle zoo* [VVG+16] contains a very large and constantly growing number of species. For this thesis, two quasiparticles will be particularly important. The polaron describes interaction between an electron and phonons excited by it in a crystal. The angulon describes interactions between a linear rotor and a bosonic bath in which the rotor is immersed. We will cover them in the following sections.

1.2.2 The polaron quasiparticle

When a negatively charged electron moves through a crystal, it distorts the positively charged crystal lattice. We illustrate such a physical system in Fig. 1.1.

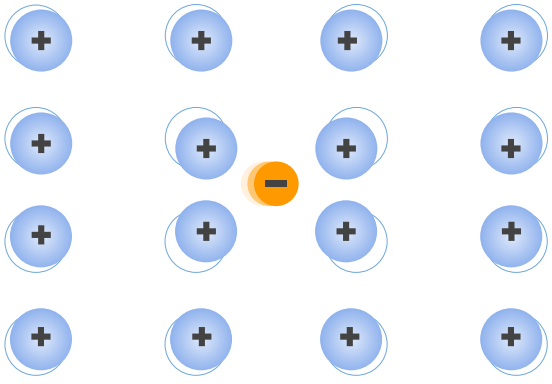


Figure 1.1: Physical system described with the polaron quasiparticle. An electron (orange) moves through and interacts with a crystal lattice (blue).

The electron interacts with excitations of the lattice, called phonons. Such interaction can be treated using the concept of quasiparticle; we imagine the travelling electron with surrounding phonons as one object – electron „dressed“ by surrounding phonons. This quasiparticle treatment of such a system is well established. The polaron dates back to work by Landau [Lan33], Pekar [Pek46], Fröhlich [Frö54] and Feynmann [Fey55].

The most general form of a Hamiltonian for such a system must assume the following form:

$$\hat{H} = \hat{H}_{\text{electron}} + \hat{H}_{\text{phonons}} + \hat{H}_{\text{electron-phonons}}. \tag{1.2}$$

One particular exemplification of this general Hamiltonian is known as Fröhlich Hamiltonian [Frö54, Dev16]. This Hamiltonian will be of importance to this thesis, so we will discuss it in detail here. The first term describes the kinetic energy of the electron. It will have the simple form

$$\hat{H}_{\text{electron}} = \frac{p^2}{2m}, \tag{1.3}$$

where m is the mass of the electron.

The second term describes the energy of the phonons written in second quantization:

$$H_{\text{phonons}} = \sum_{\mathbf{k}} \hbar\omega_{LO} \hat{a}^\dagger_{\mathbf{k}} \hat{a}_{\mathbf{k}}, \tag{1.4}$$

where ω_{LO} is the frequency of the longitudinal optical (LO) phonons in the system. The operator $\hat{a}_{\mathbf{k}}$ annihilates a phonon with wavevector $\mathbf{k}$.

The electron-phonon interactions are given by:

$$\hat{H}_{\text{electron-phonons}} = \sum_{\mathbf{k}} \left(V_{\mathbf{k}} \hat{a}_{\mathbf{k}} e^{-i\mathbf{kr}} + V^*_{\mathbf{k}} \hat{a}^\dagger_{\mathbf{k}} e^{i\mathbf{kr}} \right). \tag{1.5}$$

Here, $\mathbf{k}$ is, as already mentioned, is the phonon wavevector. The position of the electron is denoted by $\mathbf{r}$. The interaction potential $V_{\mathbf{k}}$ from Eq. (1.5) is defined by:

$$V_k = -i \frac{\hbar\omega_{\text{LO}}}{k} \left(\frac{4\pi\alpha}{V} \right)^{\frac{1}{2}} \left(\frac{\hbar}{2m_b\omega_{\text{LO}}} \right)^{\frac{1}{4}} \tag{1.6}$$

with a dimensionless coupling constant:

$$\alpha = \frac{e^2}{\hbar} \sqrt{\frac{m_b}{2\hbar\omega_{\text{LO}}}} \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0} \right). \tag{1.7}$$

Combining Eq. (1.3), (1.4), (1.5), we obtain the full Fröhlich Hamiltonian:

$$\hat{H} = \frac{p^2}{2m} + \sum_{\mathbf{k}} \hbar\omega_{LO} \hat{a}^\dagger_{\mathbf{k}} \hat{a}_{\mathbf{k}} + \sum_{\mathbf{k}} \left(V_{\mathbf{k}} \hat{a}_{\mathbf{k}} e^{-i\mathbf{kr}} + V^*_{\mathbf{k}} \hat{a}^\dagger_{\mathbf{k}} e^{i\mathbf{kr}} \right). \tag{1.8}$$

The Hamiltonian from Eq. (1.8) can be transformed for easier handling. Our intuition here is that we want to remove the impurity degrees of freedom – some methods, as we will see in Chapter 4, are better suited for interacting systems of identical entities without a distinct impurity. This goal can be achieved with the Lee-Low-Pines transformation, which brings the Hamiltonian to the impurity frame of reference. It is a unitary transformation given by the following unitary operator:

$$\hat{U} = \exp\left[\frac{i}{\hbar} \left(\mathbf{P} - \sum_k \hbar\mathbf{k} \hat{a}^\dagger_{\mathbf{k}} \hat{a}_{\mathbf{k}} \right) \cdot \mathbf{r} \right] \tag{1.9}$$

where $\mathbf{P}$ denotes the total momentum of the system. Let us then apply the transformation from Eq. (1.9) to Hamiltonian from Eq. (1.8):

$$\begin{aligned} &\hat{H}_{\text{transformed}} = \hat{U}^{-1} \hat{H} \hat{U} = \\ &\frac{1}{2m} \left(\mathbf{P} - \sum_{\mathbf{k}} \mathbf{k} \hat{a}^\dagger_{\mathbf{k}} \hat{a}_{\mathbf{k}} \right) + \sum_{\mathbf{k}} \hbar\omega_{LO} \hat{a}^\dagger_{\mathbf{k}} \hat{a}_{\mathbf{k}} + \sum_{\mathbf{k}} \left(V_{\mathbf{k}} \hat{a}_{\mathbf{k}} + V^*_{\mathbf{k}} \hat{a}^\dagger_{\mathbf{k}} \right) \end{aligned} \tag{1.10}$$

Looking at Eq. (1.10), we immediately see that, as desired, the impurity degrees of freedom, namely position **r** and momentum **p**, are no longer present in the Hamiltonian. The cost for this removal is that the first term of Eq. (1.10) now includes correlations between phonon modes, absent in the original form of the Hamiltonian in Eq. (1.8). Nevertheless, the method used in Chapter 4 is naturally able to tackle such correlations, while it is not naturally suited to deal with impurity coordinates.

1.2.3 The angulon quasiparticle

Having described the polaron quasiparticle, we now move to its rotational counterpart, the angulon. The motivation to study the system composed of a molecule immersed in a bosonic bath comes from experiment. External perturbations and thermal motion pose a big challenge to molecular spectroscopy. For precise measurement, it is desired to have the spectral lines as narrow as possible. However, the perturbations broaden these lines in two main effects called collisional broadening and Doppler broadening.

To combat these problems, it is desirable to cool down the molecules and isolate them from each other. Since 1990s, the so-called nanocryostats became a commonly used technique [TV04]. One prepares small droplets of superfluid helium and puts single molecules in them. Contrary to large portions of superfluid helium, which do not mix well with impurities [JMRW64], such droplets can stably hold the molecules.

Such nanocryostats efficiently cool down and isolate the molecules. As a result, the rotational spectra are free of both Doppler and collisional broadening. As another positive "side effect", one is able to measure spectra of molecules which in gas phase would be reactive. However, these benefits come at a cost of the energy levels being distorted by the presence of the bath. In order to properly interpret such spectra and understand the physics behind the molecule-bath interaction, one needs sufficient theoretical and/or numerical foundation.

Originally, the method of choice for analysing molecules interacting with the bath were numerical calculations [Sza08a], especially using Monte Carlo techniques. When compared with experimental data, these numerical techniques often turn out be very accurate [TM02]. However, such methods, while providing accurate numerical results, are less suitable for hinting towards physically intuitive understanding of the underlying phenomena.

This gap has been bridged by the introduction of the angulon quasiparticle [SL15, SL16, LS17, Lem17], which describes a linear rotor interacting with a many-body bosonic bath. Rotation is the only degree of freedom in the original angulon theory. Within the model, the molecule is assumed to neither move translationally, nor to possess additional angular momentum degrees of freedom such a spin or orbital angular momentum. We illustrate such a setting in Fig. 1.2.

For the derivation of angulon Hamiltonian, the bath was assumed to be a weakly-interacting Bose-Einstein condensate. The angulon Hamiltonian for such a case is:

$$\hat{H} = B\hat{J}^2 + \sum_{k\lambda\mu} \omega_k \hat{b}^\dagger_{k\lambda\mu}\hat{b}_{k\lambda\mu} + \sum_{k\lambda\mu} U_\lambda(k)\left[\hat{b}^\dagger_{k\lambda\mu}Y^\star_{\lambda\mu}(\hat{\theta},\hat{\phi}) + \hat{b}_{k\lambda\mu}Y_{\lambda\mu}(\hat{\theta},\hat{\phi})\right]. \tag{1.11}$$

Here, the first term is just the rotational energy of a rigid rotor with B being the rotational constant. As there are no other degrees of freedom, rotation is the only contribution to the total angular momentum J.

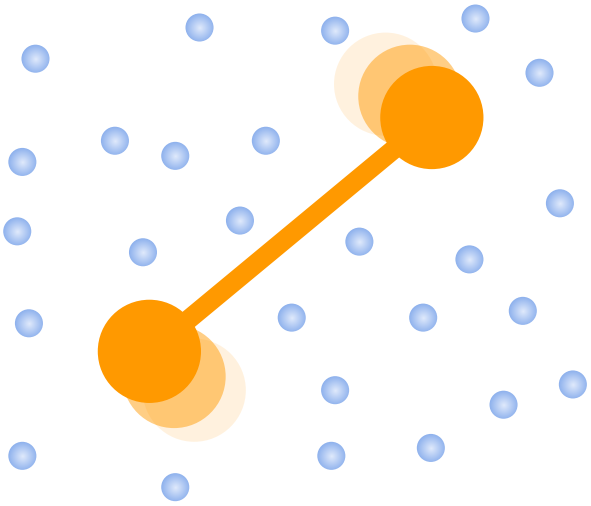


Figure 1.2: Physical system described with the angulon quasiparticle. A linear rotor (orange) interacts with a bosonic bath (blue). The molecule does not move translationally. It does not have any angular momenta (e.g. spin, orbital) other than angular momentum of the rotation.

The second term is the energy of the bath. The third term stands for the bath-impurity interaction governed by the spherical harmonic operators $Y_{\lambda\mu}(\hat{\theta},\hat{\phi})$. Here $U_\lambda(k)$ is the interaction potential in reciprocal space.

For better understanding of the underlying physics, let us sketch the derivation of the second and third term of the angulon Hamiltonian, Eq. (1.11). Let us start with the bath alone without the rotor being present, i.e. the second term. This Hamiltonian at the highest level of generality must assume the following form:

$$\hat{H}_{\text{bosons}} = H_{\text{bosons-kinetic}} + \hat{H}_{\text{boson-boson}}. \tag{1.12}$$

In our case of a weakly interacting BEC, this becomes concretely:

$$\hat{H}_{\text{bosons}} = \frac{\sum_{\mathbf{k}} |\mathbf{k}|^2 \hat{a}^\dagger_{\mathbf{k}} \hat{a}_{\mathbf{k}}}{2m} + \frac{1}{2} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} V_{\text{bb}}(\mathbf{q}) \hat{a}^\dagger_{\mathbf{k}'-\mathbf{q}} \hat{a}^\dagger_{\mathbf{k}+\mathbf{q}} \hat{a}_{\mathbf{k}'} \hat{a}_{\mathbf{k}}. \tag{1.13}$$

Here m is the mass of bosons in the bath. The quantity $V_{\text{bb}}(\mathbf{q})$ corresponds to boson-boson interaction potential, parametrized by the momentum $\mathbf{q}$ exchanged between two bosons.

To be able to handle this very general Hamiltonian in further calculations, we introduce the Bogoliubov approximation:

$$\hat{a}_{\mathbf{k}} = (2\pi)^3 \hat{\Phi}_0 \delta(\mathbf{k}) + \hat{\Phi}_{\mathbf{k}\neq 0}. \tag{1.14}$$

As of now, Eq. (1.14) is exact, but the in the next step, we will use this separation between the lowest energy $\mathbf{k} = 0$ mode and other modes. Inserting $\hat{a}_{\mathbf{k}}$ from Eq. (1.14) into Eq. (1.13), we obtain:

$$\hat{H}_{\text{bosons}} = \sum_{\mathbf{k}} \left[\frac{|\mathbf{k}|^2}{2m} + V_{\text{bb}}(\mathbf{k}) n \right] \hat{\Phi}^\dagger_{\mathbf{k}} \hat{\Phi}_{\mathbf{k}} + \frac{n}{2} \sum_{\mathbf{k}} V_{\text{b-b}}(\mathbf{k}) \left[\hat{\Phi}^\dagger_{\mathbf{k}} \hat{\Phi}^\dagger_{-\mathbf{k}} + \hat{\Phi}_{\mathbf{k}} \hat{\Phi}_{-\mathbf{k}} \right], \tag{1.15}$$

where we have neglected terms of order higher than two in $\hat{\Phi}_{\mathbf{k}}$ and dropped constant terms. This clarifies why we have used the term "approximation" when introducing the

exact transformation in Eq. (1.14). We have indeed assumed now that most bosons are in the lowest energy mode with $\mathbf{k} = 0$.

Now, we use the following transformation, known as Bogoliubov rotation, in order to express the Hamiltonian from Eq. (1.15) in terms of new operator $\hat{b}_{\mathbf{k}}$:

$$\hat{\Phi}_{\mathbf{k}} = u_{\mathbf{k}}\hat{b}_{\mathbf{k}} + v^*_{-\mathbf{k}}\hat{b}^\dagger_{-\mathbf{k}} \tag{1.16}$$

and

$$\hat{\Phi}^\dagger_{\mathbf{k}} = u^*_{\mathbf{k}}\hat{b}^\dagger_{\mathbf{k}} + v_{-\mathbf{k}}\hat{b}_{-\mathbf{k}}. \tag{1.17}$$

The goal of this transformation is to eventually diagonalize the bosonic Hamiltonian. The transformation is normalized as follows:

$$|u_{\mathbf{k}}|^2 - |v_{\mathbf{k}}|^2 = 1. \tag{1.18}$$

In our case of a stable Bose gas, the following $u_{\mathbf{k}}$, $v_{\mathbf{k}}$ diagonalize the Hamiltonian:

$$u_{\mathbf{k}} = \sqrt{\frac{\epsilon(|\mathbf{k}|) + V_{bb}(\mathbf{k})n}{2\omega(\mathbf{k})} + \frac{1}{2}} \tag{1.19}$$

and

$$v_{\mathbf{k}} = -\sqrt{\frac{\epsilon(|\mathbf{k}|) + V_{bb}(\mathbf{k})n}{2\omega(\mathbf{k})} - \frac{1}{2}}. \tag{1.20}$$

Here we have introduced $\epsilon(\mathbf{k})$ as short-hand notation for the kinetic energy:

$$\epsilon(\mathbf{k}) = \frac{|\mathbf{k}|^2}{2m}, \tag{1.21}$$

while the dispersion relation $\omega_{\mathbf{k}}$ is given by:

$$\omega_{\mathbf{k}} = \sqrt{\epsilon(\mathbf{k})(\epsilon(\mathbf{k}) + 2V_{bb}(\mathbf{k})n)}. \tag{1.22}$$

Finally, we substitute Eqs. (1.16, 1.17) into Eq. (1.15). This brings the Hamiltonian from Eq. (1.15) to the following form:

$$\hat{H} = \sum_{\mathbf{k}} \omega_k \hat{b}^\dagger_{\mathbf{k}}\hat{b}_{\mathbf{k}}, \tag{1.23}$$

thus concluding the derivation of the bosonic bath Hamiltonian.

Having outlined the derivation of the bath Hamiltonian, let us proceed to deriving the impurity-bath Hamiltonian, i.e. the third term of Eq. (1.11). In its general form, it the impurity-bath Hamiltonian will be:

$$\hat{H}_{\text{mol-bos}} = \sum_{\mathbf{k},\mathbf{q}} \hat{V}_{\text{mol-bos}}\,(\mathbf{q}, \hat{\phi}, \hat{\theta}, \hat{\gamma}) e^{-\mathbf{q}\hat{\mathbf{r}}} \hat{a}^\dagger_{\mathbf{k}+\mathbf{q}}\hat{a}_{\mathbf{k}}. \tag{1.24}$$

For our needs, it will be convenient to expand some quantities in terms of spherical harmonics [VMK88] $Y_{lm}(\theta, \phi)$. These are a set of orthogonal functions defined in spherical coordinates. They are the spatial wavefunction of the angular momentum eigenstates:

$$\langle \theta, \phi | l, m \rangle = Y_{lm}(\theta, \phi). \tag{1.25}$$

We first expand the interaction potential in terms of the spherical harmonics:

$$\hat{V}_{\text{mol-bos}} = \sum_{\lambda} V_{\lambda}(r) Y_{\lambda 0}(\theta_r, \phi_r). \tag{1.26}$$

The angles used in Equations (1.24) and (1.26) will be central for our derivation. Let us summarize definitions of all angles used in the derivation of the rotor-bath Hamiltonian in one place below:

- θ_r, ϕ_r are the spherical coordinates in the molecular frame of reference,
- Θ_R, ϕ_R are the spherical coordinates in the laboratory frame of reference,
- $\hat{\theta}, \hat{\phi}$ are angle operators of the rotation between frames of reference marked by Θ_R, Φ_R and θ_r, ϕ_r. In general, there should be third rotation angle $\hat{\gamma}$, but it is equal to zero for linear molecules. The angles θ, ϕ, γ are known in literature as Euler angles [Eul76].

We also need to expand the bosonic operators $\hat{b}^{\dagger}_{\mathbf{k}}$ and $\hat{b}_{\mathbf{k}}$ in terms of the spherical harmonics. The expression for $\hat{b}^{\dagger}_{k\lambda\mu}$ is given by [LS17]:

$$\hat{b}^{\dagger}_{k\lambda\mu} = \frac{k}{(2\pi)^{3/2}} \int d\Phi_k d\Theta_k \ \sin\Theta_k \ \hat{b}^{\dagger}_{\mathbf{k}} \ i^{-\lambda} \ Y_{\lambda\mu}(\Theta_k, \Phi_k). \tag{1.27}$$

This obeys the proper commutation relations:

$$[\hat{b}_{k\lambda\mu}, \hat{b}_{k'\lambda'\mu'}] = \delta(k - k')\delta_{\lambda\lambda'}\delta_{\mu\mu'}, \tag{1.28}$$

while the inverse transformation is given by:

$$\hat{b}^{\dagger}_{\mathbf{k}} = \frac{(2\pi)^{3/2}}{k} \sum_{\lambda\mu} \hat{b}^{\dagger}_{k\lambda\mu} i^{\lambda} Y^{*}_{\lambda\mu}(\Theta_k, \Phi_k). \tag{1.29}$$

To convert them to the molecular frame of reference, we use the following property of spherical harmonics:

$$Y_{\lambda 0}(\theta_r, \phi_r) = \sum_{\mu} \sqrt{\frac{4\pi}{2\lambda + 1}} \hat{Y}^{*}_{\lambda\mu}(\hat{\theta}_r, \hat{\phi}_r) Y_{\lambda\mu}(\Theta_R, \Phi_R). \tag{1.30}$$

Then, we convert it to momentum space using the Fourier transformation:

$$\begin{aligned} \hat{V}_{\text{mol-bos}}(\mathbf{k}, \hat{\theta}, \hat{\phi}) &= \int d^3R \hat{V}_{\text{mol-bos}}(\mathbf{R}, \hat{\theta}, \hat{\phi}) e^{-i\mathbf{k}\mathbf{R}} \\ &= \sum_{\lambda\mu} (2\pi)^{3/2} i^{-\lambda} \tilde{V}_{\lambda}(k) Y_{\lambda\mu}(\Theta_k, \Phi_k) \hat{Y}^{*}_{\lambda\mu}(\hat{\theta}, \hat{\phi}) \end{aligned} \tag{1.31}$$

We again apply the Bogoliubov approximation:

$$\hat{H}_{\text{mol-bos}} = n\hat{V}_{\text{mol-bos}}(\mathbf{k} = 0, \hat{\theta}, \hat{\phi}) + \sqrt{n} \sum_{\mathbf{k}} \hat{V}_{\text{mol-bos}}(\mathbf{k}, \hat{\theta}, \hat{\phi}) \left(\hat{\Phi}^{\dagger}_{\mathbf{k}} + \hat{\Phi}_{-\mathbf{k}} \right) \tag{1.32}$$

and Bogoliubov transformation:

$$n\hat{V}_{\text{mol-bos}}\left(\mathbf{k}=0,\hat{\theta},\hat{\phi}\right)+\sqrt{n}\sum_{\mathbf{k}}\hat{V}_{\text{mol-bos}}\left(\mathbf{k},\hat{\theta},\hat{\phi}\right)\sqrt{\frac{\epsilon(\mathbf{k})}{\omega(\mathbf{k})}}\left(\hat{b}_{\mathbf{k}}^{\dagger}+\hat{b}_{-\mathbf{k}}\right). \tag{1.33}$$

We conclude by reaching our interaction potential from Eq. (1.11):

$$\sum_{k\lambda\mu}U_{\lambda}(k)\left[\hat{b}_{k\lambda\mu}^{\dagger}Y_{\lambda\mu}^{\star}(\hat{\theta},\hat{\phi})+\hat{b}_{k\lambda\mu}Y_{\lambda\mu}(\hat{\theta},\hat{\phi})\right] \tag{1.34}$$

with

$$U_{\lambda}(k)=\left[\frac{8nk^{2}\epsilon(k)}{\omega(k)(2\lambda+1)}\right]^{1/2}\int drr^{2}V_{\lambda}(r)j_{\lambda}(kr). \tag{1.35}$$

Here $j_{\alpha}(x)$ denotes the spherical Bessel function of the first kind:

$$j_{\alpha}(x)=\sum_{m=0}^{\infty}\frac{(-1)^{m}}{m!\Gamma(m+\alpha+1)}\left(\frac{x}{2}\right)^{2m+\alpha}. \tag{1.36}$$

In this section, we have covered the original angulon theory. Many extensions have been developed over the recent years. In Chapter 3, we cover an extension developed as part of this thesis, namely an angulon with additional spin degree of freedom. The spin degree of freedom naturally couples to a magnetic field, hence we will consider both the case with and without the presence of an external magnetic field.

Other developments to date include angulon in a static electric field, dubbed „pendulon" [RL16], angulon in time-dependent electromagnetic field [YL17] or the theory of a rotating symmetric top inside a bath [CL17]. Addition of translational degree of freedom has been considered and named the "rotating polaron" problem [YMD+18, MLY20]. Multiple rotating impurities have also been studied [LYB+20]. The angulon has been proposed as microscopic explanation of the Einstein-de Haas effect [MKL19].

A path-integral treatment of the angulon has been developed [BL17]. Interesting connections to topology [YDL17, YL18, YGL+20, BLLY21a, BLLY21b] as well as to quantum group theory [YSL18] have also been identified. The dynamical properties of the angulon have been studied in collaboration with experimentalists [CBC+19, CCS+20, CBS+21] in order to explain the spectra of helium-immersed molecules excited with laser fields.

1.3 Lattice systems

The presence of a lattice is a fundamental property of crystals. A D-dimensional lattice is an infinitely stretched set of discrete points in $\mathbb{R}^{D}$, whose positions are given by:

$$\mathbf{R}=n_{1}\mathbf{a}_{1}+n_{2}\mathbf{a}_{2}+\ldots+n_{D}\mathbf{a}_{D}. \tag{1.37}$$

Here $n_{1},n_{2},\ldots,n_{D}$ are integers, while $\mathbf{a}_{1},\mathbf{a}_{2},\ldots,\mathbf{a}_{D}$ are vectors defining the primitive translations of the lattice.

In this Thesis, we will focus on the square lattice in two dimensions. The lattice vectors as defined in Eq. (1.37) are in this case simply given by:

$$\mathbf{a}_{1}=[0,1],\quad \mathbf{a}_{2}=[1,0]. \tag{1.38}$$

Let us also note that also with multiple layers, these models are essentially two-dimensional. The layers do not make the lattice three dimensional as the lattice should stretch infinitely in its dimensions.

In this section, we will concentrate on spin lattice models, where each site of the lattice is occupied by by a single spin. We will be studying the following concrete spin lattice models:

- bilayer Ising model,
- trilayer Ising model,
- Ashkin-Teller model.

We will now introduce these models briefly. More details can be found in Chapter 5 along with the results of their study.

1.3.1 Ising model

The Ising model is a paradigmatic model used to study ferromagnetism using the methods of statistical physics. The Hamiltonian of the Ising model is given by

$$\hat{H} = -J \sum_{<ij>} \sigma_i \sigma_j, \tag{1.39}$$

where $< ij >$ denotes summation over all pairs of nearest neighbours on the lattice and $\sigma_i = \pm 1$ denotes the classical Ising spin variable.

Let us note that the form of the Hamiltonian in Eq. (1.39) involves abstract summation over nearest neighbours, but does not define the lattice. This Hamiltonian is universal and hence can be studied on many types of lattices and the physical properties of the models will change depending on the lattice. However, as already said, we will focus on square lattices.

One can couple multiple square lattice Ising models to obtain a multilayer Ising model. The simplest such system is the bilayer Ising model with two layers. The Hamiltonian will be a natural extension of the Hamiltonian from Eq. (1.39) with an additional coupling constant K, which mediates the strength of coupling between the layers.

$$\hat{H} = -J \sum_{<ij>} \sigma_i \sigma_j - J \sum_{<ij>} \tau_i \tau_j - K \sum_i \sigma_i \tau_j. \tag{1.40}$$

The quadratic coupling in Eq. (1.40) is known as Yukawa coupling. The Ising spin variables $\sigma_i, \tau_i = \pm 1$ correspond to the first and second layer, respectively.

For three layers, we proceed analogously. Hence, the Hamiltonian of the trilayer Ising model on a square lattice will be:

$$\hat{H} = -J \sum_{<ij>} \sigma_i \sigma_j - J \sum_{<ij>} \tau_i \tau_j - J \sum_{<ij>} \nu_i \nu_j - K \sum_i \sigma_i \tau_i - K \sum_i \tau_i \nu_i. \tag{1.41}$$

This is, again, a Hamiltonian comprising individual single-layer Ising models coupled with Yukawa coupling. Now there are three Ising spin variables $\sigma_i, \tau_i, \nu_i = \pm 1$ corresponding to the three layers.

1.3.2 Ashkin-Teller model

The Ashkin-Teller model is inherently multilayer. It does not exist in single-layer form. It consists of two square lattice ising models with a quartic interlayer coupling. The Hamiltonian of the Ashkin-Teller model is given by

$$\hat{H}_{AT} = -J \sum_{<i,j>} \sigma_i \sigma_j - J \sum_{<i,j>} \tau_i \tau_j - K \sum_{<i,j>} \sigma_i \sigma_j \tau_i \tau_j. \tag{1.42}$$

Here σ_i and τ_i are again Ising spin variables. Due to more involved coupling, the Asking-Teller model features a rich phase diagram, partially characterized by composite order parameters. We will discuss it in more detail in Chapter 5.

1.4 Colloidal systems

Lattice systems, such as those discussed in Section 1.3, can be used to explain a wide range of phenomena in solid-state systems. However, they are less applicable to soft matter systems, which are central to our understanding of living creatures [BH03]. In this section, we will provide an introduction to the basic physics of complex fluids. We will start with fluids in general, then proceed to complex fluids, and then to colloidal systems. At this point, let us introduce the following standard terminology: "fluid" refers to both "gas" and "liquid", while "condensed matter" refers to both "liquid" and "solid".

Fluids can be divided into two basic types: simple and complex fluids. Simple fluids refer to fluids whose interaction potential is spherically symmetric and does not contain both attractive and repulsive parts. An example of such interactions can be the purely attractive Lennard-Jones potential, which we discuss later in this chapter. Examples in nature include molecular nitrogen, ammonia, methane or atomic noble gases like argon. Their phase diagrams are simple and well studied, typically consisting of just solid, liquid and vapour phases. In the gas phase, they are completely disordered; in the low density limit it is reasonable to entirely neglect the intermolecular interaction in an approximation known as ideal gas model. In the liquid phase, the substance becomes densely packed and ordered at short scales of a few intermolecular distances. Crystallization to solid phase is associated with the appearance of a long-range order.

Complex fluids are characterized by more involved interaction between the individual constituents. Their interactions can contain both attractive and repulsive parts and/or be non-spherically symmetric. Such interactions lead to an emergence of much richer phase diagrams. An example that we will discuss in this Thesis is a spherically symmetric potential with repulsive part manifested at short distances between the particles and the attractive part dominating the longer distances. Such systems are called short-range attraction, long-range repulsion (SALR). This competition between the interactions leads to a much richer phase diagram, in which the constituents of the SALR systems self-assemble into different types of patterns [GLL09, WG02].

Such competition between the interactions can be a property of a wide range of physically different systems. Such interactions are, for example, manifested in globular proteins [SSC+04a], which after folding form spherical structure. Another example, which we will consider in this thesis, are colloidal systems, where spherical insoluble are suspended in another substance.

For mathematical modelling of SALR systems, a simple and widely used model is the combination of Lennard-Jones and Yukawa potentials. Their are responsible for modelling the attractive and repulsive parts, respectively. The Lennard-Jones potential is given by:

$$v_{\mathrm{LJ}}(r) = 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right). \tag{1.43}$$

Here σ is the distance at which the potential crosses zero (sometimes referred to as atomic diameter). The parameter ϵ regulates the depth of the attractive well of the potential and is often called dispersion energy. The Lennard-Jones potential is valid if we assume spherical symmetry of the interactions. For anisotropic interactions, one can add an additional term to the potential from Eq. (1.43), for example a dipole-dipole interaction potential. In this Thesis, however, we will concentrate on spherically symmetric interactions.

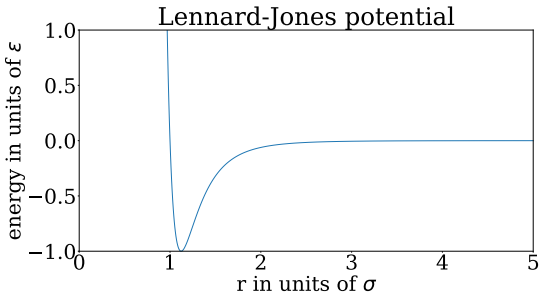
The Yukawa potential is given by

$$v_{\mathrm{Yukawa}}(r) = \frac{A}{r}e^{-r\kappa} \tag{1.44}$$

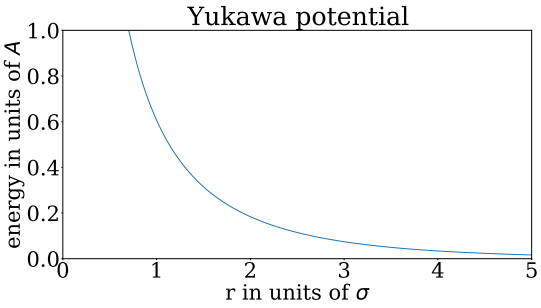
Here A is the magnitude of the potential, while the parameter κ sets the range of the potential to $1/\kappa$. The Yukawa potential is sometimes referred to as "screened Coulomb potential" as the exponential decay of the $e^{-r\kappa}$ term screens the Coulomb interaction A/r, which decays much slower.

In Fig 1.3, we illustrate both Lennard-Jones (panel a) and Yukawa (panel b) potentials as well as their sum (panel c).

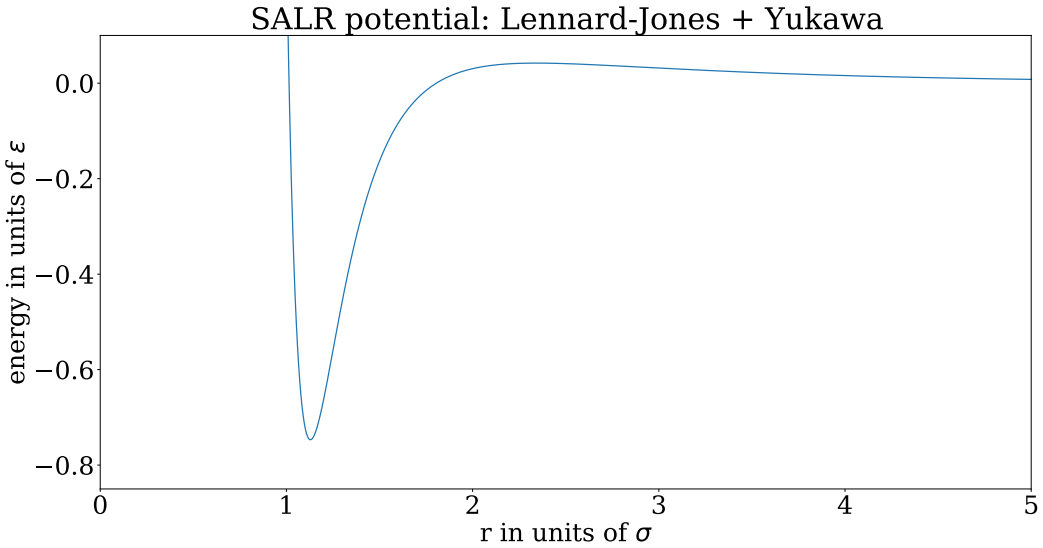
In this Thesis, we will be focusing on homogeneous systems. We will not consider interfaces with other substances. However, it is worth noting that pattern formation effects occur for example at oil-water interfaces [MFKM10], or particles adsorbed on membranes [ŠC12, ŠC13] .



(a) The attractive Lennard-Jones potential.



(b) The repulsive Yukawa potential.



(c) Short-range attraction, long range repulsion (SALR) potential resulting from summing both Lennard-Jones and Yukawa potentials.

Figure 1.3: The SALR potential used in this thesis (panel c) and its two contributors: the Lennard-Jones (panel a) and Yukawa (panel b) potentials. The y-axes have different limits in each of the panels to show the most relevant details of each potential. The x-axes are the same in all panels. See text for details.

CHAPTER 2

Methods

2.1 Variational approach

Given a Hamiltonian of a system, $\hat{H}$, solving a static quantum mechanical problem amounts to solving the time-independent Schrödinger equation [LL13, GS18]:

$$\hat{H}\,|\psi\rangle = E\,|\psi\rangle\,. \tag{2.1}$$

Solving this equation corresponds to finding a set of orthogonal eigenstates and corresponding eigenenergies:

$$\hat{H}\,|\psi_i\rangle = E_i\,|\psi_i\rangle\,, \quad \langle\psi_i|\psi_j\rangle = \delta_{ij}, \tag{2.2}$$

where δ_{ij} is the Kronecker delta.

This is exactly feasible only for few systems. For more involved systems, one needs to resort to approximate techniques. One of them is the variational approach, which we will discuss here. Among all eigenenergies of a given quantum system, the one with particular importance is the ground state, i.e. the state characterized by the lowest energy. In the variational approach, one picks up a family of candidate states parametrized by a set of parameters $\mathcal{W}$

$$|\psi_{\mathcal{W}}\rangle\,, \quad \mathcal{W} = \xi_1, \xi_2, \ldots. \tag{2.3}$$

Then, for any Hamiltonian $\hat{H}$ and given configuration of the variational parameters $\mathcal{W}$, one can compute the expected value of energy as follows:

$$E_{\mathcal{W}} = \frac{\langle\psi_{\mathcal{W}}|\,\hat{H}\,|\psi_{\mathcal{W}}\rangle}{\langle\psi_{\mathcal{W}}|\psi_{\mathcal{W}}\rangle}. \tag{2.4}$$

Naturally, this energy depends on the variational parameters. In variational approach, we are interested in the lowest energy state, so we minimize the energy $E_{\mathcal{W}}$ over all possible variational parameters $\mathcal{W}$ to obtain our variational energy $E^{\text{opt}}_{\mathcal{W}}$:

$$E^{\text{opt}}_{\mathcal{W}} = \min_{\mathcal{W}} E_{\mathcal{W}} = \min_{\mathcal{W}} \frac{\langle\psi_{\mathcal{W}}|\,\hat{H}\,|\psi_{\mathcal{W}}\rangle}{\langle\psi_{\mathcal{W}}|\psi_{\mathcal{W}}\rangle}. \tag{2.5}$$

The quantity $E_{\mathcal{W}}^{\text{opt}}$ is our final solution from the variational method. It is always larger or equal to the ground state energy:

$$E_{\mathcal{W}}^{\text{opt}} \geq E_0 = \frac{\langle\psi_0|\,\hat{H}\,|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle}, \tag{2.6}$$

where $|\psi_0\rangle$ is the true ground state and the equality holds if the true ground state belongs to the class of states represented by the variational ansatz.

Hence, the most difficult and the most impactful part of solving any Hamiltonian with variational approach is to choose a well-suited ansatz. The choice of variational ansatz is usually based on the physical intuition about the given system. In what follows, we will describe one particular type of variational ansatzes, inspired by artificial neural networks.

2.2 Variational approach with large number of structured parameters: the birth of neural-network quantum states

In the previous section, we have discussed the variational approach in general. Let us now introduce the recent developments, where variational parameters are structured into a neural-network-inspired architecture. In principle, there is no fundamental limit to the number of variational parameters. One needs, however, to take into account that with increasing number of parameters, finding the set of parameters that minimize the variational energy becomes increasingly difficult.

Recent times witnessed huge, fast-paced developments in machine learning, especially deep learning. These developments were caused by increasing ability to train large neural networks, especially using GPUs. They are sometimes referred to as the "deep learning revolution", often associated with a deep network called AlexNet [KSH12] winning the 2012 ImageNet image classification challenge with a huge 10.8 % advantage over the runner-up.

Before proceeding to variational ansatzes inspired by artificial neural networks, let us provide a brief introduction to artificial neural networks in general. They are a framework to decompose functions, often with highly-dimensional arguments (input) and multi-dimensional values (output) into a composition of smaller functions. The individual building blocks (i.e. the basic functions that are not compositions of other functions) are called neurons. They are organised sequentially in groups called layers. Such approach is inspired by life sciences and aims to provide a simplified model of the function of the brain.

In the most basic form, called multilayer perceptron, each of the layers typically takes the entire previous layer as input and implements the following function

$$\mathbf{h}_j\left(\mathbf{h}_{j-1}\right) = \sigma\left(\mathbf{W}_j^{j-1}\mathbf{h}_{j-1} + \mathbf{b}_j\right). \tag{2.7}$$

Let us explain this equation in detail. Here $\mathbf{h}_k$ denotes the k-th layer of the network. They are vectors of shape $(\text{neurons}(k), 1)$, where $\text{neurons}(k)$ is the number of neurons in the k-th layer. The matrix W_{j-1}^{j} contains trainable weights that perform a linear transformation between the layers j and $j+1$. This matrix has shape $(\text{neurons}(j), \text{neurons}(j-1))$. The vector $\mathbf{b}_j$ is the bias of the j-th layer and has shape $(\text{neurons}(j), 1)$.

The aim of the activation function $\sigma(x)$ is to introduce nonlinearity. Otherwise, multiple linear layers would just trivially collapse into a single linear transformation. Popular choices for the activation function are tanh, ReLU:

$$\mathrm{ReLU}(x) = \begin{cases} x, & x >= 0 \\ 0, & x < 0 \end{cases} \tag{2.8}$$

or the sigmoid function:

$$\mathrm{sigmoid}(x) = \frac{e^x}{e^x + 1}. \tag{2.9}$$

Over the years, numerous detailed architectures of neural networks have been developed. In this Thesis, we will be using two specific architectures that we will discuss in details in the following part of the Thesis.

The applications of artificial neural networks are very broad. Just to name a few interesting examples, they have been successful in playing games at superhuman level [SHM+16], solving partial differential equations [LKA+20], providing natural language machine translation [Sta20] or filtering e-mail spam with high accuracy [DBC+19].

The training process corresponds to finding an optimal value of cost function dependent on a large number of parameters, in deep learning often millions of them [GBC16]. This setting is naturally very similar to the variational setting in quantum mechanics. This similarity sparked the interest of quantum physics researchers as the successes of machine learning could potentially be repeated in quantum physics by harnessing the power of the machine learning concepts to boost the variational approach.

The landmark idea, that we will also develop in this Thesis, was introduced by Carleo and Troyer in Ref. [CT17]. There, the Restricted Boltzmann Machine, which is a well-established concept in computer science [Mon16], is used as a variational ansatz for spin-1/2 systems. In this introduction, we will concentrate on one-dimensional spin chains.

On top of the visible spins v_i, $i \in 1 \dots N$, the authors introduced an additional „hidden" layer of spins h_j, $j \in 1 \dots M$. The complex variational parameters are $\{a_i\}$, $\{b_j\}$ and $\{W_{ij}\}$. They can be thought of as follows:

- a_i are visible spin biases,
- b_j are hidden spin biases,
- W_{ij} are weights, which densely connect all visible spins with all hidden spins.

Then, the ansatz is:

$$\psi(\vec{s}) = \langle \vec{s} | \psi \rangle = \sum_{\{h_j\}} \exp\left(\sum_i a_i s_i + \sum_j b_j h_j + \sum_{ij} W_{ij} s_i h_j \right), \tag{2.10}$$

where $|\vec{s}\rangle$ is a Hilbert space basis vector corresponding to the visible layer (i.e. physical spins) configuration:

$$\vec{s} = s_1, s_2, \dots, s_N. \tag{2.11}$$

The sum goes over all possible hidden spin configurations. Hence, it can be computed exactly and the ansatz assumes the following form:

$$\psi(\vec{s}) = \langle \vec{s} | \psi \rangle = \sum_{\{h_j\}} \exp\left(\sum_i a_i s_i\right) \prod_j \left(b_j + \sum_i W_{ij} s_i\right). \tag{2.12}$$

We visualize this approach graphically on Fig. 2.1.

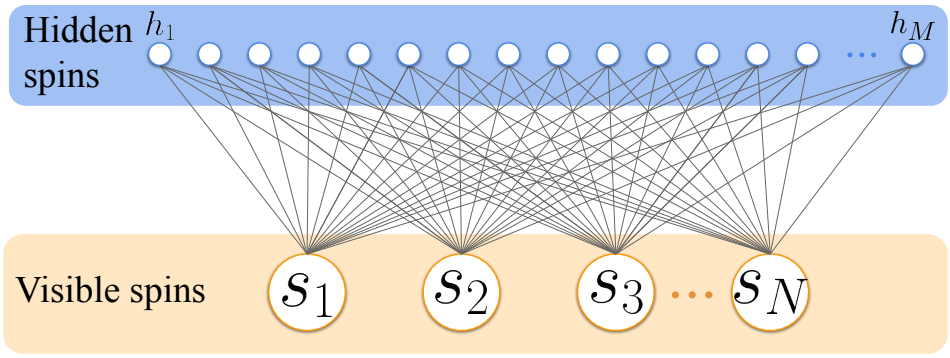


Figure 2.1: Visualization of the restricted Boltzmann Machine approach to one-dimensional spin chains, introduced in Ref. [CT17]. The variational ansatz is constructed by conceptually introducing a new layer of "hidden" spins in addition to "visible" spins forming the chain. See text for details.

Then, the aim, as in every variational approach, is to minimize the energy functional

$$E = \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle}. \tag{2.13}$$

This minimization inherently involves two subproblems:

1. calculating (estimating) the gradients of energy with respect to the variational parameters
2. minimizing the energy, at each step performing the gradient estimation procedure from point 1

Estimation of the gradients. However, as already mentioned in Chapter 1, due to the size of the Hilbert space growing exponentially with the number of spins involved, exact numerical calculation of the energy is infeasible. Therefore, one performs stochastic sampling of the system with Metropolis-Hastings algorithm, which is a Monte Carlo method commonly used to get stochastic estimates of energy and its gradients with respect to variational parameters.

Minimizing the energy After having obtained the gradient estimates, the problem to minimize energy is the standard minimization problem common in machine learning community. Therefore, one can take the full advantage of the plethora of well-developed techniques. The simplest approach is to simply update the weights by a small fraction of the gradients (stochastic gradient descent). More involved techniques include the Sorella method [SCR07a], or the modern techniques used in deep learning, such as momentum [Pol64] or Adam [KB14] optimization.

Many extensions of the original neural-network quantum states framework have been developed. Already in Ref. [CT17], the same ansatz was used to calculate dynamic properties of the system. Instead of the time-independent variational energy, the following quantity was minimized:

$$\mathrm{dist}\left(i\hbar\frac{d\left|\psi(t)\right\rangle}{dt}, \hat{H}\left|\psi(t)\right\rangle\right), \tag{2.14}$$

where

$$\mathrm{dist}(\left|\varphi\right\rangle, \left|\theta\right\rangle) = \arccos\sqrt{\frac{\left\langle\varphi|\theta\right\rangle\left\langle\theta|\varphi\right\rangle}{\left\langle\varphi|\varphi\right\rangle\left\langle\theta|\theta\right\rangle}} \tag{2.15}$$

is the so-called Bures angle [Bur69]. This approach directly corresponds to minimizing the difference between the left-hand- and right-hand-side of the time-dependent Schrödinger equation:

$$i\hbar\frac{d\left|\psi(t)\right\rangle}{dt} = \hat{H}\left|\psi(t)\right\rangle. \tag{2.16}$$

Apart from the dynamics, the neural-network quantum states approach has been extended to open quantum systems [HC19, VBRC19, YH19, NS19]. Other extensions include bosonic [Sai17] and fermionic [NDYI17] systems, as well as sytems of coupled fermions and bosons [Nom20].

2.3 Artificial neural networks as a tool to characterize phase transitions

Physical applications of artificial neural networks stretch way beyond variational optimization of ground states. Another application that we will pursue in this Thesis is the reconstruction of phase diagrams from snapshots of the system. The term "snapshots" might for example refer to Monte Carlo snapshot of a classical lattice spin system or to snapshots of a continuous system obtained with molecural dynamics methods, or even to images of a physical system obtained in an experiment.

These snapshots can be treated like images for computational purposes. Compared to image (e.g. photograph) classification tasks from computer science, the physics problems are often simpler. For instance, a layer of spin-1/2 particles will only feature binary values, $S_z = 1/2$. Compared to color images, where the standard is to use 8-bit color depth, giving 256 different possibilities per colour ($256^3 = 2^24$ possibilities in total for an RGB image), this an extremely significant decrease in the size of all possible combinations. This often allows successful with much less computational power than normally needed in the field of image classification.

One can work in one of the two basic frameworks of machine learning:

- unsupervised learning, where the phase diagram is explored by the algorithm on its own, without being fed prior training information, such as Monte Carlo snapshots associated with labels,
- supervised learning, where labeled examples are provided, the model learns from them, and is able to do further classification.

One particular architecture of the artificial neural networks that is particularly well tailored to solve discrimination problems in computer vision are convolutional neural networks (CNNs). These networks contain one or more convolutional layers. These layers perform a convolution between their input and a set of trainable filters. We illustrate a convolutional layer in Fig. 2.2. In this figure we also show the size dependences between the input and output layers.

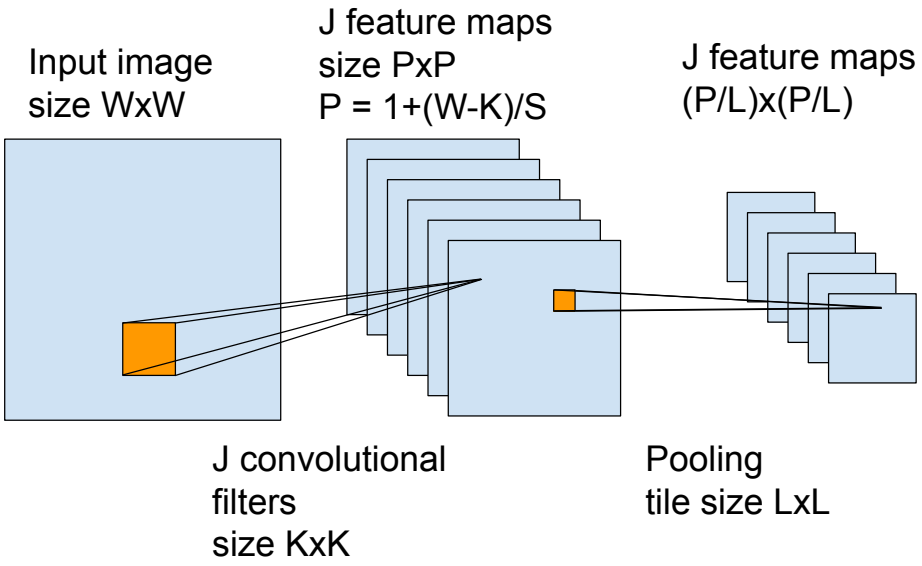


Figure 2.2: Visualization of a convolutional layer followed by a pooling layer. The input image is independently convolved with J convolutional filters with stride S, resulting in J feature maps. Each of them is then pooled, resulting in J new smaller maps. See details in the text.

Each filter convolves with the input image independently, thus resulting in multiple feature maps. Such approach allows the filters to learn different patterns in the input image. The key parameters of a CNN architecture are:

- filter size – this should be chosen such that it captures the expected size of features that one aims to detect. For example, if we expect order up to second-nearest neighbour in a square lattice physical system, it is reasonable to use a 5x5 filter.
- stride – the step size in the discrete convolution. Larger strides decrease the dimensionality of the feature maps, but might miss relevant features.
- number of filters – the larger the number of filters, the greater the representative power of the network. However, the risk of overfitting as well as memory and computational cost also grow with the number of filters.

Convolutional layer is usually followed by a pooling layer. This method of non-linear sampling has two main advantages. As a form of regularization, it prevents overfitting, a phenomenon where the network overly fits to training data and generalizes poorly to unseen data. Secondly, it reduces the feature map size, thus reducing both the memory and computational power needed. The most common pooling operation is the max pooling, where the tiles select the maximum value from their receptive field. We visualize the max-pooling layer in Fig. 2.3.

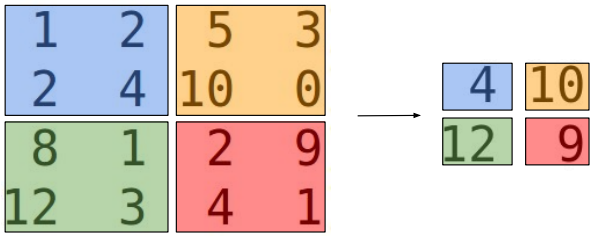


Figure 2.3: Detailed visualization of the max-pooling layer. The pool tiles of size 2x2 select the maximum from their receptive field.

Due to the nature of convolution operation, CNNs are perfectly tailored for learning local features of images. Their filters have a fixed size which then moves over the entire image. This has advantages both for image recognition as well as for physics. The filters can train to recognize local patterns, in image recognition these can be objects that we detect, in physics these can be local patterns of the lattice (for example antiferromagnetic "checkerboard" pattern).

The applications of artificial neural networks to phase recognition have been pioneered by Carrasquilla and Melko [CM17]. They have used CNNs in a supervised learning framework to distinguish between phases of Ising model. This approach has been later extended to strongly correlated fermionic systems [CCMK17, BCMT17]. After the successful supervised approaches, the more challenging unsupervised framework has also been explored, e.g. in [Wan16, WS17, HSS17]. A good review can be found in [Car20]. The developments made in this Thesis are presented in Chapter 5.

CHAPTER 3

Variational approach to spinful angulon

3.1 Chapter intro

As already discussed in Chapter 1, properties of quantum many-particle systems can be understood by considering their elementary building blocks – individual impurities coupled to a many-particle environment [Mah90, Wei12, BP02]. During the last years much effort has been focused on uncovering the physics associated with point-like impurities, possessing simple or no internal structure. Such impurity problems date back to the concept of polaron first introduced by Landau and Pekar [Lan33, Pek46, LP48].

The polaron, discussed in detail in Section 1.2.2, represents a quasiparticle consisting of an electron dressed by a cloud of crystal vibrations. It has become a standard tool to describe transport phenomena in solid state and chemical physics [Emi13, Dev16]. Recently, controllable polarons have been realized in ultracold quantum gases of bosons and fermions [MZB14, JWS$^+$16]. Another broad class of well-studied impurity problems involves a localized spin coupled to a bath of bosons [LCD$^+$87], fermions [LCND08], or other spins [PS00].

In many settings, however, quantum impurities possess additional internal degrees of freedom, such as orbital or rotational angular momentum. Such problems arise, e.g. in the context of molecules rotating in superfluid helium [TV04], ultracold alkali dimers interacting with a Bose-Einstein condensate (BEC) [JY12], or electrons whose orbital angular momentum is coupled to the crystal lattice [KRLdJ05, CGS05, ZN14, TP15, GC15a, TIH$^+$16].

Section 1.2.3 introduces the 'angulon' quasiparticle [SL15, SL16, LS17], a convenient tool to address the problems mentioned in the previous paragraph. The angulon forms out of an impurity exchanging rotational angular momentum with a many-particle bath of some sort; it can be thought of as a quantum rotor dressed in a coat of orbital bath excitations. In a way, the angulon represents a rotational counterpart of the polaron, however, the non-abelian algebra of quantum rotations and their discrete spectrum render the angulon physics remarkably different.

The angulon theory has been tested against experiments on molecules trapped in superfluid helium nanodroplets. There, it was observed that the effective moment of inertia

increases for molecules immersed in superfluid helium, as compared to free species [TV04]. This phenomenon is somewhat similar to renormalization of the effective mass of electrons interacting with a crystal lattice [Dev16]. It was recently shown that the angulon theory can reproduce the effective moments of inertia for molecules in helium nanodroplets for a broad range of species, both in the weak-coupling and strong-coupling regimes [Lem17]. Moreover, a coherent non-adiabatic rotation of angulons formed out of I_2 molecules has been experimentally demonstrated [SSC$^+$17]. The angulon theory thereby offers an alternative approach to molecules in quantum solvents, along with established numerical techniques based on quantum Monte Carlo and Density Functional Theory calculations [Sza08b, RCGLV16, ABC$^+$17].

In this Chapter we generalize the angulon theory to the case where the impurity possesses both rotational and spin–1/2 degrees of freedom and is exposed to a static magnetic field. Our main focus will be on open-shell diatomic molecules rotating in quantum solvents. It is important to emphasize that the angulon model (including Eqs. (3.20) and (3.19) below) has been originally derived for an ultracold molecule immersed in a weakly-interacting BEC, where the theory is expected to provide quantitatively accurate predictions. It has been shown, however, that one can approach the angulon Hamiltonian from a phenomenological perspective in order to describe the properties of molecules in superfluid ^{4}He, in good agreement with experiment [Lem17, CL17, SSC$^+$17]. Thus, while the theory is not designed to compete with numerical Monte Carlo calculations in accuracy, it is expected to provide qualitatively accurate predictions for molecules in liquid helium along with simple explanations for the underlying physics. Furthermore, the theory can be in principle generalized to other types of orbital impurities such as polyatomic molecules, non-spherical paramagnetic atoms, or $p-$, $d-$, or $f-$electrons.

In Section 3.2 we derive the extended angulon Hamiltonian, which includes the impurity spin and the impurity-field interaction. In Section 3.3 we derive a Dyson equation for the spinful angulon, which allows us to calculate its self-energy and spectral function. In Section 3.4 we analyze the angulon spectral function and the way it changes in a magnetic field, for various bath densities. Section 3.5 focuses on the angulon instabilities which result in resonant emission of phonons with a given value of angular momentum. In particular, we reveal the possibility to manipulate the angular momentum of phonons using a magnetic field. In Section 3.6 we study a limiting case when the magnetic field tends to zero, thus discussing the spinful angulon in the absence of external fields.

3.1.1 Authors' contributions

ML and I conceived the project. I performed the calculations under regular supervision of ML.

3.2 Hamiltonian of the spinful angulon in a magnetic field

We consider a molecular impurity with spin–1/2 and orbital angular momentum, immersed in a bosonic bath. In the presence of a magnetic field, the system can be described by the following Hamiltonian:

$$\widehat{H} = \widehat{H}_{\text{mol}} + \widehat{H}_{\text{mol-f}} + \widehat{H}_{\text{bos}} + \widehat{H}_{\text{mol-bos}}, \tag{3.1}$$

where the terms correspond to the bare molecule, molecule-field interaction, bosonic bath, and molecule-boson interaction, respectively. Note that we assume a neutral, spinless bath, such that its direct interaction with the magnetic field can be neglected. In the following subsections we describe each term of Eq. (3.1) in detail. In what follows, we use the units where $\hbar \equiv 1$.

3.2.1 Bare spinful molecule

The first term of Eq. (3.1) corresponds to a linear molecule with spin–1/2 ($^2\Sigma$ electronic state), as given by the following Hamiltonian:

$$\widehat{H}_{\text{mol}} = B\hat{\vec{L}}^2 + \gamma \hat{\vec{L}} \cdot \hat{\vec{S}}. \tag{3.2}$$

Here $B = 1/(2I)$ is the rotational constant with I being the moment of inertia, and γ defines the spin-rotation coupling [LBF04].

Since we focus on spin–1/2 molecules, in Eq. (3.2) we omit the constant shift proportional to $\hat{\vec{S}}^2$. For higher spins (as, e.g., in $^3\Sigma$ molecules), the spin-spin interaction will lead to an additional term in $\widehat{H}_{\text{mol}}$. Furthermore, here we consider an impurity whose translational motion is frozen in space, which is a good approximation for molecules in helium nanodroplets [TV04]. Our formalism, however, can be generalized to include the above mentioned terms, as well as to treat more complex impurities, such as polyatomic molecules [Ber05].

We denote the eigenstates of the bare molecular Hamiltonian (3.2) as $|J = L \pm 1/2, L, M_J\rangle$ with the corresponding eigenenergies

$$E^0_{JL} = BL(L+1) + \gamma \left[(J-L)\left(L + \frac{1}{2}\right) - \frac{1}{4} \right]. \tag{3.3}$$

Here J is the total (rotation+spin) angular momentum, L is the total rotational angular momentum and M_J is the projection of J on the quantization axis. All three numbers are good quantum numbers. Note that the eigenstates can be written in the uncoupled basis as

$$\begin{aligned} &\left| J = L \pm \frac{1}{2}, L, M_J \right\rangle = \\ &C^{J,M_J}_{L,M_J-\frac{1}{2};\frac{1}{2},\frac{1}{2}} \left| L, M_J - 1/2 \right\rangle \left| \frac{1}{2}, \frac{1}{2} \right\rangle + C^{J,M_J}_{L,M_J+\frac{1}{2};\frac{1}{2},-\frac{1}{2}} \left| L, M_J + 1/2 \right\rangle \left| \frac{1}{2}, -\frac{1}{2} \right\rangle. \end{aligned} \tag{3.4}$$

3.2.2 Molecule-field interaction

The term $\widehat{H}_{\text{mol-f}}$ describes a static magnetic field applied to the system:

$$\widehat{H}_{\text{mol-f}} = B\eta \hat{S}_Z, \tag{3.5}$$

where the dimensionless molecule-field interaction parameter is given by:

$$\eta = \frac{\mathscr{H} g_s \mu_B}{B}. \tag{3.6}$$

Here $g_s \approx 2.0023$ is the gyromagnetic ratio, μ_B is the Bohr magneton, and $\mathscr{H}$ gives the magnitude of the magnetic field.

Assuming that the magnetic field affects only the spin degree of freedom (i.e. neglecting rotational magnetism in the case of molecules [LBF04]), the operator (3.5) couples only levels with the same L. In such a case, the eigenstates of the molecule+field Hamiltonian, $\widehat{H}_{\text{mol}} + \widehat{H}_{\text{mol-f}}$, represent field-dependent superpositions of the states $|J, L, M_J\rangle$ [FH00]:

$$\begin{aligned}\left|\tilde{J}, L, M_J\right\rangle =& a_{\tilde{J}LM_J}(\eta)\,|J = L + 1/2, L, M_J\rangle + \\ &+ b_{\tilde{J}LM_J}(\eta)\,|J = L - 1/2, L, M_J\rangle\,.\end{aligned} \tag{3.7}$$

In the presence of a field, L and M_J are good quantum numbers, while J is not. However, in Eq. (3.7) we use $\tilde{J}$ as an adiabatic (approximately good) quantum number, such that

$$\left|\tilde{J}, L, M_J\right\rangle \underset{\eta\to 0}{\longrightarrow} |J, L, M_J\rangle\,. \tag{3.8}$$

The exact form of the $a_{\tilde{J}LM_J}(\eta)$ and $b_{\tilde{J}LM_J}(\eta)$ coefficients of Eq. (3.7) can be calculated as follows. For given L, M_J, the molecular and magnetic part of the Hamiltonian in the matrix form reads:

$$\widehat{H}_{\text{mol}} + \widehat{H}_{\text{mol-f}} = \begin{bmatrix} BL(L+1) + \frac{L\gamma}{2} + \frac{B\eta M_J}{2\left(L+\frac{1}{2}\right)} & -\frac{B\eta}{2}\sqrt{1 - \frac{M_J^2}{\left(L+\frac{1}{2}\right)^2}} \\ -\frac{B\eta}{2}\sqrt{1 - \frac{M_J^2}{\left(L+\frac{1}{2}\right)^2}} & BL(L+1) - \frac{\gamma}{2}(L+1) - \frac{B\eta M_J}{2\left(L+\frac{1}{2}\right)} \end{bmatrix}. \tag{3.9}$$

Upon diagonalization we find eigenenergies as given by Eq. (3.13) and corresponding eigenstate coefficients given by:

$$a_{\tilde{J}LM_J}(\eta) = \frac{-\phi_1 + 2(\tilde{J} - L)\sqrt{\phi_1^2 + \phi_2^2}}{\sqrt{\phi_2^2 + (-\phi_1 + 2(\tilde{J} - L)\sqrt{\phi_1^2 + \phi_2^2})^2}}, \tag{3.10}$$

and

$$b_{\tilde{J}LM_J}(\eta) = \frac{\phi_2}{\sqrt{\phi_2^2 + (-\phi_1 + 2(\tilde{J} - L)\sqrt{\phi_1^2 + \phi_2^2})^2}}, \tag{3.11}$$

where

$$\phi_1 = \frac{B\eta M_J}{L + \frac{1}{2}} + \gamma\left(L + \frac{1}{2}\right), \quad \phi_2 = B\eta\sqrt{1 - \frac{M_J^2}{\left(L + \frac{1}{2}\right)^2}}. \tag{3.12}$$

The eigenenergies of the $\left|\tilde{J}, L, M_J\right\rangle$ states are given by [FH00]:

$$E^0_{\tilde{J}LM_J} = BL(L+1) - \frac{\gamma}{4} + (\tilde{J} - L)\gamma(L + \frac{1}{2})\xi_{BLM_J}, \tag{3.13}$$

where

$$\xi_{BLM_J} = (1 + 2\alpha_{LM_J}X_{BL} + X_{BL}^2)^{\frac{1}{2}}, \tag{3.14}$$

with

$$\alpha_{LM_J} = \frac{M_J}{L + 1/2},\ X_{BL} = \frac{B\eta}{\gamma(L + 1/2)}. \tag{3.15}$$

3.2.3 Bosonic bath energy

The term $\widehat{H}_{\text{bos}}$ corresponds to the kinetic energy of the bosonic excitations in a quantum solvent, such as phonons, rotons, and ripplons in superfluid ^{4}He [SL06]. In its diagonal form, the bosonic bath Hamiltonian reads:

$$\widehat{H}_{\text{bos}} = \sum_{k\lambda\mu} \omega_k \hat{b}^\dagger_{k\lambda\mu} \hat{b}_{k\lambda\mu}, \tag{3.16}$$

where ω_k is the dispersion relation. The creation and annihilation operators of Eq. (3.16) are conveniently expressed in the angular momentum basis [LS17]:

$$\hat{b}^\dagger_{k\lambda\mu} = \frac{k}{(2\pi)^{3/2}} \int d\Omega_k \; i^{-\lambda} \; Y_{\lambda\mu}(\Omega_k) \; \hat{b}^\dagger_{\mathbf{k}}, \tag{3.17}$$

$$\hat{b}_{k\lambda\mu} = \frac{k}{(2\pi)^{3/2}} \int d\Omega_k \; i^{\lambda} \; Y^*_{\lambda\mu}(\Omega_k) \; \hat{b}_{\mathbf{k}}. \tag{3.18}$$

Here $\hat{b}^\dagger_{\mathbf{k}}$ and $\hat{b}_{\mathbf{k}}$ are the creation and annihilation operators defined in Cartesian space, and $Y_{\lambda\mu}(\Omega_k) \equiv Y_{\lambda\mu}(\theta_k, \varphi_k)$ are the spherical harmonics [VMK88]. The quantum numbers $k = |\mathbf{k}|$, λ, and μ, label, respectively, the linear momentum of phonons, the angular momentum of phonons, and the projection of the phonon angular momentum onto the laboratory-frame z-axis.

In Eq. (3.16), the form of the dispersion relation ω_k depends on the particular system under consideration. Here, without loss of generality, we chose the dispersion relation corresponding to Bogoliubov excitations in a weakly-interacting BEC [PS16]:

$$\omega_k = \sqrt{\epsilon_k(\epsilon_k + 2g_{bb}n)}. \tag{3.19}$$

Here $\epsilon_k = k^2/(2m)$ is the boson kinetic energy and $g_{bb} = 4\pi a_{bb}/m$ parametrizes the interactions between the bosons of mass m, where a_{bb} gives the boson-boson scattering length. While the Bogoliubov dispersion (3.19) does not provide a quantitatively good approximation to the properties of superfluid helium, in the regime of large values of g_{bb} and n it qualitatively describes the properties of a dense superfluid for small momenta k. Furthermore, the theory can be extended to other types of excitations, such as rotons or lattice phonons.

3.2.4 Molecule-boson interaction

The last term of Eq. (3.1) determines the interaction between the molecule and the bosonic bath, as given by [SL15]:

$$\widehat{H}_{\text{mol-bos}} = \sum_{k\lambda\mu} U_\lambda(k) \; [Y^*_{\lambda\mu}(\hat{\theta}, \hat{\phi}) \hat{b}^\dagger_{k\lambda\mu} + Y_{\lambda\mu}(\hat{\theta}, \hat{\phi}) \hat{b}_{k\lambda\mu}]. \tag{3.20}$$

As already outlined in Section 1.2.3, the form of this term stems from expanding the Hamiltonian in fluctuations around a homogeneous BEC of density n and applying the Bogoliubov approximation and transformation (a constant mean-field shift is omitted). In such a case the Fourier-space interaction potentials, $U_\lambda(k)$, can be obtained in closed form [SL15]:

$$U_\lambda(k) = \sqrt{\frac{8nk^2\epsilon_k}{\omega_k(2\lambda+1)}} \int dr r^2 V_\lambda(r) j_\lambda(kr), \tag{3.21}$$

where $j_\lambda(kr)$ is the spherical Bessel function, and $V_\lambda(r)$ give the Legendre moments of the two-body interaction between the molecule and an atom from the BEC in the molecular frame:

$$V_{\text{mol-at}}(\mathbf{r}) = \sum_\lambda V_\lambda(r) Y_{\lambda 0}(\theta_r, \phi_r). \tag{3.22}$$

The spherical harmonic operators, $Y_{\lambda\mu}(\hat{\theta}, \hat{\phi})$, in Eq. (3.20) arise due to rotation of the molecule-atom interaction potential from the molecular to the laboratory frame. The explicit dependence on the molecule angle operators, $(\hat{\theta}, \hat{\phi})$, makes the angulon problem substantially different from other impurity problems such as the Bose polaron [Dev16] and spin-boson [LCD$^+$87] models. It is important to note that while we consider the closed-form coupling of Eqs. (3.20) and (3.21) for simplicity, we expect the model to provide qualitative predictions beyond the range of applicability of the Bogoliubov approximation. For other types of impurities, such as electrons or non-spherical atoms, $U_\lambda(k)$ will assume a different form. Furthermore, the coupling constants $U_\lambda(k)$ are taken to be independent on the μ quantum number, which is the case, e.g. for linear molecular impurities [SL15, LS17]. Treating more complex, nonlinear molecules requires μ-dependent potentials [CL17]. However, the microscopic details of the impurity-bath interaction are not expected to alter the effects discussed in this Chapter qualitatively. Therefore, in what follows we use the coupling given by Eq. (3.21).

Note that the long-wavelength behavior of Eq. (3.21) is given by:

$$U_\lambda(k \to 0) \approx \zeta k^{\lambda+3/2} + \mathcal{O}(k^{\lambda+7/2}), \tag{3.23}$$

where ζ is a constant independent of λ. The term $\zeta k^{\lambda+3/2}$ contributes to the rise of $U_\lambda(k)$ for small values of k, which is a consequence of the centrifugal barrier emerging for collisions with finite angular momentum. This behavior is illustrated in Fig. 3.1, where $U_\lambda(k)$ is plotted for $\lambda = 0$ and $\lambda = 1$ at several densities.

3.3 The angulon self energy and spectral function

In order to uncover the behavior of spinful angulons in a magnetic field, we make use of the equivalence between the variational and diagrammatic approaches to the angulon problem, see Refs. [SL15, LS17] for details. We start from the variational ansatz constructed of field-dependent molecular states and taking into account single-phonon excitations:

$$\begin{aligned}\left|\psi_{\tilde{J}LM_J}\right\rangle = Z^{1/2}_{\tilde{J}LM_J} \left|0\right\rangle \left|\tilde{J}, L, M_J\right\rangle \\ + \sum_{\substack{k\lambda \\ \tilde{j}lm_j\mu}} \beta_{\lambda\tilde{j}l}(k) C^{\tilde{J},M_J}_{\tilde{j},m_j;\lambda,\mu} \hat{b}^\dagger_{k\lambda\mu} \left|0\right\rangle \left|\tilde{j}, l, m_j\right\rangle,\end{aligned} \tag{3.24}$$

where $|0\rangle$ is the vacuum of bosonic excitations, and $Z^{1/2}_{\tilde{J}LM_J}$ and $\beta_{\lambda\tilde{j}l}(k)$ are the variational parameters obeying the following normalization condition:

$$Z_{\tilde{J}LM_J} + \sum_{k\lambda\tilde{j}l} |\beta_{\lambda\tilde{j}l}(k)|^2 = 1. \tag{3.25}$$

Note that, despite the presence of the field, our variational coefficients are independent of m_j, which comes from the fact that the interaction potentials $U_\lambda(k)$ are independent

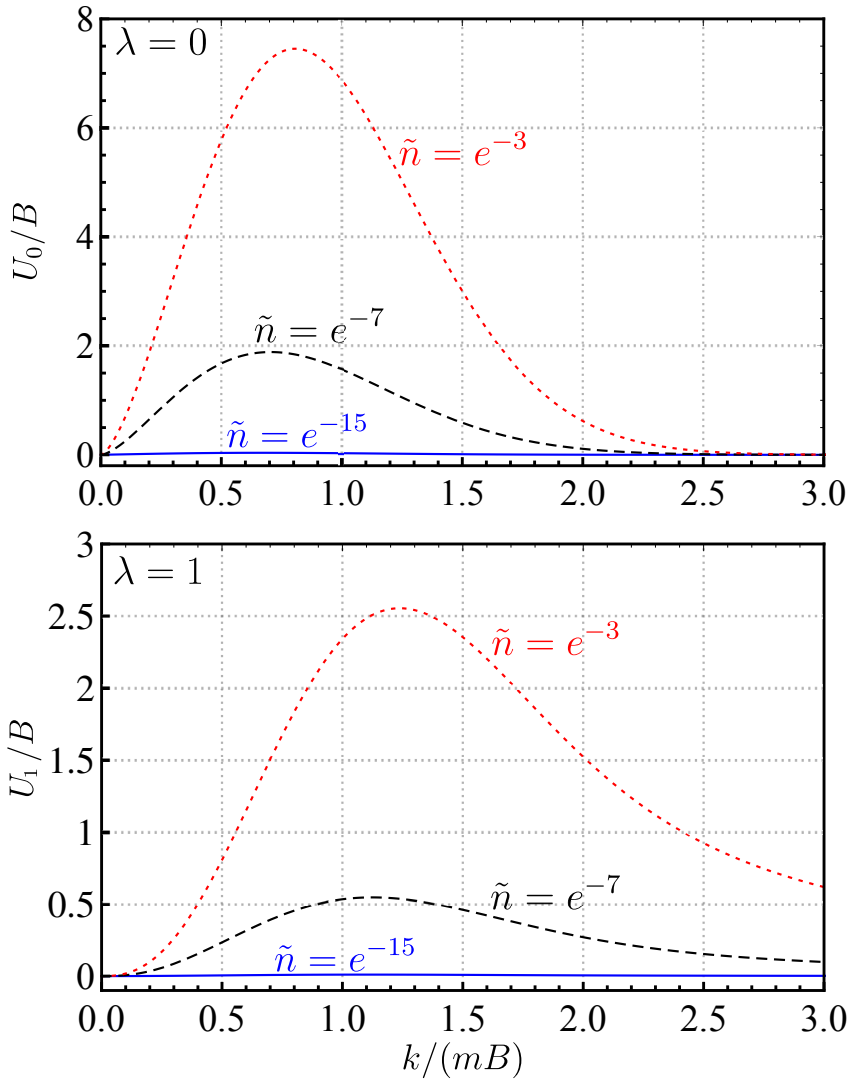


Figure 3.1: Dependence of the molecule-boson couplings U_0 and U_1 on k, for the parameters defined in Section 3.4. Each coefficient $U_\lambda(k)$ is shown at three different densities: $\tilde{n} = \exp(-15)$, $\tilde{n} = \exp(-7)$, $\tilde{n} = \exp(-3)$. Reproduced from Ref. [RL18], with the permission of AIP Publishing.

of μ. Moreover, in the presence of a field, M_J is the only good quantum number of the system. In the variational ansatz of Eq. (3.24), conservation of M_J is accounted for by the Clebsch-Gordan coefficient $C^{\tilde{J},M_J}_{\tilde{j},m_j;\lambda,\mu}$ [VMK88]. The number $\tilde{J}$, despite its presence in the Clebsch-Gordan coefficients is not a good quantum number since the ansatz is constructed on top of states which are not eigenstates of $\hat{\mathbf{J}}^2$. This means that in Eq. (3.24) we neglect the processes where $\tilde{J}$ changes due to the molecule-bath interactions, which is a good approximation away from crossings of levels with different $\tilde{J}$. This approximation becomes exact in the limit of $\eta \to 0$, where J is a good quantum number.

In Refs. [SL15, LS17] it has been shown, that by minimizing the functional

$$F = \left\langle \psi_{\tilde{J}LM_J} \middle| \widehat{H} - E \middle| \psi_{\tilde{J}LM_J} \right\rangle \tag{3.26}$$

over $Z_{\tilde{J}LM_J}$ and $\beta_{\lambda\tilde{j}l}(k)$, one can derive the following Dyson equation for the angulon:

$$G^{\mathrm{ang}}_{\tilde{J}LM_J}(E)^{-1} = G^{0}_{\tilde{J}LM_J}(E)^{-1} - \Sigma_{\tilde{J}LM_J}(E), \tag{3.27}$$

where $G^0_{\tilde{J}LM_J}(E)$ is the free molecule Green's function

$$G^0_{\tilde{J}LM_J}(E) = \frac{1}{E^0_{\tilde{J}LM_J} - E}, \tag{3.28}$$

and

$$\Sigma_{\tilde{J}LM_J}(E) = \sum_{\substack{k\lambda \\ \tilde{j}lm_j m'_j}} \frac{U_\lambda(k)^2 K^{\tilde{J}LM_J}_{\tilde{j}lm_j\lambda}(\eta) K^{\tilde{J}LM_J}_{\tilde{j}lm'_j\lambda}(\eta)}{\omega_k + \sum_{m''_j}(C^{\tilde{J},M_J}_{\tilde{j},m''_j;\lambda,M_J-m''_j})^2 E^0_{\tilde{j}lm''_j} - E} \tag{3.29}$$

is the angulon self energy. Here $K^{\tilde{J}LM_J}_{\tilde{j}lm_j\lambda}(\eta)$ is a coefficient resulting from the relevant angular momentum algebra, dependent upon Clebsch-Gordan coefficients, $a_{\tilde{J}LM_J}(\eta)$, and $b_{\tilde{J}LM_J}(\eta)$.

Let us now derive Eq. (3.27) and provide explicit form of the coefficients $K^{\tilde{J}LM_J}_{\tilde{j}lm_j\lambda}(\eta)$. We denote the terms of the ansatz from Eq. (3.24) as $\left|\psi_{\tilde{J}LM_J}\right\rangle = \left|\psi^1_{\tilde{J}LM_J}\right\rangle + \left|\psi^2_{\tilde{J}LM_J}\right\rangle$, where:

$$\left|\psi^1_{\tilde{J}LM_J}\right\rangle = Z^{1/2}_{\tilde{J}LM_J} \left|0\right\rangle \left|\tilde{J}, L, M_J\right\rangle, \tag{3.30}$$

and

$$\left|\psi^2_{\tilde{J}LM_J}\right\rangle = \sum_{\substack{k\lambda \\ \tilde{j}lm_j\mu}} \beta_{\lambda\tilde{j}l}(k) C^{\tilde{J},M_J}_{\tilde{j},m_j;\lambda,\mu} \hat{b}^\dagger_{k\lambda\mu} \left|0\right\rangle \left|\tilde{j}, l, m_j\right\rangle. \tag{3.31}$$

We minimize the functional

$$\begin{aligned} F &= \left\langle \psi_{\tilde{J}LM_J} \right| \widehat{H} \left| \psi_{\tilde{J}LM_J} \right\rangle - E \left\langle \psi_{\tilde{J}LM_J} \middle| \psi_{\tilde{J}LM_J} \right\rangle = \\ &= \left\langle \psi^1_{\tilde{J}LM_J} \right| \widehat{H} \left| \psi^1_{\tilde{J}LM_J} \right\rangle + \left\langle \psi^2_{\tilde{J}LM_J} \right| \widehat{H} \left| \psi^2_{\tilde{J}LM_J} \right\rangle + \\ &+ \left[\left\langle \psi^1_{\tilde{J}LM_J} \right| \widehat{H} \left| \psi^2_{\tilde{J}LM_J} \right\rangle + c.c. \right] - E \left\langle \psi_{\tilde{J}LM_J} \middle| \psi_{\tilde{J}LM_J} \right\rangle. \end{aligned} \tag{3.32}$$

To calculate F, we sequentially evaluate its terms. First,

$$\left\langle \psi_{\tilde{J}LM_J} \middle| \psi_{\tilde{J}LM_J} \right\rangle = |Z_{\tilde{J}LM_J}| + \sum_{\substack{k\lambda \\ \tilde{j}l}} |\beta_{\lambda\tilde{j}l}(k)|^2, \tag{3.33}$$

where we summed over m_j and μ using the orthogonality relations for Clebsch-Gordan coefficients [VMK88]. Then

$$\left\langle \psi^1_{\tilde{J}LM_J} \right| \widehat{H} \left| \psi^1_{\tilde{J}LM_J} \right\rangle = |Z_{\tilde{J}LM_J}| E^0_{\tilde{J}LM_J}, \tag{3.34}$$

and

$$\left\langle \psi^2_{\tilde{J}LM_J} \right| \widehat{H} \left| \psi^2_{\tilde{J}LM_J} \right\rangle = \sum_{k\lambda\tilde{j}l} |\beta_{\lambda\tilde{j}l}(k)|^2 \left[\sum_{m_j} (C^{\tilde{J},M_J}_{\tilde{j},m_j;\lambda,M_J-m_j})^2 E^0_{\tilde{j}lm_j} + \omega_k \right]. \tag{3.35}$$

The contribution of the interaction (fourth term of Eq. (3.1)) vanishes in Eqs. (3.33), (3.34), (3.35), unlike the $\left\langle \psi^1_{\tilde{J}LM_J} \right| \widehat{H} \left| \psi^2_{\tilde{J}LM_J} \right\rangle$ term, which is nonzero solely due to the interaction. We express the field-dependent states in the basis of Eq. (3.7), and decompose these states further as in Eq. (3.4). Now we can act with the spherical harmonics operators on the rotational kets, with the matrix elements given by [VMK88, LS17]:

$$\left\langle L', M' \right| Y_{\lambda\mu}(\hat{\theta}, \hat{\phi}) \left| L, M \right\rangle = \sqrt{\frac{(2L+1)(2\lambda+1)}{4\pi(2L'+1)}} C^{L'M'}_{LM;\lambda\mu} C^{L'0}_{L0;\lambda0}. \tag{3.36}$$

Then, making use of the spin kets orthogonality we arrive at the final form of $\left\langle \psi^1_{\tilde{J}LM_J} \right| \widehat{H} \left| \psi^2_{\tilde{J}LM_J} \right\rangle$:

$$\left\langle \psi^1_{\tilde{J}LM_J} \right| \widehat{H} \left| \psi^2_{\tilde{J}LM_J} \right\rangle = (Z^*_{\tilde{J}LM_J})^{1/2} \sum_{k\lambda\tilde{j}lm_j} \beta_{\lambda\tilde{j}l}(k) U_\lambda(k) K^{\tilde{J}LM_J}_{\tilde{j}lm_j\lambda}(\eta), \tag{3.37}$$

where we denoted

$$\begin{aligned} K^{\tilde{J}LM_J}_{\tilde{j}lm_j\lambda}(\eta) = C^{\tilde{J},M_J}_{\tilde{j},m_j;\lambda,M_J-m_j} \sqrt{\frac{(2\lambda+1)}{4\pi}} (-1)^\lambda C^{l0}_{L0\lambda0} \times \\ \times \Big\{ \left[C^{L+\frac{1}{2},M_J}_{L,M_J-\frac{1}{2};\frac{1}{2},\frac{1}{2}} a_{\tilde{J}LM_J}(\eta) + C^{L-\frac{1}{2},M_J}_{L,M_J-\frac{1}{2};\frac{1}{2},\frac{1}{2}} b_{\tilde{J}LM_J}(\eta) \right] \left[C^{l+\frac{1}{2},m_j}_{l,m_j-\frac{1}{2};\frac{1}{2},\frac{1}{2}} a_{\tilde{j}lm_j}(\eta) + C^{l-\frac{1}{2},m_j}_{l,m_j-\frac{1}{2};\frac{1}{2},\frac{1}{2}} b_{\tilde{j}lm_j}(\eta) \right] C^{L,M_J-\frac{1}{2}}_{l,m_j-\frac{1}{2};\lambda,M_J-m_j} + \\ \left[C^{L+\frac{1}{2},M_J}_{L,M_J+\frac{1}{2};\frac{1}{2},-\frac{1}{2}} a_{\tilde{J}LM_J}(\eta) + C^{L-\frac{1}{2},M_J}_{L,M_J+\frac{1}{2};\frac{1}{2},-\frac{1}{2}} b_{\tilde{J}LM_J}(\eta) \right] \left[C^{l+\frac{1}{2},m_j}_{l,m_j+\frac{1}{2};\frac{1}{2},-\frac{1}{2}} a_{\tilde{j}lm_j}(\eta) + C^{l-\frac{1}{2},m_j}_{l,m_j+\frac{1}{2};\frac{1}{2},-\frac{1}{2}} b_{\tilde{j}lm_j}(\eta) \right] C^{L,M_j+\frac{1}{2}}_{l,m_j+\frac{1}{2};\lambda,M_J-m_j} \Big\}. \end{aligned} \tag{3.38}$$

with $a_{\tilde{J}LM_J}(\eta)$ and $b_{\tilde{J}LM_J}(\eta)$ given by Eq. (3.10) and Eq. (3.11), respectively.

Having calculated the functional F, we can now compute its derivatives with respect to the variational parameters:

$$\frac{\partial F}{\partial (Z^*_{\tilde{J}LM_J})^{1/2}} = (Z_{\tilde{J}LM_J})^{1/2}(E^0_{\tilde{J}LM_J} - E) + \sum_{k\lambda\tilde{j}lm_j} \beta_{\lambda\tilde{j}l}(k) U_\lambda(k) K^{\tilde{J}LM_J}_{\tilde{j}lm_j\lambda} \equiv 0, \tag{3.39}$$

$$\begin{aligned} &\frac{\partial F}{\partial \beta^*_{\lambda\tilde{j}l}(k)} = \\ &= \beta_{\lambda\tilde{j}l}(k) \left[\sum_{m_j} (C^{\tilde{J},M_J}_{\tilde{j},m_j;\lambda,M_J-m_j})^2 E^0_{\tilde{j}lm_j} - E + \omega_k \right] + (Z_{\tilde{J}LM_J})^{1/2} U_\lambda(k) \sum_{m_j} K^{\tilde{J}LM_J}_{\tilde{j}lm_j\lambda} \equiv 0. \end{aligned} \tag{3.40}$$

Then, upon substitution of $\beta_{\lambda\tilde{j}l}(k)$ from Eq. (3.40) into Eq. (3.39), $Z^{1/2}_{\tilde{J}LM_J}$ cancels out and we arrive at the Dyson equation as given by Eq. (3.41).

Furthermore, in the field-free limit, $\eta \to 0$, the expression for the self-energy (3.29) can be simplified, see Section 3.6 for a detailed derivation.

By casting the variational problem in terms of the Dyson equation (3.27), we are able to access the energies of the excited states of the system by solving the following equation:

$$E = E^0_{\tilde{J}LM_J} - \Sigma_{\tilde{J}LM_J}(E), \tag{3.41}$$

as well as to calculate the spectral function of the angulon, which is defined as:

$$\mathcal{A}_{\tilde{J}LM_J}(E) = \mathrm{Im}[G^{\,\mathrm{ang}}_{\tilde{J}LM_J}(E + i0^+)]. \tag{3.42}$$

3.4 Zeeman effect for the angulon

In this section we study the angulon spectral function, Eq. (3.42), in the presence of a magnetic field. In order to describe the effects quantitatively, we use the Gaussian-shaped potentials

$$V_\lambda(r) = u_\lambda (2\pi)^{-3/2} e^{-r^2/(2r_\lambda^2)} \tag{3.43}$$

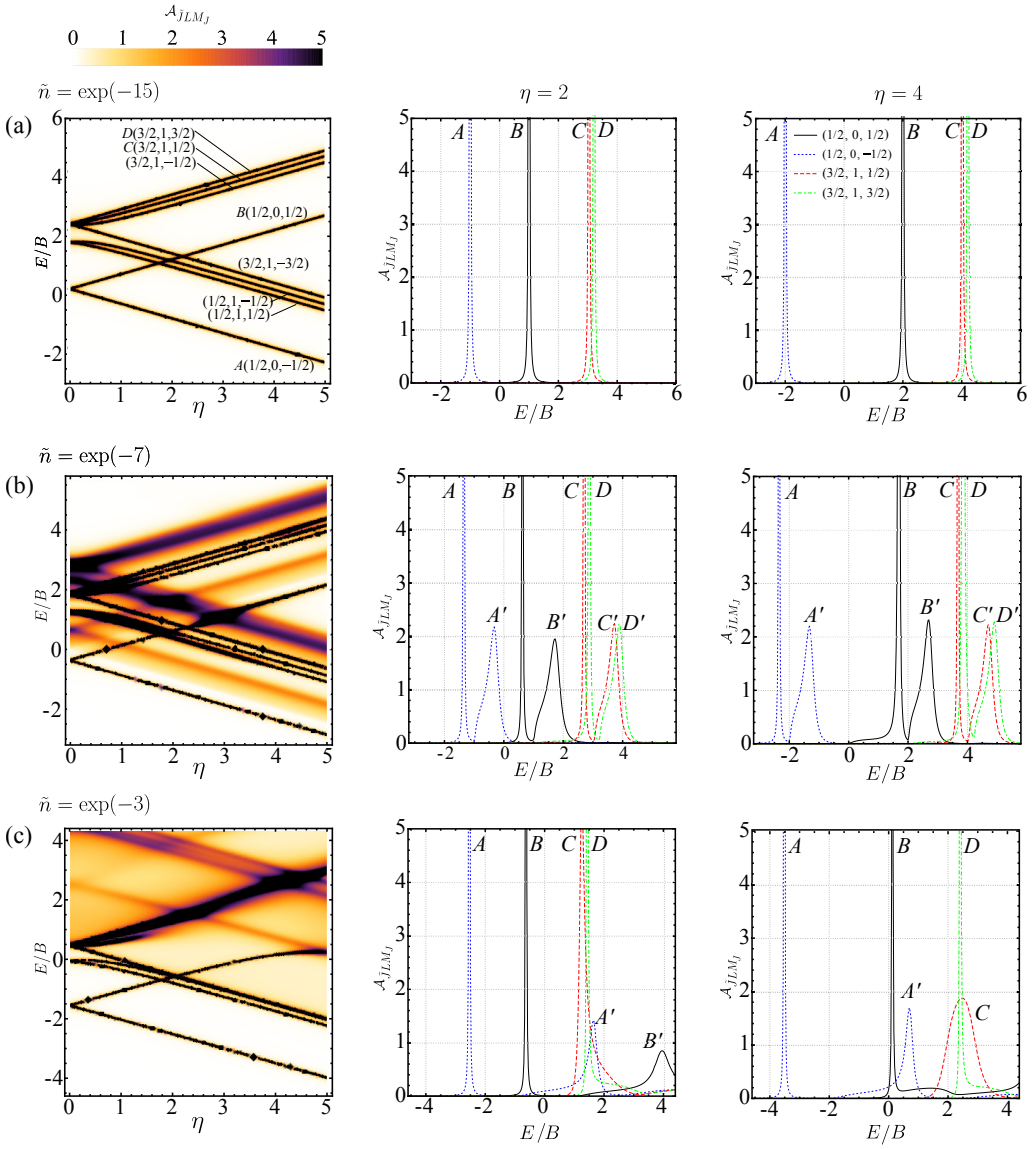


Figure 3.2: The angulon spectral function, $\mathcal{A}_{\tilde{J}LM_J}(E)$, for different dimensionless densities: (a) $\tilde{n} = \exp(-15)$, which approximately corresponds to bare molecular states in a magnetic field as in Ref. [FH00], (b) $\tilde{n} = \exp(-7)$, and (c) $\tilde{n} = \exp(-3)$. In the left column, the spectral functions are shown as a density plot, for all eight sublevels of $L = 0$ and $L = 1$, as a function of the magnetic field strength, η, and energy, E/B (these states are shown one by one in Fig. 3.3). The middle and right columns show the energy dependence of the spectral function for two selected values of $\eta = 2$ and $\eta = 4$ depicted for four states (1/2,0,-1/2), (1/2,0,1/2), (3/2,1,1/2), and (3/2,1,3/2) labeled as A, B, C, and D, respectively. The primed letters A', B', C', and D' mark additional fine structure emerging due to the interaction with the bath (see text). Reproduced from Ref. [RL18], with the permission of AIP Publishing.

and the following parameter values in dimensionless units: $r_0 = r_1 = 1.5(mB)^{-1/2}$ and $a_{bb} = 3.3(mB)^{-1/2}$, $u_0 = 1.75u_1 = 218B$, as previously used in Ref. [SL15]. Our choice of spin-rotation coupling is $\gamma = 0.418B$, which is the value used in Ref. [FH00].

In what follows, we focus on the substates belonging to $L = 0$ and $L = 1$ manifolds of rotational angular momentum. Fig. 3.2 shows the dependence of the angulon spectral

function on the field-strength parameter, η. Fig. 3.2(a) corresponds to a vanishingly small density of the bath, and therefore reproduces the structure of bare molecular states in a magnetic field [FH00, LF09].

The $L = 0$ and $L = 1$ manifolds of rotational angular momentum contain eight bare molecular Zeeman levels $(\tilde{J}, L, M_J)$ with their parity given by $P = (-1)^L$. As the magnetic field couples only levels with the same L, the parity remains unchanged in the presence of the field. For extreme values of the projection M_J, i.e. $M_J = \pm\tilde{J}$ with $\tilde{J} = L + 1/2$, the field-dependence of the levels, as given by Eqs. (3.13)–(3.15), reduces to the linear one. For other states, linearity of the energy dependence on the field occurs in the high-field regime due to Paschen-Back uncoupling of spin from rotation of the molecule.

Figures 3.2(b) and (c) reveal that for a finite bath density, the molecular levels are shifted towards lower energies. This effect, known as polaron shift, is a result of isotropic interactions between the impurity and the bath and has been widely studied for structureless impurities [Dev16]. Apart from the polaron shift, Figures 3.2(b) and (c) reveal a complex spectral structure emerging from the molecule-bath interaction. Namely, a lot of metastable states (shades of yellow) appear in between the stable angulon states (dark lines).

In order to understand this fine structure in detail, in the middle and right columns of Fig. 3.2 we present the spectral functions for the four selected states, $A(1/2, 0, -1/2)$, $B(1/2, 0, 1/2)$, $C(3/2, 1, 1/2)$, and $D(3/2, 1, 3/2)$, at two selected values of the magnetic field strength, $\eta = 2$ (middle column) and $\eta = 4$ (right column).

The observed features are qualitatively similar to that predicted in Ref. [SL15] for the spinless angulon in the absence of a magnetic field. For the vanishingly low density of the bath, $\tilde{n} = \exp(-15)$, Fig. 3.2(a), each of the states is given by a sharp peak, which approximately coincides with the molecule spectrum in the absence of a bath. For a finite density, $\tilde{n} = \exp(-7)$, an additional fine structure emerges in the spectrum, as labeled by the primed letters A', B', C', and D'. This is the so-called Many-Body-Induced Fine Structure (MBIFS) of the first kind [SL15], which emerges due to dressing of the stable angulon state with a phonon excitation carrying zero angular momentum, $\lambda = 0$ – this effect is described in more detail below. In a magnetic field, the position of this phonon wing changes in the same way as that of the stable angulon state: the states B, C, and D are shifted towards higher energies for larger η, while the energy of the A state decreases. For an even larger density, $\tilde{n} = \exp(-3)$, Fig. 3.2(c), the splitting between the stable angulon peak and the attached phonon continuum increases further. Note that the features C' and D' for $\eta = 2$ and B', C', and D' for $\eta = 4$ move outside the range of the corresponding plots. Moreover, the further the phonon branch is from the main angulon line, the broader is the spectral feature associated with it.

To provide a detailed analysis of the physical phenomena happening in the presence of the bath and a magnetic field, in Fig. 3.3 we study the states one by one. The left column presents bath-free bare molecular states. While moving to the middle and right columns, first we notice the splitting of the levels due to the MBIFS which leads to splitting of lines in each plot into a doublet. This splitting results from the isotropic term $U_0(k)$ of the molecule-bath interaction in Eq. (3.20) and can be understood (approximately) as a splitting between the states $\left|\tilde{J}, L, M_J\right\rangle \otimes |\text{no phonons}\rangle$ and $\left|\tilde{J}, L, M_J\right\rangle \otimes |\text{one phonon with } \lambda = 0\rangle$. This effect is reminiscent of the phonon wings predicted for acetylene molecules in He nanodroplets in Ref. [ZKW04], and will not be the main focus of our studies.

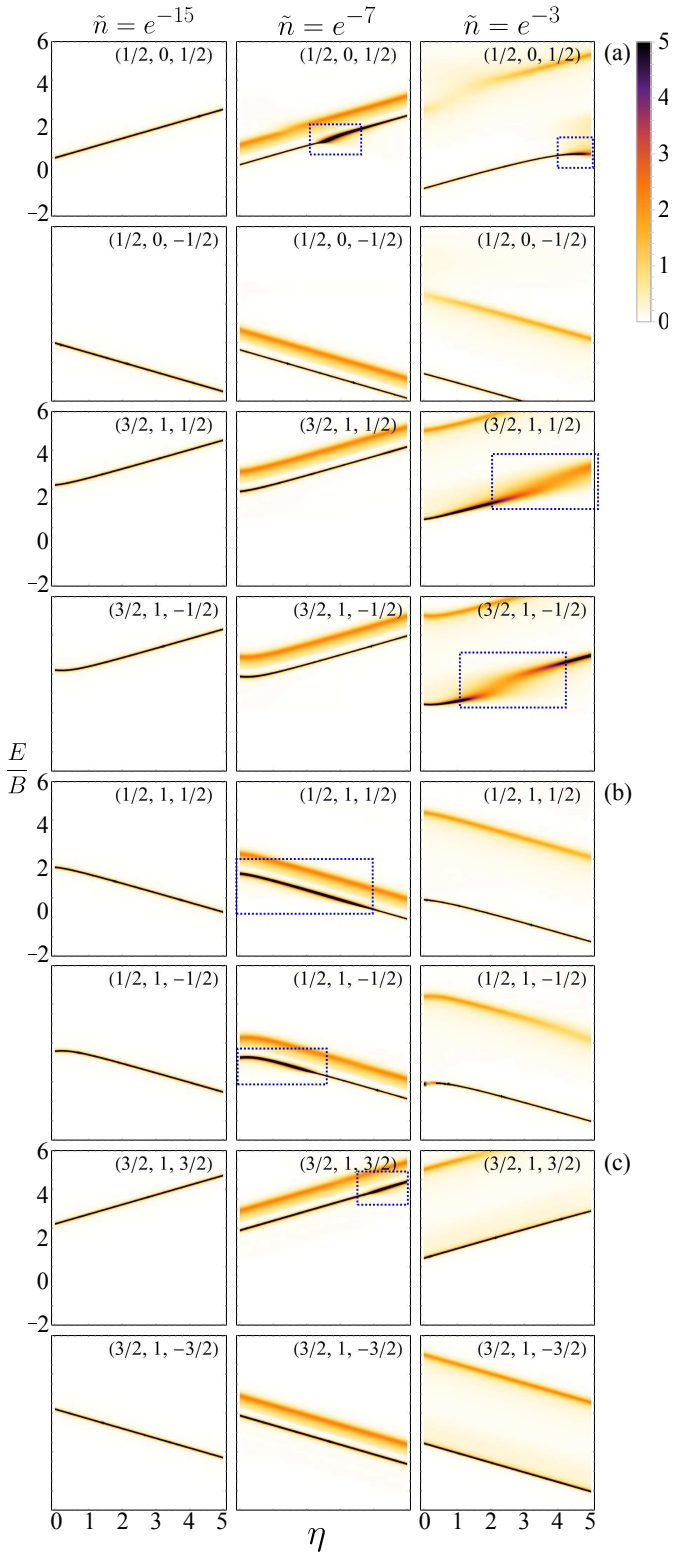


Figure 3.3: Spectral functions of all eight $L = 0$ and $L = 1$ states vs. magnetic field strength and energy for densities $\tilde{n} = \exp(-15)$, $\tilde{n} = \exp(-7)$, $\tilde{n} = \exp(-3)$. The first column on the left (density of $\tilde{n} = \exp(-15)$) corresponds to bare molecular states in magnetic field as in Ref. [FH00] The labels (a), (b), and (c) on the right correspond to panels in Fig. 3.5. The blue dotted frames mark clearly visible angulon instabilities. Reproduced from Ref. [RL18], with the permission of AIP Publishing.

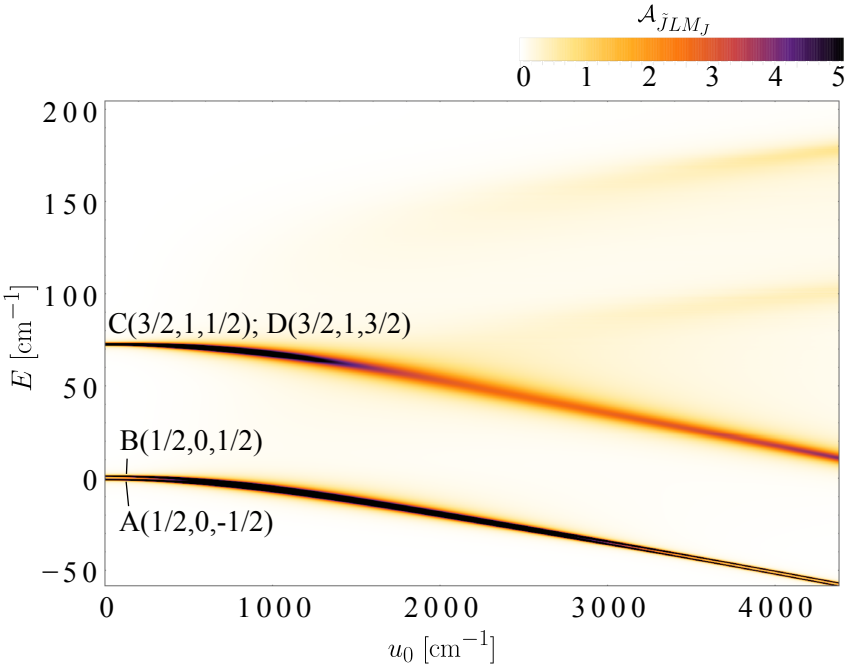


Figure 3.4: An example of a spectral function in absolute units, plotted for a molecule with $B = 14.603$ cm^{-1} (that of CH), spin-rotation coupling $\gamma = 0.4B$, in a field of $\mathscr{H} = 2$ T, as a function of the effective molecule-helium interaction u_0 and the energy E. The molecule is immersed in superfluid ^{4}He (superfluid density $n = 10^{22}$ cm^{-3}). The states C and D are labeled jointly as they lie too close to be resolved. Reproduced from Ref. [RL18], with the permission of AIP Publishing.

Secondly, we observe the emergence of instabilities which results from anisotropic interactions, $U_1(k)$ of Eq. (3.20). The instabilities are highlighted in Fig. 3.3 by the blue dotted frames. It has been previously shown that such instabilities are accompanied by the transfer of angular momentum from the impurity to the bath, i.e. the resonant emission of phonons with nonzero angular momentum. Since in our model we include only the leading anisotropic interaction term, $U_1(k)$, the emitted phonon carries angular momentum of $\lambda = 1$, however excitations with higher angular momentum are also possible [RL16]. As an example, the state $(1/2, 0, 1/2)$ features the angulon instabilities around $\eta = 2.8$ at the density of $\tilde{n} = \exp(-7)$ and around $\eta = 4.8$ at the density $\tilde{n} = \exp(-3)$, as can be seen from the first row of Fig. 3.3. These instabilities occur due to the interaction with the phonon continua of the states (1/2,1,1/2) and (1/2,0,-1/2), respectively. Another example is an instability of state $(3/2, 1, -1/2)$, taking place around $\eta = 2.5$ at the density of $\tilde{n} = \exp(-3)$, which is due to interaction with the phonon continuum of the state $(1/2, 0, -1/2)$. It was previously shown that the instabilities discussed above lead to anomalous screening of the impurities' dipole moments and polarizabilities [YL17], and can be manipulated using an external electrostatic field [RL16]. Crucially, from Fig. 3.3 one can see that the position of the instabilities depends on the magnitude of the magnetic field η as well. This paves the way to control the resonant emission of phonons with a given angular momentum, as discussed in the following section. As an example, in Fig. 3.4, plotted in absolute units, we illustrate the emergence of angulon instabilities for a molecule with the rotational constant $B = 14.603$ cm^{-1} (equal to that of CH), and the spin-rotation coupling $\gamma = 0.4B$, in a field of $\mathscr{H} = 2$ T. Although Fig. 3.4 is plotted for a high solvent density of $n = 10^{22}$ cm^{-3}, we also quantitatively account

for lower interactions (such as those occurring at lower densities of weakly-interacting BEC's) by including the regime of small interaction parameter u_0 in the plot (the ratio of $u_0/u_1 = 1.75$ is kept constant).

3.5 Angular momentum of the bath phonons and normalization of their populations

In this section we study the dependence of the phonon populations in individual λ channels,

$$\tilde{\beta}_\lambda(k) = \sum_{\tilde{j}l} |\tilde{\beta}_{\lambda\tilde{j}l}(k)|^2, \tag{3.44}$$

on the magnitude of the magnetic field. Here, the phonon populations $\tilde{\beta}_{\lambda\tilde{j}l}(k)$ can be obtained from the variational calculations as follows:

$$\tilde{\beta}_{\lambda\tilde{j}l}(k) = -\frac{U_\lambda(k) K^{\tilde{J}LM_J}_{\tilde{j}lm_j\lambda}}{\sum\limits_{m''_j} (C^{J,M_J}_{j,m''_j;\lambda,M_J-m''_j})^2 E^0_{jlm''_j} - E + \omega_k}, \tag{3.45}$$

and the energy E found as a solution of Eq. (3.41). The phonon populations $\tilde{\beta}_{\lambda\tilde{j}l}(k)$ are related to the variational coefficients of Eq. (3.24) by the following normalization relation:

$$\beta_{\lambda\tilde{j}l}(k) = \frac{\tilde{\beta}_{\lambda\tilde{j}l}(k)}{\left(1 + \sum\limits_{k\lambda\tilde{j}l} |\tilde{\beta}_{\lambda\tilde{j}l}(k)|^2\right)^{1/2}}. \tag{3.46}$$

This normalization is problematic, thus we discuss the unnormalized quantities as they fully show the proportion between phonon populations for different λ channels. Let us, though, discuss the details of the normalization issues. The quantity $\tilde{\beta}_{\lambda\tilde{j}l}(k)$ defined in Eq. (3.45) simply follows from Eq. (3.40) if one temporarily assumes for the calculation that the quasiparticle weight $Z_{\tilde{J}LM_J} = 1$. This results in normalization condition of Eq. (3.46) for $\beta_{\lambda\tilde{j}l}(k)$ and complementary normalization condition for $Z_{\tilde{J}LM_J}$:

$$Z_{\tilde{J}LM_J} = \frac{1}{1 + \sum\limits_{k\lambda\tilde{j}l} |\tilde{\beta}_{\lambda\tilde{j}l}(k)|^2}. \tag{3.47}$$

As we can see, the normalizations of both phonon population $\beta_{\lambda\tilde{j}l}(k)$ and quasiparticle weight $Z_{\tilde{J}LM_J}$ inevitably involve an integral of the following type:

$$\int_0^\infty dk \frac{f(k)}{(k-k_0)^2}. \tag{3.48}$$

At the angulon instabilities, there is a pole and the integral is divergent. This physically results in $|\beta_{\lambda\tilde{j}l}(k)|^2$ close to one for a given combination of $\lambda, \tilde{j}, l$ parameters and the rest of $\beta_{\lambda\tilde{j}l}(k)$ coefficients and the quasiparticle weight $Z_{\tilde{J}LM_J}$ being close to 0. Although there exist techniques such as Hadamard regularization, in our problem we need to calculate the value of the integral, not its finite part.

Let us also note that the problem presented above does not influence the calculations of spectral functions. There, the integrals are of type $\int_0^\infty dk \frac{f(k)}{k-k_0}$. There still might be

a pole, but as we calculate the spectral function, the form of Eq. (3.42) causes us to introduce a small but finite imaginary part to the denominator and take the real part of the integral. Mathematically, we know from Sokhotski-Plemelj theorem:

$$\lim_{\epsilon\to 0^+}\int_a^b dk\frac{f(k)}{k-k_0\pm i\epsilon} = \mp i\pi f(k_0) + \mathcal{P}\int_a^b dk\frac{f(k)}{k-k_0}, \tag{3.49}$$

where $\mathcal{P}$ denotes the Cauchy principal value and $a < k_0 < b$, that by taking ϵ small enough, in our numerical method we approach the Cauchy principal value of the integral. Hence, had we been able to calculate the analytical form of spectral function from Eq. (3.42), it would differ from the one obtained numerically in this Chapter (with an introduction of a finite imaginary part) only by sharper spectral features.

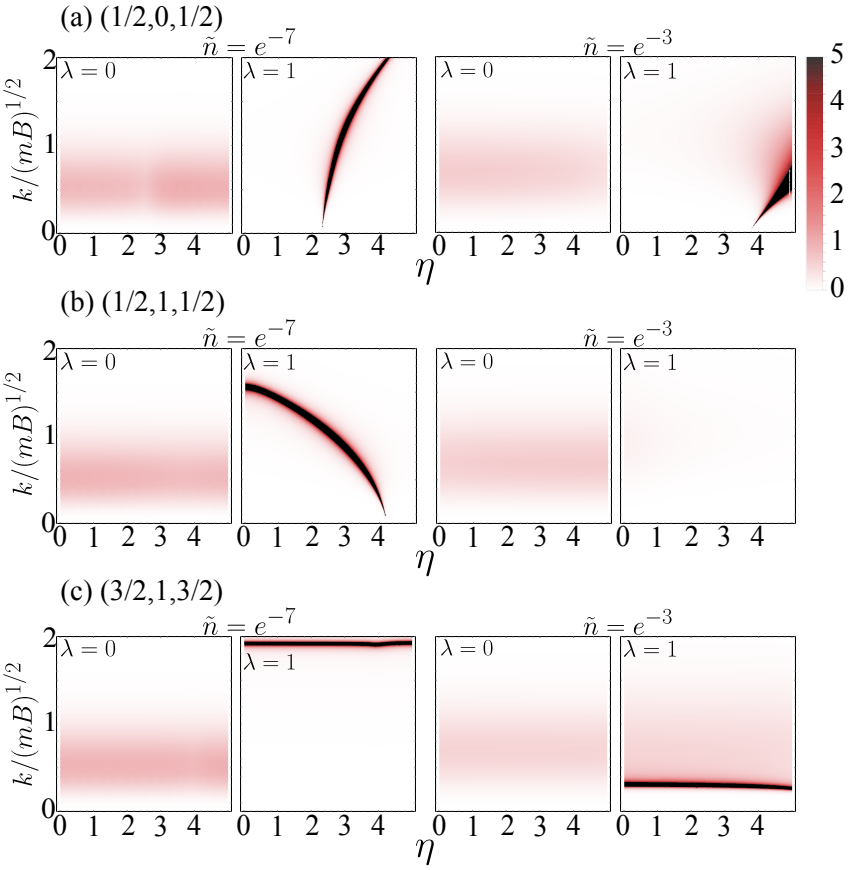


Figure 3.5: Dependence of the phonon populations $\tilde{\beta}_\lambda(k)$ on the magnitude of the magnetic field η and wavevector k for two first channels $\lambda = 0, 1$ in two different densities $\tilde{n} = e^{-7}$ and $\tilde{n} = e^{-3}$. The three panels feature different states: (a) (1/2,0,1/2), (b) (1/2,1,1/2), (c) (3/2,1,3/2). Reproduced from Ref. [RL18], with the permission of AIP Publishing.

Fig. 3.5 shows the phonon populations $\tilde{\beta}_\lambda(k)$ for two lowest channels, $\lambda = 0, 1$, depending on the magnetic field strength η and the wavevector k. As discussed above, the phonons with $\lambda = 0$ emerge as a phonon wing surrounding the main angulon state, while the $\lambda = 1$ phonons are resonantly emitted at the angulon instability. This reflects itself in moderate emission of phonons into the $\lambda = 0$ channel across all studied densities and magnetic field and infinitely growing $\tilde{\beta}_1(k)$ at the instabilities. By comparing Figs. 3.3 and 3.5 we can see that the strong emissions in $\lambda = 1$ channel occur for the same range of the field as the instabilities in the spectral function. Let us note that the instabilities can look

differently on the spectral function plots – they can be manifested as line discontinuities or line broadening. Moreover, stable shape of the spectral function of (1/2,1,1/2) state for $\tilde{n} = e^{-3}$ corresponds to almost no emission to $\lambda = 1$ for this state and density.

Thus, by manipulating the angulon energies with a magnetic field, one can control the emission of phonons with a given angular momentum λ, and fine-tune the phonon populations in different λ-channels.

3.6 Spinful angulon in the absence of external fields

In this Section, we describe an intermediate case between that described by Hamiltonian derived in Section 3.2 and the original angulon theory introduced in Ref. [SL15]. Namely, we consider a zero-field limit, $\eta = 0$, in the presence of the spin-rotation coupling, $\gamma \neq 0$.

The new variational ansatz reads:

$$|\psi_{JLM_J}\rangle = \underbrace{Z^{1/2}_{JLM_J}|0\rangle\,|J,L,M_J\rangle}_{:=\left|\psi^1_{JLM_J}\right\rangle} + \underbrace{\sum_{\substack{k\lambda\mu\\ jlm_j}} \beta_{\lambda jl}(k) C^{J,M_J}_{j,m_j;\lambda,\mu}\hat{b}^\dagger_{k\lambda\mu}|0\rangle\,|j,l,m_j\rangle}_{:=\left|\psi^2_{JLM_J}\right\rangle}. \tag{3.50}$$

In full analogy with the case when the field is present, we minimize the following functional

$$\begin{aligned} F &= \langle\psi_{JLM_J}|\,\widehat{H}\,|\psi_{JLM_J}\rangle - E\,\langle\psi_{JLM_J}|\psi_{JLM_J}\rangle \\ &= \left\langle\psi^1_{JLM_J}\right|\widehat{H}\left|\psi^1_{JLM_J}\right\rangle + \left\langle\psi^2_{JLM_J}\right|\widehat{H}\left|\psi^2_{JLM_J}\right\rangle + \\ &+ \left(\left\langle\psi^1_{JLM_J}\right|\widehat{H}\left|\psi^2_{JLM_J}\right\rangle + c.c\right) - E\,\langle\psi_{JLM_J}|\psi_{JLM_J}\rangle\,. \end{aligned} \tag{3.51}$$

We calculate F term-by-term:

$$\langle\psi_{JLM_J}|\psi_{JLM_J}\rangle = |Z_{JLM_J}| + \sum_{k\lambda jl}|\beta_{\lambda jl}(k)|^2, \tag{3.52}$$

and

$$\left\langle\psi^1_{JLM_J}\right|\widehat{H}\left|\psi^1_{JLM_J}\right\rangle = |Z_{JLM_J}|E^0_{JL}. \tag{3.53}$$

where E^0_{JL} is the energy of the $|J,L,M_J\rangle$ state as defined by Eq. (3.3). The next term reads

$$\left\langle\psi^2_{JLM_J}\right|\widehat{H}\left|\psi^2_{JLM_J}\right\rangle = \sum_{k\lambda jl}|\beta_{\lambda jl}(k)|^2\left(\omega_k + E^0_{jl}\right). \tag{3.54}$$

Furthermore, we have

$$\left\langle\psi^1_{JLM_J}\right|\widehat{H}\left|\psi^2_{JLM_J}\right\rangle = \left\langle\psi^1_{JLM_J}\right|\widehat{H}_{\text{mol-bos}}\left|\psi^2_{JLM_J}\right\rangle, \tag{3.55}$$

from which we obtain:

$$\left\langle\psi^1_{JLM_J}\right|\widehat{H}\left|\psi^2_{JLM_J}\right\rangle = (Z^*_{JLM_J})^{1/2}\sum_{k\lambda jl}U_\lambda(k)\beta_{\lambda jl}(k)K^{JL}_{jl\lambda}, \tag{3.56}$$

where

$$\begin{aligned} K^{JL}_{jl\lambda} =& C^{L,0}_{l,0;\lambda,0}\sqrt{\frac{(2\lambda+1)(2l+1)}{4\pi}}\times \\ &\times 4(J-L)(j-l)(-1)^{1/2+l+\lambda+J}\sqrt{(2(l+J-L)+1)}\begin{Bmatrix} 1/2 & l & j \\ \lambda & J & L \end{Bmatrix}. \end{aligned} \tag{3.57}$$

Now we vary F with respect to the parameters, obtaining:

$$\frac{\partial F}{\partial (Z^{1/2}_{JLM_J})^*} = Z^{1/2}_{JLM_J}(E^0_{JL} - E) + \sum_{k\lambda jl} U_\lambda(k) K^{JL}_{jl\lambda} \equiv 0, \tag{3.58}$$

$$\frac{\partial F}{\partial \beta_{\lambda jl}(k)^*} = \beta_{\lambda jl}(k)(\omega_k + E^0_{jl} - E) + Z^{1/2}_{JLM_J} U_\lambda(k) K^{JL}_{jl\lambda} \equiv 0. \tag{3.59}$$

Then, substituting β from Eq. (3.59) into Eq. (3.58) we arrive at the following Dyson equation:

$$0 = E^0_{JL} - E - \sum_{k\lambda jl} \frac{U^2_\lambda(k)(K^{JL}_{jl\lambda})^2}{\omega_k + E^0_{jl} - E}, \tag{3.60}$$

The last term in Eq. (3.60) is the self-energy, from which one can obtain the Green's function and the spectral function.

3.7 Chapter conclusions and outlook

In this Chapter, we used the angulon theory to study a quantum impurity with spin$-1/2$ and rotational angular momentum, immersed in a many-particle bath of bosons, in the presence of an external magnetic field. We have shown that the field can be used to manipulate the exchange of angular momentum between the impurity and the bath, as mediated by spin-rotation coupling. In turn, this paves the way to manipulate the positions of the angulon instabilities [SL15, SL16, LS17, YL17, RL16], and thereby control the angular momentum of phonons in the bath. Recently, the signatures of the angulon instabilities were found in spectra of molecules in helium nanodroplets [CL17], which opens up a possibility to test the presented predictions in experiment. In particular, we expect that the effects predicted in this Chapter can be detected through electron spin resonance measurements on molecules in superfluid helium nanodroplets [KACE09]. For typical $^2\Sigma$ molecules, such as CaF, SrF, and CH, the respective rotational constants are 0.338 cm^{-1} [FHT75], 0.250 cm^{-1} [DSH77], and 14.603 cm^{-1} [HH79] which makes $\eta = 1$ correspond to a field of 0.36 T, 0.27 T, and 15.62 T, respectively. Table 3.1 lists the quantities used throughout Chapter 3 in absolute units for $^2\Sigma$ electronic ground states of CaF, SrF, and CH molecules. Substantially smaller field magnitudes are expected to be required for molecules containing highly magnetic elements, such as erbium [AFM$^+$12] and dysprosium [LBYL11]. We note that in the present work translational motion of molecules in a superfluid has been neglected. The rotation-translation coupling can lead to additional inhomogeneous broadening of the spectroscopic lines [Leh99] and we are currently extending the angulon model to account for it.

Finally, we would like to note that the formalism presented in this Chapter is quite general, and can be applied to any spinful impurity possessing rotational or orbital angular momentum, immersed into, in principle, any kind of a bosonic bath. Therefore, the predicted effects can find potential applications not only for molecules trapped in superfluid helium nanodroplets [TV04] and ultracold gases [MTSL16], but also to Rydberg excitations in Bose-Einstein condensates [BKG$^+$13] and non-equilibrium magnetism in solids [KRLdJ05, CGS05, ZN14, TP15, GC15a, TIH$^+$16].

Table 3.1: Values of energy, bath density, and magnetic field in absolute units for CaF, SrF, and CH molecules immersed in superfluid ^{4}He.

Molecule	CaF	SrF	CH
E=1	0.338 cm^{-1}	0.250 cm^{-1}	14.603 cm^{-1}
$\tilde{n} = \exp(-15)$	$2.46 \cdot 10^{13}$ cm^{-3}	$1.56 \cdot 10^{13}$ cm^{-3}	$6.98 \cdot 10^{15}$ cm^{-3}
$\tilde{n} = \exp(-7)$	$7.32 \cdot 10^{16}$ cm^{-3}	$4.66 \cdot 10^{16}$ cm^{-3}	$2.08 \cdot 10^{19}$ cm^{-3}
$\tilde{n} = \exp(-3)$	$4.00 \cdot 10^{18}$ cm^{-3}	$2.54 \cdot 10^{18}$ cm^{-3}	$1.14 \cdot 10^{21}$ cm^{-3}
$\eta = 1$	0.36 T	0.27 T	15.62 T

CHAPTER 4

Neural-network quantum states approach to the polaron Hamiltonian

4.1 Chapter intro

In Chapter 2, we have discussed an integration of machine learning ideas into the study of quantum physics. As mentioned there, such approaches recently attracted great interest, owing to the new possibilities it offers to tackle challenging problems in quantum physics [BWP+17, CCC+19, BC21].

Pioneered by Carleo and Troyer [CT17], whose approach we discussed in detail in Section 2.2, a particularly appealing approach is to represent the quantum many body wave function by an artificial neural network. This was first demonstrated to quantum spin systems in one and two dimensions and subsequently generalized to bosonic [Sai17, SK17] and fermionic [NDYI17, CL18, CMC20] systems. Moreover, beyond pure quantum states, artificial neural networks can also accurately represent mixed quantum states in open systems [TM18, HC19, YH19, VBRC19] and quantum systems at finite temperature [IS20].

However, all these examples involve additive many-body systems, for which by definition the total number of particles is conserved. On the other hand, there is an important class of physical systems which does not satisfy particle number conservation. This comprises the important class of inherently non-additive quantum impurity systems. Such systems include for example:

- Fröhlich [Frö54] model, discussed in detail in Section 1.2.2, a paradigmatic model of a polaron – a moving impurity inside a crystal,
- Holstein [Hol59] model, a second important polaron model. Compared to Fröhlich model, it focuses on *small polaron*, in which the *polaron radius* is comparable with the crystal lattice size. The polaron radius is the uncertainty of the electron's position if using the phonon field as the measuring device,
- Su-Schrieffer-Heeger (SSH) [SSH79, SSH80], a simple two-sublattice model displaying topological properties,

- Dicke model [HL73, Dic54], a simple model of light-matter interaction, where a single light mode interacts with a collection of two-level systems,
- Anderson impurity model [And61], which describes magnetic impurities in metals, by treating them as a two-level system interacting with the conduction electrons

These models are quite simple, yet they display qualitatively rich physics. As such, they play a crucial role in the understanding of quantum many-body systems. However, analytic solutions are often available only in limiting cases. To address the full complexity of the problem in a wide gauge of physical parameters, numerical methods are necessary.

Neural-network quantum states are, in principle, a prospective candidate for efficient representation of such complex quantum many-body wavefunctions. For example, the restricted Boltzmann Machine (RBM), one of the simplest and most widely used architectures [MCCC19], exhibits volume law entanglement and can represent even models with long-range interactions [DLDS17]. Analogous to variational Monte Carlo (VMC) approaches to the Holstein and SSH model [OI14, KTSB17, FVB20], recently the electron-phonon correlation factor was represented using an RBM, while keeping a Jastrow correlation factor for the electron subsystem. Lattice polarons have also been tackled with Gaussian process regression capable of extrapolating across their phase transitions [VHSBK18]. In addition, the Anderson impurity model has been addressed with machine learning methods to find the Green function [ALBvLM14] and to derive its low-energy effective model [RM20]. However, so far no neural network states exists which directly provide an unbiased estimate of the full many-body wave function of non-additive systems.

In this Chapter we show that efficient artificial neural network quantum states for non-additive systems can be constructed as a feed-forward neural network with outputs inspired by coherent states well-known from quantum optics [Gla63]. We investigate the efficiency of this architecture by considering the Fröhlich model featuring long-range interactions between the phonon degrees of freedom. We benchmark our architecture, termed *neural coherent states*, against exact diagonalization and mean field solutions. In all cases studied, we find that this approach outperforms the standard mean-field coherent state solution, in particular when impurity-induced phonon-phonon correlations are strong.

In Section 4.3 we introduce a novel neural network architecture, the neural coherent states, as a tool to express wavefunctions of non-additive systems with artificial neural networks. We justify the necessity to introduce this architecture by proving that the restricted Boltzmann machine quantum states are not capable of expressing the polaron ground state even in a limiting case. In Section 4.2, we adapt the Hamiltonian described in Section 1.2.2 for numerical treatment and discuss its approximate mean-field solution. In Section 4.4.1 we describe the details of variational optimization, including the inherent effects characteristic to non-additive systems. In Section 4.5 we discuss our numerical results and benchmark them against the mean-field and exact diagonalization solutions.

4.1.1 Authors' contributions

JM and ML conceived the project. I designed the neural coherent states architecture and performed the numerical experiments, exchanging ideas closely with JM and ML.

4.2 Impurity Hamiltonian and its approximate solution

To test the efficiency of the NCS for non-additive systems, we focus on the Fröhlich Hamiltonian already discussed in Section 1.2.2. Here, we just remind the form of this Hamiltonian already after transformation to the impurity frame, which is achieved by the Lee-Low-Pines transformation [LLP53]. In the sector of zero total momentum, the Hamiltonian is:

$$\hat{H} = \frac{\left(-\sum_{\mathbf{k}} \hbar\mathbf{k}\hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}}\right)^2}{2m} + \sum_{\mathbf{k}} \hbar\omega_0 \hat{a}^\dagger_{\mathbf{k}}\hat{a}_{\mathbf{k}} + \sum_{\mathbf{k}} \left(V_{\mathbf{k}}\hat{a}_{\mathbf{k}} + V^*_{\mathbf{k}}\hat{a}^\dagger_{\mathbf{k}}\right). \tag{4.1}$$

The Lee-Low-Pines transformation removes the impurity degrees of freedom from the Hamiltonian. This maps the problem to a pure problem of interactions between the bosonic modes, at the price of introducing effective interactions between the phonon modes, described by the first term of the transformed Hamiltonian. The transformed impurity Hamiltonian problem is closer to the lattice boson problems such as the Bose-Hubbard problem, studied earlier with different NQS architectures [Sai17, SK17]. However, the problem mentioned earlier that the total number of bosons is not conserved, persists.

To make the Hamiltonian more convenient for numerical computation, we measure energy in units of $\hbar\omega_0$. Moreover, we discretize the k-grid to include N points k_i ranging from $-k_0$ to k_0 with step Δk. This finally puts the Hamiltonian into the following form in one dimension:

$$\hat{H} = \frac{\left(-\sum_i \hbar k_i \hat{a}^\dagger_{k_i}\hat{a}_{k_i}\right)^2}{2m} + \sum_i \hat{a}^\dagger_{k_i}\hat{a}_{k_i} + \sum_i \left(V_{k_i}\hat{a}_{k_i} + V^*_{k_i}\hat{a}^\dagger_{k_i}\right). \tag{4.2}$$

In the calculations we further restrict the maximum number of bosons at each mode at a value $n_{\max}$, ranging between 3 and 8, depending on the parameter regime.

To benchmark our results, we compare them with two approaches – exact diagonalization, and the mean field approach [Dev16], where the ground state $|\psi_{\mathrm{MF}}\rangle$ is a direct product of coherent states, resulting in energy E_{MF}:

$$|\psi_{\mathrm{MF}}\rangle = \bigotimes_i \left|-\frac{V_{k_i}}{1+\frac{\hbar^2k^2}{2m}}\right\rangle, E_{\mathrm{MF}} = -\sum_k \frac{|V_k|^2}{1+\frac{\hbar^2k^2}{2m}}. \tag{4.3}$$

By such choice of benchmarks, we are able to quantify the correlations expressed with our ansatz.

4.3 Neural-network architecture

Our novel ansatz is necessitated by the fact an RBM operates in a fixed particle number sector. Let us provide an explicit proof that an RBM is not capable of representing the coherent state. This is relevant because the ground state of Fröhlich hamiltonian is coherent in the infinite mass limit and in the mean-field approximation. Let us, without loss of generality, consider just one k-point, i.e. $\mathbf{n} \equiv n$. The ground state $|-V\rangle$ is coherent and its wavefunction is given by:

$$\psi_{\mathrm{GS}}(n) = \langle n|\psi_{\mathrm{GS}}\rangle = \exp\left(-\frac{|V|^2}{2}\right)\frac{(-V^*)^n}{\sqrt{n!}} \tag{4.4}$$

It needs to be represented by the RMB variational ansatz:

$$\psi_{\mathrm{RBM}}(n) = \exp(an)(\exp(b + Wn) + \exp(-b - Wn)) \tag{4.5}$$

The task is to find complex numbers a, b, W such that Eq. (4.5) matches Eq. (4.4).

This is infeasible. Proof (by contradiction): assume that one can find a, b, W such that $\psi_{\mathrm{RBM}}(n) = \psi_{\mathrm{GS}}(n)$. Then equating $\psi_{\mathrm{RBM}}(n+1)/\psi_{\mathrm{RBM}}(n)$ with $\psi_{\mathrm{GS}}(n+1)/\psi_{\mathrm{GS}}(n)$ we obtain:

$$-\frac{V^\star}{n+1} = e^a \frac{\cosh(b + Wn + W)}{\cosh(b + Wn)} \tag{4.6}$$

The left and right hand sides of this equation have different $n \to \infty$ limits. The left hand side decreases to 0 with rising n while right hand side tends to 1 with rising n. This ends the proof. Let us now move to our new architecture.

We use a basis corresponding to bosonic occupations of the system with $\mathbf{n}$ denoting a single bosonic configuration of the whole system:

$$|\mathbf{n}\rangle \equiv |n_1, n_2, \ldots, n_i, \ldots, n_N\rangle\,, \tag{4.7}$$

where N is the number of discrete phonon modes considered. In an RBM architecture, N is equal to the number of visible neurons. However, direct application of an RBM to non-additive systems is not efficient. This can easily been seen by writing the neural-network quantum state as $|\psi(\mathbf{n})\rangle = \psi(\hat{\mathbf{n}})\,|\mathbf{n}\rangle$; we have $[\psi(\hat{\mathbf{n}}), \hat{N}] = 0$, $\hat{N} = \sum_i \hat{n}_i$ for a function $\psi(\mathbf{n}) \sim \exp(E(\mathbf{n}))$, E being linear in $\mathbf{n}$. Hence, by construction an RBM operates in a sector of a given total number of particles[1].

To bypass this problem, we propose a neural network inspired from coherent states, which may be termed neural coherent states (NCS) and is illustrated in Fig. 4.1. Analogous to a standard coherent state, which for a given n returns an output proportional to $\lambda^n/\sqrt{n!}$, with λ being the parameter representing the coherent state, we construct a (mutilayer) feedforward neural network taking $\mathbf{n}$ as input. For each configuration, N output numbers λ_i are generated, which are subsequently transformed according to: $\lambda_i \to {\lambda_i}^{n_i}/\sqrt{n_i!}$. Then these numbers are multiplied to form the wavefunction. If the solution is exactly a coherent state λ for each mode i, the network simply learns that $\lambda_i \equiv \lambda$ regardless of n_i. Correlated solutions are represented by perturbing the numbers λ_i, such that they depend on the input vector $\mathbf{n}$ in arbitrary way. This yields the variational ansatz expressed as:

$$\langle \mathbf{n}|\psi\rangle = \psi(\mathbf{n}) = \psi(n_1, n_2, \ldots, n_N) = \prod_{i=1}^{N} \frac{\lambda_i^{n_i}}{\sqrt{n_i!}}, \tag{4.8}$$

where λ_i is the output of a feedforward neural network (multilayer perceptron) with M layers, i.e.:

$$\boldsymbol{\lambda} = \{\lambda_1, \lambda_2, \ldots, \lambda_N\} = h_M(h_{M-1}(\ldots h_1(\mathbf{n}))), \tag{4.9}$$

with each of the hidden layer transformations h_j acting on output $\mathbf{h}_{j-1}$ of the previous layer:

$$h_j\left(\mathbf{h}_{j-1}\right) = \sigma\left(\mathbf{W}_j^{j-1}\mathbf{h}_{j-1} + \mathbf{b}_j\right) \tag{4.10}$$

where $\mathbf{W}_j^{j-1}$ and $\mathbf{b}_j$ are weights and bias of j-th layer, while $\sigma(x)$ is an activation function. This function, which introduces nonlinearity in the network, can be chosen from a wide range of classes already discussed in Section 2.2, such as tanh (the choice made in this work), sigmoid or ReLU [NH10].

[1] See the Supplemental Material for more information.

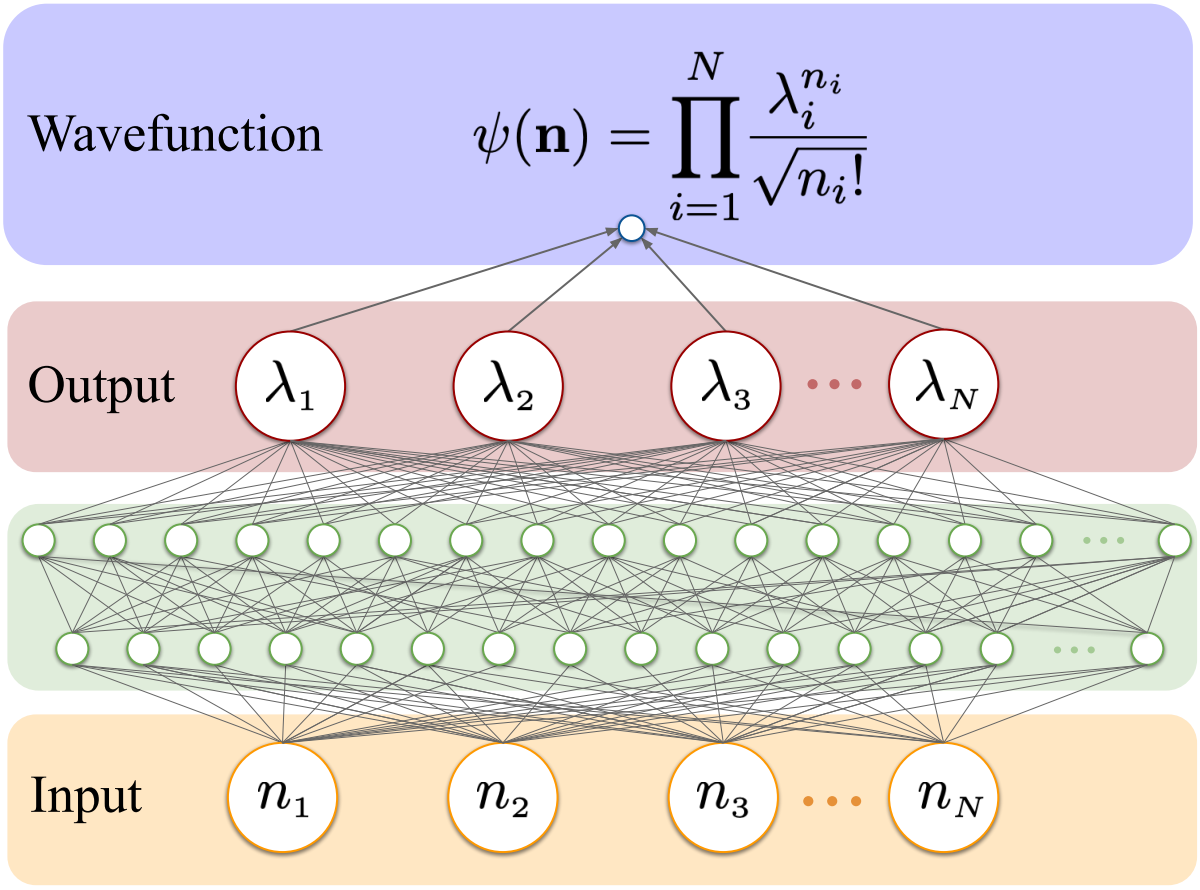


Figure 4.1: Visualization of the NCS approach. The input consists of Fock occupation numbers n_i for each of bosonic modes $i = 1, \ldots, N$, corresponding to discrete k values k_i. The input is fed to a multilayer perceptron with arbitrary number and size of the hidden layers, see text for details. The number of neurons in the output layer is equal to the number of inputs (k-points). Each of the outputs λ_i, is transformed using the information from the input, $\lambda_i \to \lambda_i^{n_i}/\sqrt{n_i!}$. These numbers are multiplied to form the wavefunction ψ. All neurons in the hidden layers are densely connected to all neurons in the neighbouring layers; for clarity of the picture not all of them are visualized with grey lines. Figure adapted from Ref. [RLM21].

4.4 Optimization of variational energy

With the ansatz from Eq. (4.8), we optimize the variational energy

$$E = \frac{\langle\psi|\,\hat{H}\,|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\sum\limits_{\mathbf{n},\mathbf{n}'} \langle\psi|\mathbf{n}\rangle \left\langle\mathbf{n}\middle|\hat{H}\middle|\mathbf{n}'\right\rangle \langle\mathbf{n}'|\psi\rangle}{\sum\limits_{\mathbf{n}} \langle\psi|\mathbf{n}\rangle \langle\mathbf{n}|\psi\rangle} = \frac{\sum\limits_{\mathbf{n},\mathbf{n}'} \psi^*(\mathbf{n}')\hat{H}_{\mathbf{n}'\mathbf{n}}\psi(\mathbf{n})}{\sum\limits_{\mathbf{n}} |\psi(\mathbf{n})|^2} \tag{4.11}$$

using Variational Monte Carlo, by sampling the probability distribution given by $|\psi(\mathbf{n})|^2$.

The derivatives of energy with respect to variational parameters ("forces") $\mathcal{F}_\xi(\mathcal{W}^{(p)})$, are given by:

$$\mathcal{F}_\xi = \left\langle E_{\mathrm{loc}} O_\xi^* \right\rangle - \langle E_{\mathrm{loc}} \rangle \left\langle O_\xi^* \right\rangle, \quad E_{\mathrm{loc}}(\mathbf{n}) = \frac{\sum\limits_{\mathbf{n}'} \hat{H}_{\mathbf{n}'\mathbf{n}}\psi(\mathbf{n}')}{\psi(\mathbf{n})}. \tag{4.12}$$

Here O_ξ are the logarithmic derivatives of the wavefunction with respect to variational parameters, which are indexed by a collective index ξ running across all weights and biases of the model:

$$O_\xi = \frac{1}{\psi(\mathbf{n})}\frac{\partial\psi(\mathbf{n})}{\partial\xi} = \frac{\partial \log\left(\psi(\mathbf{n})\right)}{\partial\xi}. \tag{4.13}$$

The quantity $E_{\mathrm{loc}}(\mathbf{n})$ is commonly referred to as local energy of the state $\mathbf{n}$. If the Hamiltonian matrix is sparse, local energy can be efficiently computed numerically.

4.4.1 Monte Carlo sampling

The braces $\langle\cdot\rangle$ in Eq. (4.12) refer to weighted averages over the probability distribution given by the wavefunction; the average of any quantity $x(\mathbf{n})$ is given by:

$$\langle x(\mathbf{n})\rangle = \frac{\sum\limits_{\mathbf{n}} |\psi(\mathbf{n})|^2 x(\mathbf{n})}{\sum\limits_{\mathbf{n}} |\psi(\mathbf{n})|^2}. \tag{4.14}$$

Such averages are not tractable as the sum runs over an exponential number of all possible states of the system. One has to resort to stochastic techniques. To sample the system for a given set of variational parameters and get the estimates $\langle\cdot\rangle$ of the logarithmic derivatives and local energies, we use the Metropolis-Hastings algorithm. At first, we choose a random initial configuration $\mathbf{n}^{(0)}$. Here, we choose $\mathbf{n}^{(0)}$ by randomly assigning 0 or 1 with probabilities 1/2 to each of the k-points k_i. This, in average, results in starting with a vector with mean occupation equal to $N/2$, i.e. half of the k-modes. We observed that allowing to start from higher mean occupations can lead to unstable sampling, i.e. that instead of reaching the high-probablity regions of the distribution, the sampling reaches states with mean occupation rising in an uncontrolled way. The possibility of such a behavior comes from the already mentioned fact that the non-additive Hamiltonian connects states with different total numbers of bosons, contrary to e.g. the Bose-Hubbard model.

Given a sample $\mathbf{n}^{(i)}$, finding the next sample $\mathbf{n}^{(i+1)}$ in the Metropolis-Hastings algorithm consists of two steps:

1. choosing a candidate for $\mathbf{n}^{(i+1)}$
2. calculating the acceptance probability $a(\mathbf{n}^{(i+1)}|\mathbf{n}^{(i)})$ and accepting the candidate $\mathbf{n}^{(i+1)}$ with that probability. If the candidate is not accepted, we remain in state $\mathbf{n}^{(i)}$.

For step 1, every choice is in principle possible. However, one needs to take into account the tradeoff between efficient exploration of the state space and still having a large acceptance probability. A choice of state $\mathbf{n}^{(i+1)}$ that differs too little from $\mathbf{n}^{(i)}$ might require too many steps to explore the state space, while a choice of state differing too much leads to lower acceptance ratio. The solution that we adopt here is the so-called Hamiltonian sampling, where the candidate $\mathbf{n}^{(i+1)}$ is chosen with uniform probability from all states connected by the Hamiltonian to $\mathbf{n}^{(i)}$ (i.e. having nonzero Hamiltonian matrix element $\hat{H}_{\mathbf{n}^{(i)}\mathbf{n}^{(i+1)}}$). We found other intuitive solutions, such as for example one that can be applied to Bose-Hubbard model – choosing randomly a pair of k-sites and swapping a

phonon between them, to work much worse. In particular, this solution conserves the boson number, while our Hamiltonian does not.

For step 2, the acceptance probability is proportional to the probability ratio of accepted states:

$$a\left(\mathbf{n}^{(i+1)}|\mathbf{n}^{(i)}\right) \propto \frac{p\left(\mathbf{n}^{(i+1)}\right)}{p\left(\mathbf{n}^{(i)}\right)} = \frac{\left|\psi(\mathbf{n}^{(i+1)})\right|^2}{\left|\psi(\mathbf{n}^{(i)})\right|^2}. \tag{4.15}$$

Moreover, as already noticed e.g. for Jastrow-type approaches [OI14], the detailed balance condition [NB99] needs to be fulfilled:

$$\frac{p\left(\mathbf{n}^{(i+1)}|\mathbf{n}^{(i)}\right)}{p\left(\mathbf{n}^{(i)}|\mathbf{n}^{(i+1)}\right)} = \frac{p\left(\mathbf{n}^{(i+1)}\right)}{p(\mathbf{n}^{(i)})}. \tag{4.16}$$

Observing that

$$p\left(\mathbf{n}^{(i+1)}|\mathbf{n}^{(i)}\right) = a\left(\mathbf{n}^{(i+1)}|\mathbf{n}^{(i)}\right) g\left(\mathbf{n}^{(i+1)}|\mathbf{n}^{(i)}\right), \tag{4.17}$$

where $g(\mathbf{n}^{(i+1)}|\mathbf{n}^{(i)})$ is the probability that state $\mathbf{n}^{(i+1)}$ was proposed given $\mathbf{n}^{(i)}$, and combining with the equality from Eq. (4.15) we arrive at:

$$a\left(\mathbf{n}^{(i+1)}|\mathbf{n}^{(i)}\right) = \frac{|\psi(\mathbf{n}^{(i+1)})|^2}{\left|\psi(\mathbf{n}^{(i)})\right|^2} \frac{g\left(\mathbf{n}^{(i)}|\mathbf{n}^{(i+1)}\right)}{g\left(\mathbf{n}^{(i+1)}|\mathbf{n}^{(i)}\right)}. \tag{4.18}$$

Our choice of update proposal yields

$$g(\mathbf{n}^{(i+1)}|\mathbf{n}^{(i)}))) \propto \frac{1}{\langle\langle \mathbf{n}^{(i)} \rangle\rangle}, \tag{4.19}$$

where $\langle\langle \mathbf{n} \rangle\rangle$ is the number of states connected to state $\mathbf{n}$ by the Hamiltonian. Hence, the probability of accepting the sample proposal $\mathbf{n}^{(i+1)}$ given current sample $\mathbf{n}^{(i)}$ is given by

$$a\left(\mathbf{n}^{(i+1)}|\mathbf{n}^{(i)}\right) = \frac{\left|\psi(\mathbf{n}^{(i+1)})\right|^2}{\left|\psi(\mathbf{n}^{(i)})\right|^2} \frac{\langle\langle \mathbf{n}^{(i)} \rangle\rangle}{\langle\langle \mathbf{n}^{(i+1)} \rangle\rangle} \tag{4.20}$$

The need for the term $\langle\langle \mathbf{n}^{(i)} \rangle\rangle / \langle\langle \mathbf{n}^{(i+1)} \rangle\rangle$ again results from the non-conservation of the total number of bosons and is absent (i.e. equal to 1) in systems such as the Bose-Hubbard model or Heisenberg spin. To improve the sampling stability, we run several Monte Carlo chains in parallel. Samples from each of them are collated to form the set of samples used at a given step of the optimization.

4.4.2 Gradient descent optimization

Having estimated the gradients as given by Eq. (4.12), we move to minimizing the variational energy. One is free to choose any optimization method, either inspired by physics, like stochastic reconfiguration [SCR07b] and linear method [FK21] or an algorithm chosen from a rich field of approaches originating from the field of machine learning. In this work, we choose the Adam algorithm [KB14]. This is an adaptive algorithm, where the variational parameters are updated as follows:

$$\xi^{(t+1)} = \xi^{(t)} - \eta \frac{\hat{m}_\xi^{(t)}}{\sqrt{\hat{v}_\xi^{(t)}} + \epsilon} \tag{4.21}$$

Here η is the learning rate. The quantities $\hat{m}_\xi$ ($\hat{v}_\xi$) are the estimates of the first (second) moment of the gradients. The small number ϵ provides numerical stability by preventing the denominator going to zero.

First, the bias-uncorrected estimates are obtained as follows

$$\begin{aligned} m_\xi^{(t+1)} &= \beta_1 m_\xi^{(t)} + (1-\beta_1)\, F_\xi^{(t)} \\ v_w^{(t+1)} &= \beta_2 v_\xi^{(t)} + (1-\beta_2)\left(F_\xi^{(t)}\right)^2 \end{aligned} \tag{4.22}$$

with β_1 and β_2 being freely adjustable parameters of the algorithm. The estimates are corrected as follows:

$$\begin{aligned} \hat{m}_\xi &= \frac{m_\xi^{(t+1)}}{1-\beta_1^{t+1}} \\ \hat{v}_\xi &= \frac{v_\xi^{(t+1)}}{1-\beta_2^{t+1}}. \end{aligned} \tag{4.23}$$

Intuitively, this corresponds to an extension of the momentum algorithm [Qia99]. In addition to having an estimate of the first moment (mean), like in the momentum algorithm, we divide the updates by an estimate of the second moment (variance).

The typical values of the hyperparameters tend to be agreed upon in the community. The small number ϵ is usually taken to be 10^{-8}. The β_1, which can be thought of as the "forgetting factor" for the first moment is usually set to $\beta_1 = 0.9$. The forgetting factor for the second moment is often $\beta_2 = 0.999$.

The choice of learning rate η is usually changed much more frequently from one application to another. The task here is to find an optimal tradeoff between the speed of learning and the stability of this process.

During the training, we have observed that some values of energy are outliers, i.e. they lie off the trend set by preceding and following energies. The phenomenon is illustrated in Fig. 4.2. We attribute this to a failure of a Monte Carlo chain, i.e. exploring not the entire distribution, but only a region of it. We go around this by always choosing a number (e.g. 100) of subsequent optimization steps that does not contain an energy outlier. Then, we choose the mean from these subsequent steps as the final energy and the standard deviation over these steps as the error of the final energy. Designing a method that excludes such points at runtime, for example by analyzing the discrepancy of energies obtained from the different Monte Carlo chains running in parallel, could be a valuable extension of our work. We have also observed that avoiding very low V regime, increasing the number of hidden nodes and decreasing the learning rate all contribute to a reduction of the frequency and magnitude of such instabilities.

4.4.3 Implementation

To facilitate reproduction of the results and experimentation with the proposed approach, we make the code available online under the following URL:
https://github.com/wrzadkow/ncs.

The code uses Jax [BFH+18] and Flax [HLO+20] libraries. The Jax library allows execution on CPUs, GPUs and TPUs without change and boasts a high level of control

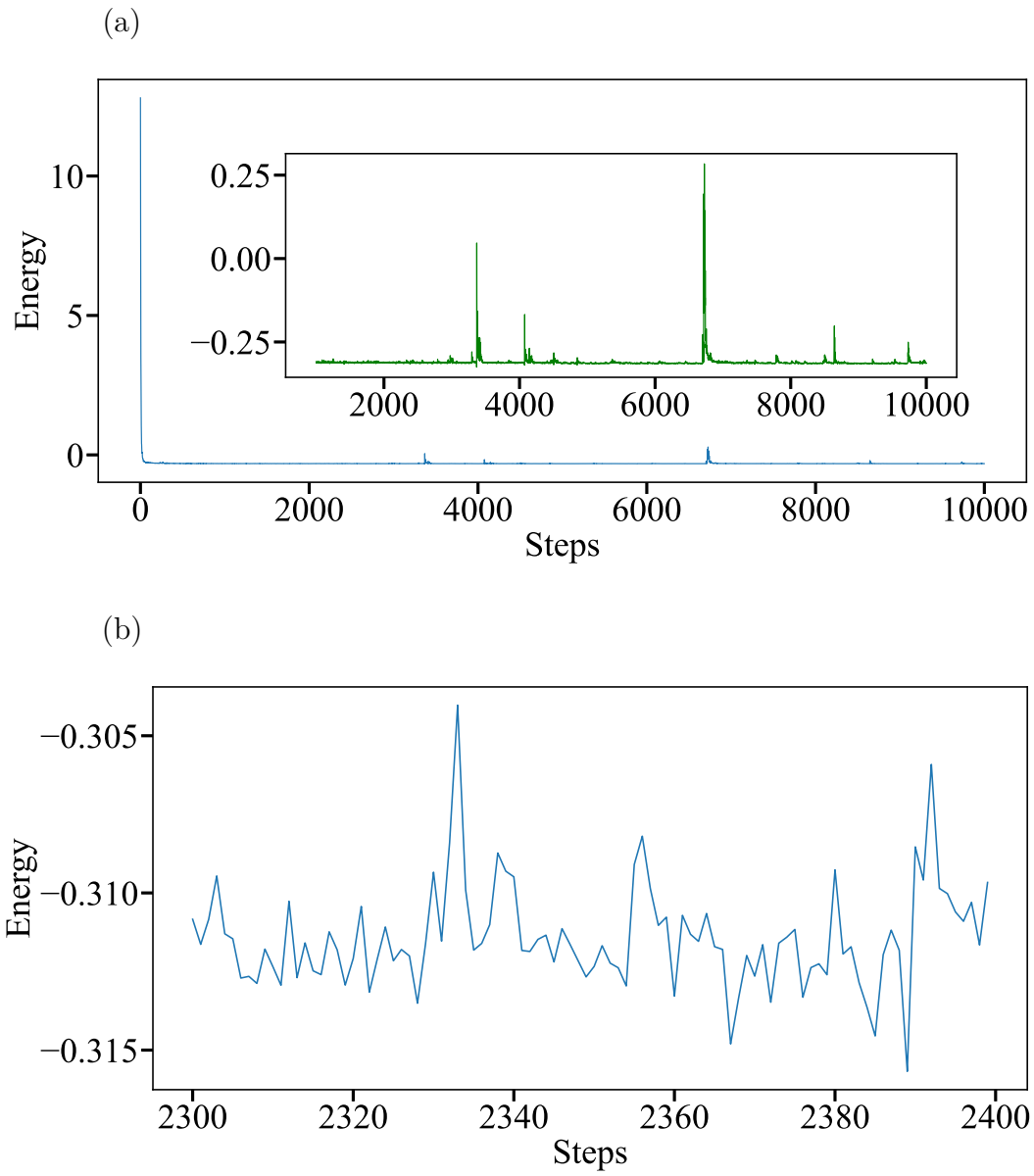


Figure 4.2: Typical example of occasional sampling failures. Panel (a) shows entire 10 000 training steps starting from random parameters, with inset omitting first 1000 steps. Panel (b) shows exemplary 100 continuous training steps from a stable part of the optimization, free of instabilities. Figure adapted from Ref. [RLM21].

over random number generation with the so-called Threefry counter pseudorandom number generator [SMDS11]. Moreover, the native vectorization functionality provided in Jax is used to parallelize the Monte Carlo chain. We use Flax, which is one of Jax's neural network libraries, to implement the multilayer perceptron neural network model.

The code is divided into modules. The energy module implements the Hamiltonian and local energy calculation. The wavefunction module contains functionality related to the variational ansatz and its optimization. Finally, Monte Carlo sampling is implemented in the sampler module. Further information about the code can be found in the documentation provided directly in the form of docstrings.

4.5 Numerical results

As the first test, we take a small system with $N = 2$ k-points $-k_0$ and k_0. We fix the impurity-bath potential at $V/(\hbar\omega_0) = 0.2$. Moreover, using convenient unit $m_0 = \hbar k_0^2/(2\omega_0)$ for the mass, such that $\hbar^2 k_0^2/(2m_0) = \hbar\omega_0$, we fix the inverse mass at $1/m = 0.6 \cdot (1/m_0)$. We vary the number of nodes in the single hidden layer, thus changing the number of variational parameters and, consequently, the representative power of the network. For each number of nodes, we optimize the energy and compare the obtained energy with the ED and mean field energy mentioned above. The results are shown in Fig. 4.3. We observe that with the number of nodes high enough, the variational energy

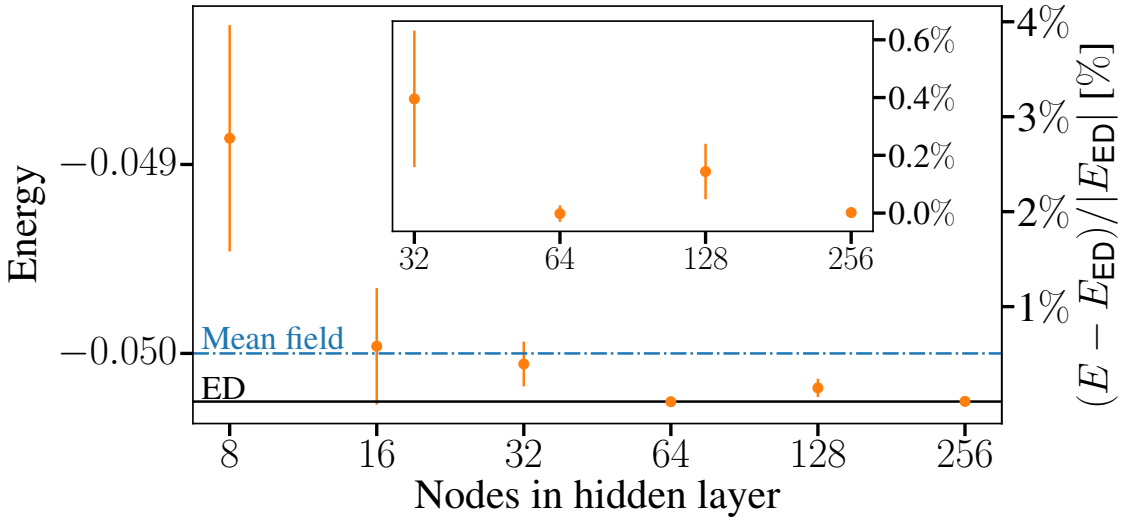
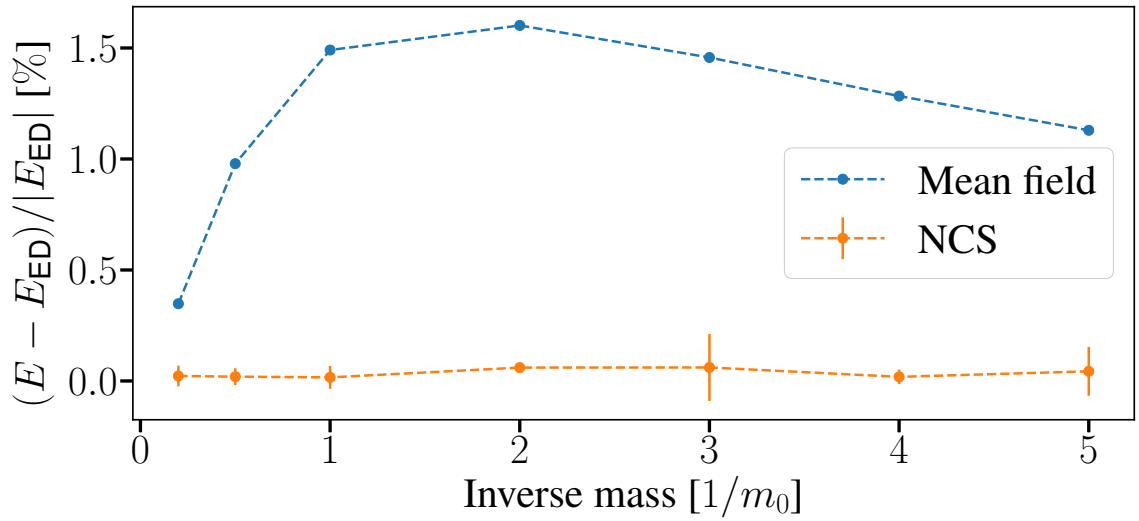


Figure 4.3: Representative power of the proposed approach. The optimized variational energy (orange dots, in units of $\hbar\omega_0$) is compared with exact diagonalization (ED), see right y-axis for percent scale relative to ED; and mean-field result for a system with 2 bosonic modes. The inset shows percent difference to ED for the four largest network sizes. Figure adapted from Ref. [RLM21].

clearly goes below the mean-field one, proving the capability of our approach. For 64 and 256 nodes in the hidden layer, we have obtained an agreement with the ED result within the stochastic error of our variational approach.

To further evaluate the ability of NCS to express correlations between different bosonic modes, we study the performance of our approach with different impurity mass. Low mass is associated with high correlation level, while high mass brings the Hamiltonian closer to the infinite mass regime, where an analytic solution in the form of coherent state exists. We fix V at $V/(\hbar\omega_0) = 0.2$. Then we optimize a network containing 1024 neurons in the hidden layer for different mass values and compare the result with the mean-field approach. The percentage deviation, $100\% \cdot (E - E_{\mathrm{ED}})/|E_{\mathrm{ED}}|$, for both the mean-field and NCS approach is shown in Fig. 4.4(a) as a function of the inverse mass. Here, we observe very stable performance – the NCS is able to outperform mean-field approach across the range of (inverse) mass tested. Even at the intermediate regime, where mean-field solution lies $\approx 1.5\%$ above the ED, the NCS is able to accurately express the correlations and agree with ED within the stochastic error.

(a)



(b)

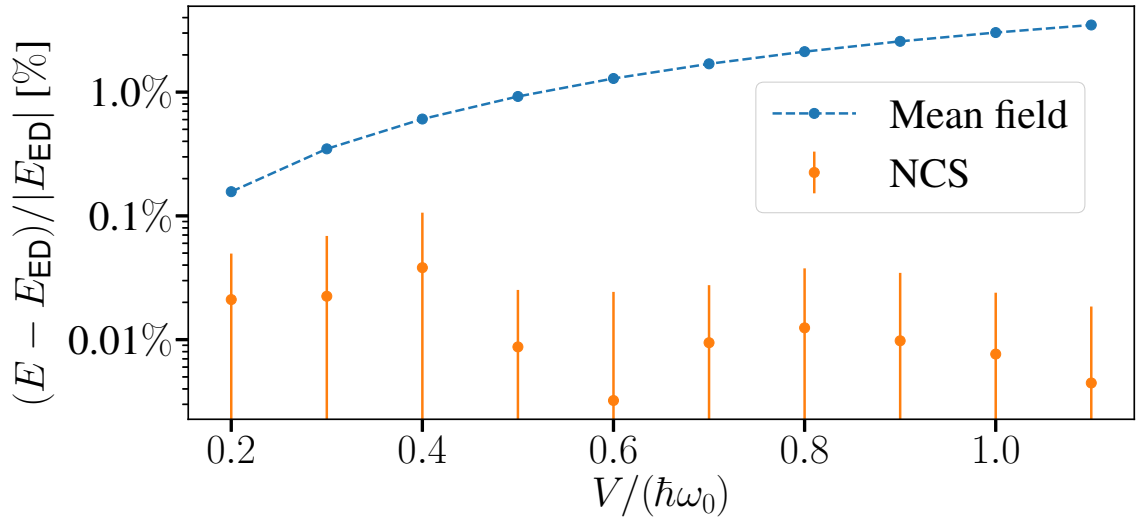


Figure 4.4: The percent difference with respect to ED as a function of: (a) inverse mass in units $1/m_0$, $m_0 = \hbar k_0^2/(2\omega_0)$ (b) impurity-bath coupling $V/\hbar\omega_0$. Our NCS approach (orange) is compared with mean field (blue). Error bars correspond to uncertainty of stochastic estimates of the energy. In panel (b), which features log scale, the error bars stretch infinitely down as the ED result lies within error bars. The dashed lines guide the eye. Figure adapted from Ref. [RLM21].

Next, we study the performance at different impurity-bath couplings V. We fix the inverse mass at $1/m = 0.2 \cdot (1/m_0)$. Then we optimize the same network with 1024 neurons in the hidden layer for different values of V. The percentage deviation from ED, for both NCS and mean-field approach is shown in Fig. 4.4(b). Here we observe consistently good performance and clear advantage over mean-field across all values of impurity-bath coupling $V > 0.1$. Data for $V < 0.1$ is not shown, $E_{MF} - E_{ED} < 10^{-5}$

and numerical errors in the gradients start to dominate the optimization, making it hard to reach more accurate results. We attribute this with a property of the NCS itself. For such small V, the system is very close to the vacuum state. When approaching the coherent state with $|\lambda| = \epsilon \ll 1$, leads to a dominance of states $\lambda_i = \epsilon$ for which $\epsilon^0 = 1$, independent on ϵ. Importantly, our approach easily extends outside the weak-coupling regime, reaching equally accurate results for all V, even in the regime $V \sim 1$.

So far, all results are obtained for high maximum occupation numbers (up to $n_{\max} = 8$) but only a small number of phonon modes. Next we gradually increase the number of k-points to benchmark the ability of the NCS approach to express the correlations between a larger number of bosonic modes, beyond a regime where ED is available. To this end we take k being an equidistant grid between $-k_0$ and k_0 with a varying number of points. The constant impurity-bath interaction potential is $V_k \equiv 0.3$, which corresponds to contact interactions in real space, which is reasonable for one-dimensional systems. We fix the mass at $1/m = 2 \cdot (1/m_0)$. Instead of the total energy, we are now interested in energy per number bosonic modes, to avoid a trivial scaling with the number of modes. In Fig. 4.5 we show the results of a benchmark against the mean-field approach and, where feasible, exact diagonalization. We observe that the difference between the energy

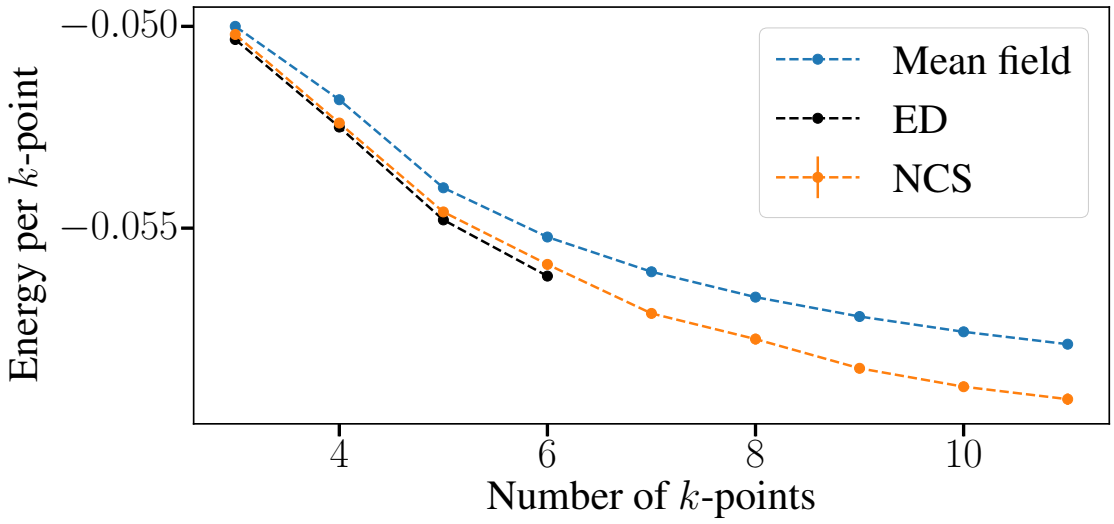


Figure 4.5: The energy divided by number of k-points calculated with the NCS approach as a function of the number of k-points on an equidistant grid between $k = -1$ and $k = 1$. Error bars for NCS approach are smaller than point size. Dashed lines guide the eye. Figure adapted from Ref. [RLM21].

reached using our NCS state and the mean field solution increases with increasing the number of phonon modes, which is consistent with the fact that the amount of modes that are coupled to each other increases as well. Moreover, within the range where ED is feasible, we observe that the NCS results closely follow ED, while the mean-field energy is systematically higher. The number of network parameters in the single hidden layer network with 1024 hidden neurons used for training this system is much smaller (∼2000 times smaller for 11 k-points) than the dimension of the Hilbert space, suggesting great potential to exploit this approach beyond the regime accessible with exact diagonalization.

4.6 Chapter conclusions and outlook

In this Chapter, we introduced a new approach to solve non-additive systems with artificial neural networks. By benchmarking against exact diagonalization and mean field, we obtained accurate results for small systems and all parameter regimes studied. In particular, we were able to capture the challenging intermediate coupling regime at the same accuracy as weak-coupling results, illustrating that this method provides an unbiased approach to strong correlations in non-additive systems.

In our research, we have used a constant impurity-bath interaction potential, corresponding to contact interactions. This renders our work to be more of a proof-of-principle of a novel ML approach, rather then a tool to study the impurity Hamiltonian with a more involved interaction potential. A natural extension would be using the interaction potential from Eq. (1.6). To accurately model such potential, a dense grid of k-points will be needed.

Other natural next steps include benchmarks against other methods for impurity systems and generalizations to other neural network architectures and more complex impurity models, such as the angulon quasiparticle [SL15, Lem17, RL18] which is the rotational counterpart of the polaron. The main complication is the non-commutative SO(3) algebra describing quantum rotations, which is inherently involved in the angulon problem. Some work in similar direction has already been done for spin models [VCN$^+$19], where irreducible representations of SU(2) were considered as inputs for the network. An appealing feature of variational neural network algorithms is their direct extension to unitary quantum dynamics of the system [CT17, CGG18, FM19, SH20, HFM$^+$21]. This requires generalizing the current approach to complex valued network parameters, yielding the possibility of extension of the presented work to the case of impurity dynamics, understanding of which is a subject of intensive ongoing research [ASBn$^+$18, Sch10, GMF$^+$20, SSJ$^+$21, VTHS13, CBC$^+$19].

Our approach is naturally suited for systems whose ground state, in an approximation, can be described as the coherent state. Therefore, it is also suited to studies of arrays of interacting harmonic oscillators. A paradigmatic Hamiltonian of such arrays is

$$\hat{H} = \sum_{r=1}^{n} \left(\frac{\hat{p}_r^2}{2m} + \frac{m\omega^2}{2} \hat{q}_r^2 \right) + \sum_{r=1}^{n} \frac{cm}{2} \left(\hat{q}_r - \hat{q}_{r+1} \right)^2 . \tag{4.24}$$

This Hamiltonian is written for the simplest 1D case, with $\hat{p}_r$ ($\hat{q}_r$) being the momentum (position) operator for the r-th site on the chain. This system has been a subject of studies, for example with mathematical methods [LSVdJ08]. It would be of interest to see how neural coherent states perform in this case, especially with different dimensionalities of the oscillator lattice (1D, 2D, and finally 3D).

CHAPTER 5

Characterizing phase transitions with convolutional neural networks

5.1 Chapter intro

In this Chapter, we study two types of systems. In Section 5.2, we focus on multilayer classical spin models, while in Section 5.3 we study colloidal systems. They are physically very different, as already outlined in Chapter 1. However, they both can be simulated numerically to produce "snapshots" of these systems. For classical spin models, this is done with Monte Carlo algorithms. This results in a set of snapshots, whose distribution matches the Boltzmann distribution at given temperature. Colloidal systems with no fixed lattice can be simulated with molecular dynamics methods. These methods are based on numerically solving Newton equations of motion for particles interacting with given interparticle potential. As a result, one obtains a set of positions for each particle in the simulation.

These snapshots can be essentially treated as images for the purpose of machine learning approaches. The Monte Carlo samples of classical spin model can be expressed as maps of pixels corresponding to individual spins. Each "pixel" can be, for instance, black for spin up and white for spin down. In the case of multilayer models, the layers are equivalent to color channels of images. The molecular dynamics data has to be processed before being treated as images. For instance, one can prepare a 2-D grid with spatial slots – each time one finds a particle in the slot, this pixel gets a given color, say black, each time a particle is not found, a different color can be used, for example white.

For both of them, we use convolutional neural networks. As already signaled in Chapter 2, these networks are perfectly tailored to learn local features of images. The trainable filters, moved over entire images through convolution operation, can then recognize characteristic features, thus distinguishing between different phases of the system.

5.1.1 Motivation behind the choice of systems

Our motivation to study layered spin models is two-fold. First, when multiple models are coupled, new phases may emerge, depending on the strength and the form of the coupling

between them. Let us consider, for instance, two magnetic systems with a tunable coupling between each other. Let us also suppose that when the coupling is zero, each system separately undergoes a conventional ferromagnetic phase transition [Car96]. For nonzero coupling, on the other hand, the order parameter may involve, in the general case, some non-trivial combination of spins of both systems. Let us consider a specific example, the Ashkin-Teller model, introduced in Section 1.3. When the interlayer coupling between the variables σ and τ is zero, the phase diagram of the model is characterized only by the single-layer order parameters $\langle\sigma\rangle$ and $\langle\tau\rangle$. However, when the interlayer coupling is large enough in comparison with the intralayer term, a new non-trivial phase with a composite order parameter $\langle\sigma\tau\rangle$ emerges.

Further examples of such emergence of new order parameters due to the coupling between different layers of the system include the 'metallic superfluid' phase [BSA04, SBP15] and the recently reported BKT-paired phase in two coupled two-dimensional XY models [BDN$^+$19]. At last, let us consider again two square-lattice Ising models with spin variables σ and τ, now coupled via a term of the form $\sigma\tau$. It is interesting whether there is a phase with composite order parameter $\langle\sigma\tau\rangle$. As discussed in literature for the bilayer configurations, such phase is expected *not* to exist.

The phase diagram of the 2D Ashkin-Teller model and of some its variations can be determined analytically [Bax07, DG04]. Together with the Ising model, considered to be a classical workhorse of statistical mechanics [Car96, Mus10], they provide an ideal benchmark to attempt detection of composite order parameters in an unsupervised way.

One could can also ask which new composite order parameters emerge in configurations involving more than two layers, such as the trilayer one. Although in the bilayer Ashkin-Teller model the composite order parameter can be easily recognized, a more complex spin model with several layers, with both short- and long-range interlayer couplings, could be much more challenging to be addressed with simple physical considerations. Many, possibly competing, composite order parameters may be present and determining the one which actually breaks the symmetry and generates a novel phase is a highly non-trivial task. From this point a view, an unsupervised approach able to correctly reproduce the phase diagram of layered models, regardless of the nature of underlying order parameters, is highly desirable.

Our second motivation lies in applicability – layered models tend to emerge in a wide range of realistic physical situations. The most basic bilayer structure, where two two-dimensional systems are coupled, has been studied in a number of cases, ranging from graphene [NGM$^+$04] to ultracold dipolar gases [BDPZ12]. Another major example is the layered superconductors, both naturally occurring and artificially created. The most important natural compounds are compounds of transition-metal dichalcogenides layers intercalated with organic molecules [GDKG70] and cuprates [Tin96]. Examples of artificial structures are alternating layers of graphite and alkali metals [HGM$^+$65] or samples with layers of different metals [RBB80]. Neutral layered superfluids can be engineered with quantum gases by using a one-dimensional optical lattice with ultracold fermions [IFT12] or bosons [CIG07]. Therefore, it is important to develop general approaches capable of dealing with coupled interacting systems. In particular, given the importance of layered physical systems and their ubiquitous presence in a variety of contexts, a general approach to reconstruct their phase diagram would provide an important tool of investigation.

Now let us move to our motivation to study colloidal systems with competing interactions. These competing interactions, discussed in Chapter 1, were shown to appear

between colloidal particles [GE97] or globular proteins [SSC$^+$04b] with a proper balance between the electrostatic and solvent-mediated forces. Theory and computer simulations predict SALR particles to self-assemble into clusters and elongated aggregates. They form either disordered glassy and gel states [STZ05, TSZ09] or periodically ordered morphologies such as fcc ordered clusters, hexagonally ordered columns, double gyroid structure or the lamellar phase [ZZC16, ZC16a, ZC16b, CG10, Arc08, Cia08, DCDGF$^+$06]. Experimentally, however, the ordered structures are observed only for SALR particles adsorbed at two-dimensional (2d) interfaces [GE97, LAS$^+$13, SG99], while in a three-dimensional (3d) bulk the efforts to form ordered structures have been so far unsuccessful [Roy18]. Although the realization of the SALR potential in 3d real space is quite challenging, formation of 3d ordered structures by isotropic particles is of significant interest.

One way to gain stability enhancement of an ordered structure is to confine the system with walls whose geometry fits the symmetries of the ordered pattern. In the case of the SALR system it was shown that as long as the distance between two parallel walls fits to the period of the stripe structure, the stability of the ordered phase is enhanced [APC16, PBC19]. Shorter wall-wall separation corresponds to a higher melting temperature according to a Kelvin-like equation [APC16]. Thus, confinement can effectively suppress thermally induced topological defects, but an accurate choice of the distance between the walls is hard to realize.

Another possible method to induce ordering is to apply shear. In the case of homogeneous systems shear was shown to either suppress [PMSL95, BAFL04] or enhance crystallization [AP88, ART$^+$00] depending on system and conditions [CMB00]. Inhomogeneous systems, such as liquid crystals [LdJ04] or diblock copolymers [HAC$^+$00, NDM$^+$14], when exposed to steady shear were shown to order into layered structures. The focus of the present study is on whether also the inhomogeneous SALR fluid orders when exposed to steady shear. In the case of SALR systems, the first reports on shear effects were provided by Imperio [IRZ08], but only ordered monolayers were sheared. In 3d, the shear effects on SALR systems have been studied only recently [RFGZ19, SR18]. In Ref. [RFGZ19] the authors showed that the equilibrium gel structure after being exposed to steady shear exhibits local, short length-scale anisotropies; however no global ordering was found. In Ref. [SR18], by means of the classical density functional theory, the effects of shear on ordered states were described, showing e.g. shear induced transition between the double gyroid and the cylindrical phases. Here, we describe a 3d SALR system that forms an ordered structure when exposed to steady shear, even above the melting temperature of the ordered phase.

Although SALR systems can spontaneously form a variety of microphases [CPG13, Cia08], we focus on the lamellar phase only. The equilibrium properties of the SALR lamellar phase were intensively studied with a variety of methods at different approximation levels. In 2d bulk the lamellar phase consists of parallel stripes that were shown to melt in a step-wise manner [APC14]. First, the translationally ordered low temperature structure undergoes a transition into a molten lamella phase with only orientational ordering. Further heating results in a transition into a disordered, yet inhomogeneous fluid. This finding from computer simulations agrees with the mean-field calculation obtained in a 1d approximation of the lamellar system that shows coexistence of two ordered phases characterized by different amplitudes of the density profiles [PCA13]. In 3d however, no signs of the step-wise melting of the lamellar phase were reported [ZZC16, IR04, Arc08, DCDGF$^+$06]. We will show here that also in 3d SALR particles

can spontaneously form two distinct structures that are both lamellar.

Another question that motivates us to study colloidal systems, is how far the level of description complexity can be reduced so that different 3d structures can still be distinguished and transitions between them localized. To do that, instead of analyzing 3d particle configurations, we will use binary matrices obtained by mapping the 3d structure onto a 2d discretized surface. The maps will be then utilized as input data for different machine-learning based methods aimed at pattern recognition.

5.1.2 Motivations behind the choice of methods

In layered models one has a certain degree of freedom in the way the MC data to be analyzed are fed to the neural networks. Possible ways could be:

1. learning from each layer independently,
2. learning from any arbitrarily chosen combination of the layers (e.g. product, sum, sum of squares, etc.). This choice would need to result from physical intuition about the system,
3. the most challenging case, learning from a multilayer model directly, treating it similarly to a color image with all layers fed into the nwtwork simultaneously

Method 1 would be entirely ineffective in detecting composite order parameters, i.e. parameters defined across multiple layers of a layered system. To use the Ashkin-Teller model as an example, it is known to have an order parameter defined as the product of spins from both layers. Method 2 requires prior knowledge of the physics of the system. For known paradigmatic models such as the Ashkin-Teller models, the order parameters are known, however, this is not the case for other models. We, in this Chapter, would like to test if method 3 will succeed on the challenging multi-layer models.

Another question that will be addressed in this Chapter, is how far the level of description complexity can be reduced so that different 3d structures can still be distinguished and transitions between them localized. To do that, instead of analyzing 3d particle configurations, we will use binary matrices obtained by mapping the 3d structure onto a 2d discretized surface. The maps will be then utilized as input data for different machine-learning based methods aimed at pattern recognition. It is interesting from the machine learning point of view if two-dimensional maps can be still used for learning structures that inherently occur in three dimensions.

5.1.3 Authors' contributions

In Ref. [RDC+20], all authors developed the theoretical concept of the study. GB performed the Monte Carlo simulations. I implemented the convolutional neural network. All authors analysed the the results.

In Ref. [PRP20], JP conceived the project. I designed the artificial neural networks approach. I performed the artificial neural networks experiments together with JP. JP generated all samples with molecular dynamics and performed the Principal Component Analysis. AP provided guidance throughout the project.

5.2 Identification of phase transitions in multi-layer lattice models

5.2.1 Machine learning phase transitions in classical spin models

Let us consider a general case of a spin system whose Hamiltonian is defined by two parameters, J and K. We aim to devise a procedure to depict the phase diagram in the $K - J$ plane. To this extent we discretize a portion of the $K - J$ plane on a grid with steps ΔJ and ΔK. For each point on the grid we generate a number of uncorrelated MC snapshots using standard algorithms [SW87, SS96, Wol89]. Unless otherwise specified we shall work on a $32 \times 32 \times N_l$ square lattice, N_l being the number of layers to be specified later, and we shall generate a number of 600 snapshots for each point in the phase diagram. Periodic boundary conditions are used on each layer throughout all the simulations.

The training of the convolutional neural network attempts at learning to distinguish snapshots belonging to the two different points, (J_1, K_1) and (J_2, K_2), in the phase diagram. Intuitively, when this training fails, the two points present nearly identical features, thus belonging to the same phase. On the other hand, if it succeeds, the two points should belong to two different phases. In order to carry out this plan, at first, we divide the data in a standard way, taking 80% of snapshots from each of the two points as training data, while keeping the other 20% as validation data. Then, we train the network on the training data and quantify the classification accuracy on the validation set as the fraction φ of correctly labeled examples from the validation set. Based on that, we introduce the following quasidistance between the two phase diagram points (J_1, K_1) and (J_2, K_2):

$$d((J_1, K_1), (J_2, K_2)) = 2(\varphi - 0.5)\Theta(\varphi - 0.5) \ , \tag{5.1}$$

where $\Theta(x)$ is the Heavyside step function, preventing d from assuming negative values. Then perfect discrimination $\varphi = 1$ (signaling different phases) corresponds to $d = 1$, while perfect confusion $\varphi = 0.5$ (signaling the same phase) corresponds to $d = 0$. We use the term 'quasidistance' since it does not respect triangular inequality. However, this fact plays no role as far as all the applications in the present Chapter are concerned.

We feed the raw Monte Carlo snapshots directly to the convolutional neural network, with spin down encoded as 0 and spin up encoded as 1, no preprocessing applied. The network architecture is optimized for the task of classifying two phases: after convolutional and fully connected layers the final layer consists of two softmax output neurons outputting the labels. The convolutional filters span both layers, which is the feature enabling the network to learn composite order parameters. Hence, both layers are simultaneously fed into the network.

We used the following architecture of the network. The first layer is a convolutional layer with 32 filters of size 2 by 2 and unit stride in both directions. Then the 'max pooling' [SMB10] operation is applied with pool size 3 by 3, stride 2 in both directions and same padding. The results is then fully connected to a hidden layer with 100 neurons. The binary classification is finally done in the output softmax layer with two neurons. Both the convolutional and hidden fully connected layers are activated by rectified linear units (ReLU) [NH10]. The network is visualized in Fig. 5.1. We train the network by minimizing the cross-entropy using the Adam [KB14] (described in Section 4.4.2) adaptive optimization algorithm with 7 epochs and minibatch size 25. Such choice leads to a fast

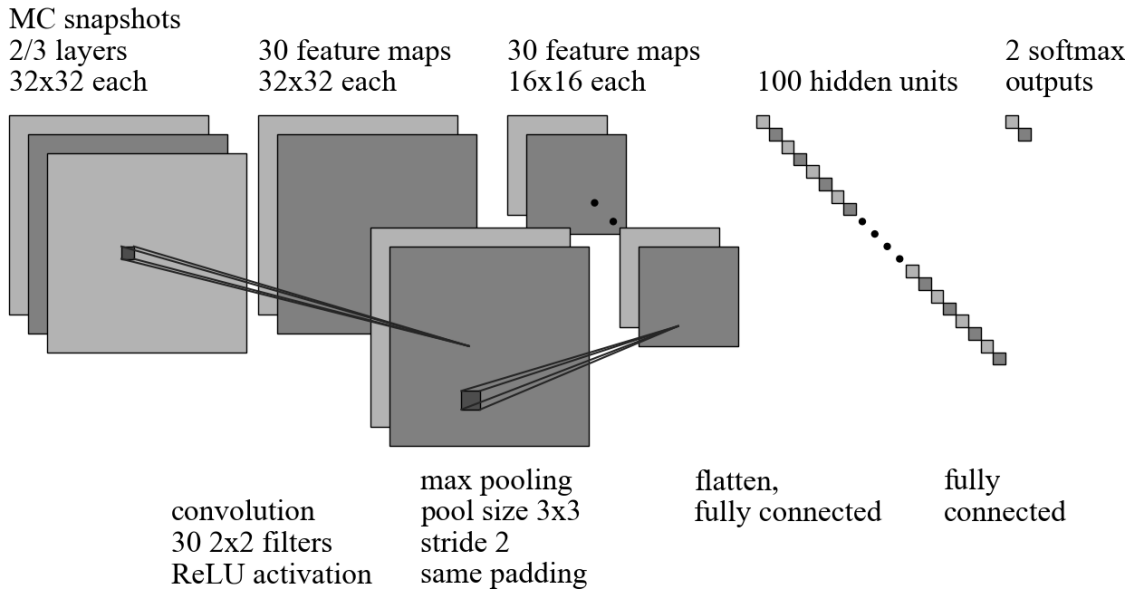


Figure 5.1: Visualization of the convolutional neural network used. Lower labels describe the layer operations. Upper labels describe the shapes of tensors before and after each operation. Figure adapted from Ref. [RDC+20].

training – the amount of training is much lower than in computer vision applications, routinely requiring hundreds or thousands of epochs – as well as prevention of overfitting by early stopping, hence eliminating the need for other measures such as dropout [HSK+12]. We use the following Adam algorithm parameters: learning rate 0.001 and standard choices of $\beta_1 = 0.9$ and $\beta_2 = 0.999$. We use Tensorflow [ABC+16] for the implementation.

At last, we make use of the distances defined in Eq. (5.1) to construct a field $u(J, K)$ defined on the phase diagram through its finite-difference lattice gradient

$$\nabla u(J,K) = \begin{pmatrix} (u(J+\Delta J, K) - u(J,K))/\Delta J \\ (u(J, K+\Delta K) - u(J,K))/\Delta K \end{pmatrix} \equiv \begin{pmatrix} d((J+\Delta J, K),(J,K))/\Delta J \\ d((J, K+\Delta K),(J,K))/\Delta K \end{pmatrix} . \tag{5.2}$$

Clearly ∇u will be constant in regions of the phase diagram belonging to the same phase, since we expect that the difficulty of telling first neighbors apart should be uniformly quite high. On the other hand, we expect the value of ∇u to abruptly change in the vicinity of a phase transition, suggesting that the phase diagram should be naturally characterized by looking at the finite-difference lattice Laplacian

$$\begin{aligned} \nabla^2 u(J,K) \approx & \frac{1}{(\Delta J)^2} \sum_{i=0}^{n} (-1)^i \binom{n}{i} u(J + (n/2-i)\Delta J, K) \\ & + \frac{1}{(\Delta K)^2} \sum_{i=0}^{n} (-1)^i \binom{n}{i} u(J, K + (n/2-i)\Delta K)) , \end{aligned} \tag{5.3}$$

with the $n = 2$, $n = 3$ and $n = 4$ cases corresponding to a 5-point, 9-point or 13-point stencil, respectively. The stencil includes $(n-1)$ nearest neighbors in the J and K directions. We stress that the summations can be rearranged so that they involve only differences of the u field evaluated between first, second and third neighbors, that can in turn be expressed in terms of the quasidistance d. From the discussion above, it is clear that a sudden rise in the value of $\nabla^2 u$ means that the CNN can distinguish with increased precision arbitrarily close points in the phase diagram, thus signaling a phase

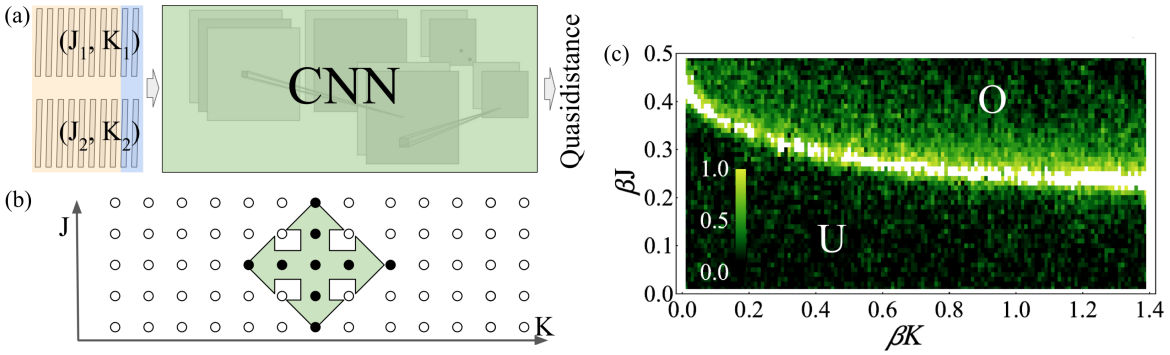


Figure 5.2: An overview of the proposed method. (a) The convolutional neural network is able to assess quasidistance which determines whether the phase diagram points (J_1, K_1) and (J_2, K_2) belong to the same or different phase. This is done by attempting to learn to distinguish between individual Monte Carlo snapshots with orange and blue visualizing training and validation MC snapshots, respectively, see main text. (b) Using distances between first, second and third neighbors, one can evaluate the Laplacian across the phase diagram. (c) Large values of the Laplacian signal the presence of phase transition. Plotting the Laplacian reconstructs the phase diagram, which is now parametrized by the dimensionless combinations βJ and βK, with β the inverse temperature. Here we show the reconstructed phase diagram for the square-lattice Ising bilayer model with the transition between ordered (O) and unordered (U) phases, see main text. Figure adapted from Ref. [RDC+20].

transition. We anticipate that including high-order finite-differences besides the obvious 5-point stencil taking into account first-neighbors stencil considerably increases the quality of the reconstructed phase diagram. This point will be analyzed in detail later. Moreover, using the stencil as opposed to always just comparing two neighboring points of the phase diagram immunizes the algorithm in the case of very dense grid. In such a case, it would be progressively difficult to find neighboring points belonging to different phases. With our approach, we are assured that using a large enough stencil will circumvent this problem for any grid density.

Calculation of $\nabla^2 u(J, K)$ for the entire phase diagram is by far the most time-consuming step of the algorithm. Using N nearest-neighbours, i.e. $[4(N-1)+1]$-point stencil, it requires $M \cdot 4(N-1)$ calculations of the quasidistance. There,

$$M = \frac{J_{\max} - J_{\min}}{\Delta J} \cdot \frac{K_{\max} - K_{\min}}{\Delta K} \tag{5.4}$$

is the total number of discretized (J, K) pairs in the phase diagram.

In conclusion of the present Section, we compare our scheme with other related approaches. As opposed to other machine learning schemes, in the present work we do not need the evaluation of any observable quantity to establish a distance [BAT17], rather directly relying on the MC snapshots. Moreover, as opposed to other approaches [BAT17] the scheme we introduce in this Chapter fully takes advantage of the two-dimensional nature of a two-parameter phase diagram, as the local information is reconstructed by taking into account neighbours in all directions. Extensions to three- or higher-dimensional phase diagrams are straightforward [TAAS13]. Finally, our approach requires only the evaluation of a fixed number of neighbors for each point in the phase diagram, ensuring

that the computational effort required for training scales linearly with the number of points in the discretized phase diagram.

5.2.2 Multilayer Ising models

We now use the framework described in the previous Section to characterize the phase diagram of different coupled spin models.

Let us start from a bilayer Ising system, simplest multilayer system described in Section 1.3. Its Hamiltonian features a quadratic coupling term (sometimes referred to as the Yukawa coupling):

$$H_{\text{bilayer}} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - J \sum_{\langle ij \rangle} \tau_i \tau_j - K \sum_i \sigma_i \tau_i \ , \tag{5.5}$$

where $\sigma_i, \tau_i = \pm 1$ are Ising variables on a two-dimensional square lattices, whose sites are denoted by the indices i, j. The sums in Eq. (5.5) are over nearest-neighbor sites. When $K = 0$, the system reduces to two uncoupled Ising models, having a phase transition at the Onsager critical point $(\beta J)_c = \ln\left(1 + \sqrt{2}\right)/2$ [Ons44, Mus10], β being the inverse temperature. This critical point is shifted by the presence of a finite interlayer coupling K. The resulting Ising critical line separating the paramagnetic and ferromagnetic phases as a function of K has been studied in the literature [OE75, HLIM93, BOST95]. It is clear that the bilayer system (5.5) is the classical counterpart of two coupled quantum Ising chains in a transverse field, a system that has been studied both in relation to its spectrum, phase transitions and possibility to determine an integrable line in the space of parameters [DM98, FGN00, Tsv07, KA09]. The classical bilayer system and the quantum coupled chains can be also related to each other by an exact mapping.

From our point of view the model described by Eq. (5.5) is an excellent starting point for our investigations, especially in order to check the existence of a composite order parameter and its relation to the phase diagram. It is now natural to parametrize the phase diagram in term of the dimensionless combinations βJ and βK, discretizing it for values of $\beta J \in [0, 0.5]$ and $\beta K \in [0, 1.4]$, with discretization steps $\Delta\beta J = \Delta\beta K = 0.01$. We then apply the phase diagram reconstruction procedure described in the previous Section to precisely determine the phase boundaries in the βK-βJ phase diagram, shown in Fig. 5.2(c). The phase transition occurs at $(\beta J)_c \approx 0.44$ in the uncoupled $\beta K = 0$ case, in agreement with analytical results [Ons44, Mus10]. The errors of our method on the determination of transition points are discussed in Section 5.2.4. Then the critical temperature gradually decreases to the strong-coupling critical temperature $(\beta J)'_c = (\beta J)_c/2$. The width of the peak is essentially due to the the finite-size ($32 \times 32 \times 2$) of the lattice used for Monte Carlo simulations, whose snapshots we feed to the neural network. The result is that it appears that only two phases are found, with order parameter $\langle \sigma \rangle = \langle \tau \rangle$. From our treatment of data we cannot determine the behavior of the order parameter inside the two phases, whose study would be an interesting future continuation of the present results.

Next, we consider a trilayer system, whose Hamiltonian is a natural extension of the one of Eq. (5.5). This Hamiltonian, already introduced in Section 1.3, is given by:

$$H_{\text{trilayer}} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - J \sum_{\langle ij \rangle} \tau_i \tau_j - J \sum_{\langle ij \rangle} \upsilon_i \upsilon_j - K \sum_i \sigma_i \tau_i - K \sum_i \tau_i \upsilon_i \ , \tag{5.6}$$

and the new variable υ_i is also an Ising spin. This is the first non-trivial example, and of course representative of properties of the multilayer Ising model with Yukawa coupling.

The central natural question is whether a composite order parameter emerges. Moreover the model of Eq. (5.6) is interesting since it paves the way to the investigation of the N layers case, which shall be trivial with the method presented here. Indeed the N layer case may serve to investigate how the (three-dimensional) limit of infinite layers is retrieved, an issue in the context of layered models, see *e.g.* Ref. [SS98].

The investigation of the model described by Eq. (5.6) follows the same line as the one of the bilayer case, we are able to reconstruct the phase diagram as shown in Fig. 5.3, recovering that strong-interlayer-coupling critical temperature that in this case is $(\beta J)_c'' = (\beta J)_c/3$, marked by a red dashed line. The main result exhibited in Fig. 5.3 is that no composite order parameter appears even for the trilayer case. Therefore, our technique has been able to correctly recover the phase diagram of the bilayer Ising model, where we do not expect any additional order to appear [SSBS05, SB16], while it also predicts the same picture for the trilayer case, for which no previous expectation exist up to our knowledge. The generalization to the N-layer case shall be straightforward, but more numerically demanding, while based on the present results no additional phases are expected to appear. Therefore, in the following we are going to investigate a different case where a composite order parameter appears by construction.

5.2.3 Reconstructing composite order parameters: the Ashkin-Teller model

We now turn to the square-lattice Ashkin-Teller model. We introduced it in Section 1.3. It is described by the following Hamiltonian

$$H_{\mathrm{AT}} = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - J \sum_{\langle ij \rangle} \tau_i \tau_j - K \sum_{\langle ij \rangle} \sigma_i \sigma_j \tau_i \tau_j \tag{5.7}$$

with $\sigma_i, \tau_i = \pm 1$. Compared to Hamiltonians (5.5)-(5.6) one sees that the coupling is now quartic in spins. Since in the Ising model there are only two scaling fields relevant in renormalization group sense [Car96, Mus10], the magnetization and the energy, one sees that in the models (5.5) and (5.7) one has basically the two natural ways of having respectively magnetization-magnetization and energy-energy couplings, higher order coupling terms being irrelevant. The Ashkin-Teller model is also related to the four state planar Potts model, and several variations of it, also in three dimensions, have been investigated [Wu82].

The Ashkin-Teller model features a rich phase diagram, and remarkably in two dimensions can be studied analytically [Bax07, DG04]. Here we consider the case of ferromagnetic couplings, $J, K \geq 0$. It is known that three different phases exist [Bax07]. Besides an ordered phase, denoted by I, characterized by $\langle \sigma \rangle \neq 0 \neq \langle \tau \rangle$ and a disordered phase, II, characterized by $\langle \sigma \rangle = \langle \tau \rangle = 0$ one also finds the peculiar phase III in which the single spins σ and τ are disordered, whereas a composite order parameter given by their product is ferromagnetically ordered, i.e. $\langle \sigma \tau \rangle \neq 0$.

Whereas the previous investigation of Ising-like models makes us confident that the ML procedure we have introduced is able to correctly characterize the transition between phase I and phase II, it is not *a priori* clear that phase III can be correctly identified. As shown in the small inset of Fig. 5.4, MC snapshots show disordered spins both in phase II and in phase III, the transition being determined by the $\sigma\tau$ composite variable, that we do not directly feed to the CNN. In order to learn the existence of the II-III phase

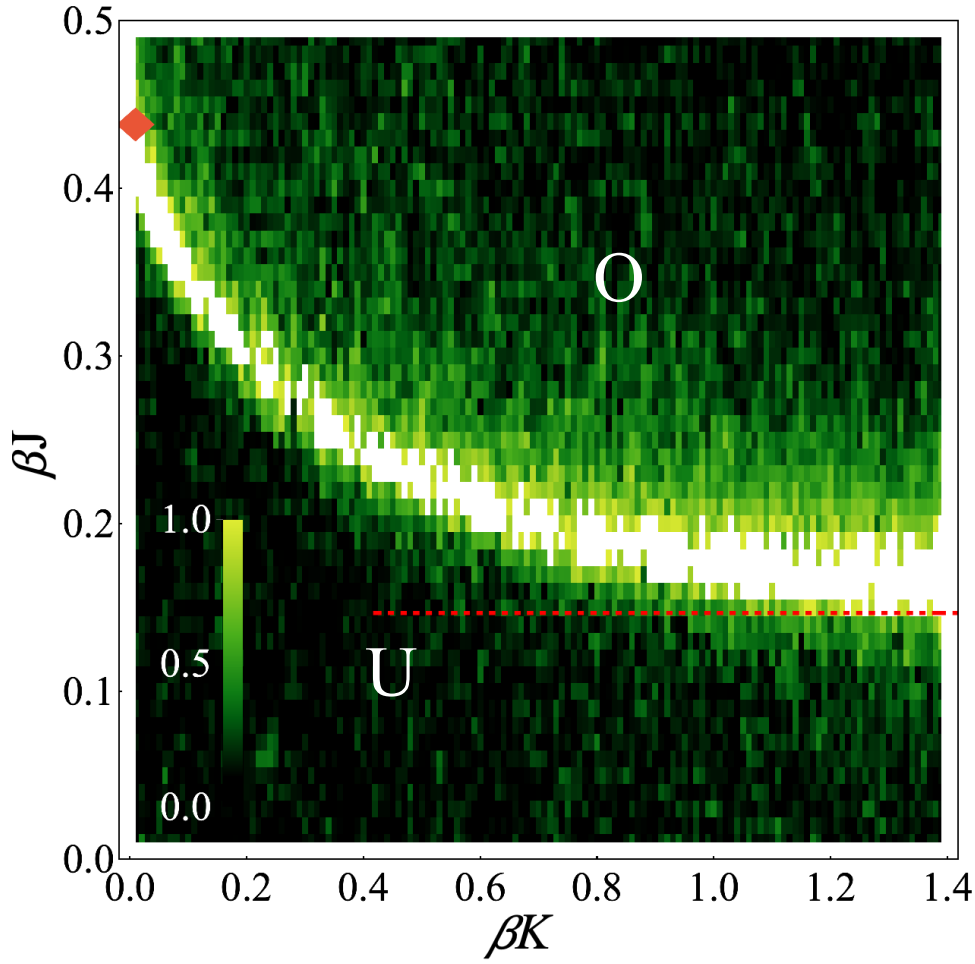


Figure 5.3: Reconstructed phase diagram for the square-lattice Ising trilayer model, showing a phase transition between an unordered, high-temperature phase (U) to an ordered, low-temperature phase (O). Note that as the interlayer interaction βK is increased, the critical temperature decreases from the analytic limit $(\beta J)_c = \ln\left(1+\sqrt{2}\right)/2 \approx 0.44$, marked by a red diamond, to the strong-interlayer-coupling limit $\beta K \to \infty$ where $(\beta J)'_c = (\beta J)_c/3$, marked by a red, dashed line. Figure adapted from Ref. [RDC+20].

transition the CNN must learn to reconstruct the composite order parameter. We find that our framework successfully performs this task, owing to the convolutional filters which are convolved in 2D spanning across the layers and are able to learn even elusive interlayer correlations.

The reconstructed phase diagram of Fig. 5.4 shows that indeed our approach is able to correctly learn the phase transitions in the ferromagnetic Ashkin-Teller model. Whereas the transition line corresponding to the magnetization of σ and τ, as separated variables, corresponds to a prominent peak, whose width is essentially determined by finite-size effects, the line corresponding to the magnetization of the composite $\sigma\tau$ order parameter corresponds to a smaller peak, displaying that the characterization of this transition line is more demanding to the CNN, but still possible.

We can compare the obtained phase diagram we obtain with exact results available in the literature. In the $K \to 0$ the model reduces to a square-lattice Ising model with coupling constant J, with critical temperature $(\beta J)_c = \ln\left(1+\sqrt{2}\right)/2 \approx 0.44$ [Ons44, Mus10], whereas in the $K \to \infty$ limit the model reduces to a square-lattice Ising model with

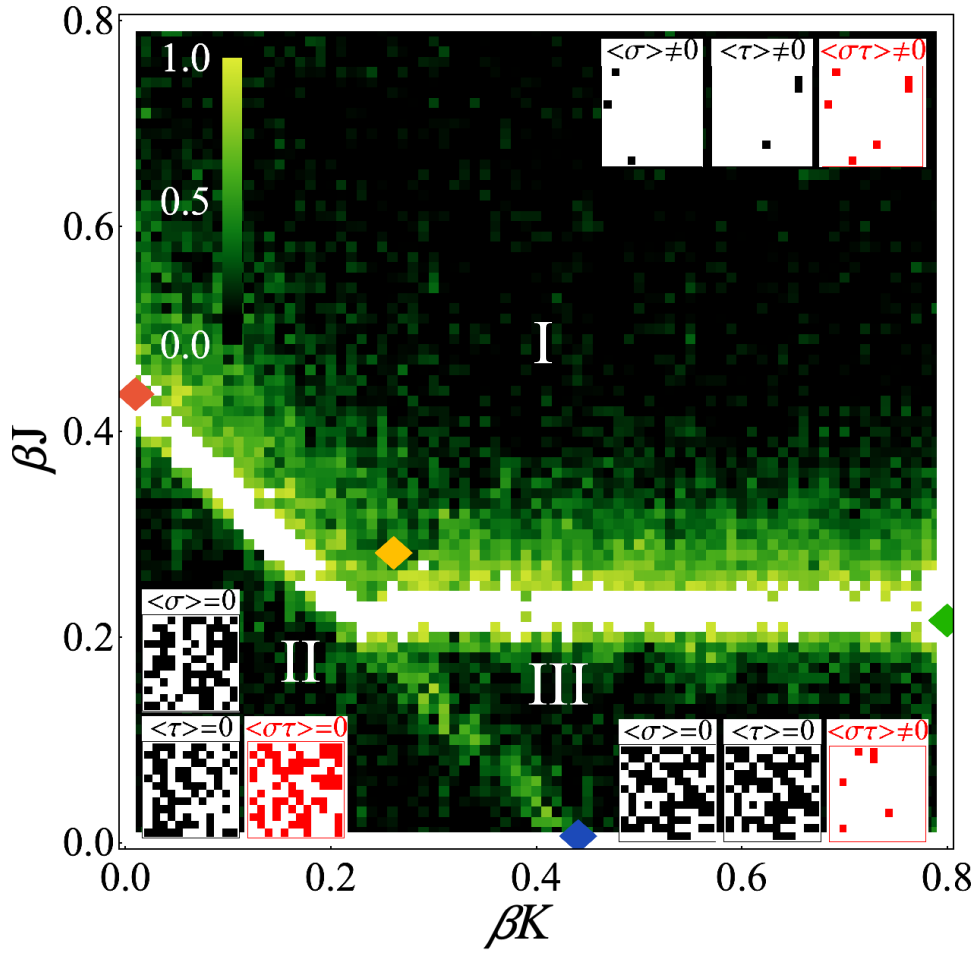


Figure 5.4: Reconstructed phase diagram for the square-lattice Ashkin-Teller model: the red, blue yellow and green diamonds show analytically-determined phase transition points, see main text. Our approach identifies three phases, in agreement with the theory of the Ashkin-Teller model. The insets show representative configurations of the σ, τ spins and of the 'composite' spin $\sigma\tau$, for each phase: note that the transition between phase II and phase III does not correspond to any apparent difference in the σ and τ layers that we feed to the CNN. We stress that the $\sigma\tau$ 'composite' variable, marked in red, is not fed to the CNN. Figure adapted from Ref. [RDC+20].

coupling constant $2J$ and critical temperature $(\beta J)'_c = \ln\left(1+\sqrt{2}\right)/4 \approx 0.22$. Finally for $J = 0$ the system again undergoes an Ising-like phase transition for the composite order parameter, at $(\beta K)_c = \ln\left(1+\sqrt{2}\right)/2 \approx 0.44$. These three points are marked by a red, green and blue diamond, respectively, in the phase diagram of Fig. 5.4, showing an excellent agreement between the analytical results and the reconstructed phase diagram, even in the latter case when the composite order parameter $\sigma\tau$ drives the transition. Finally, the yellow diamond marks the bifurcation point as determined analytically in Ref. [Bax07]; we attribute the difference with respect to the bifurcation point in our reconstructed phase diagram to finite size effects. We also mention that the critical lines separating the different phases are retrieved with a precision up to $\sim 20\% - 30\%$, except for vanishing βJ. Again we attribute this to finite size effects; proceeding as extensively discussed in the literature [CCC+19] one could obtain a quantitative agreement on the location of the critical lines. Here, our emphasis is on the possibility of retrieving the phases with composite order parameters and to ascertain their existence, as we also did

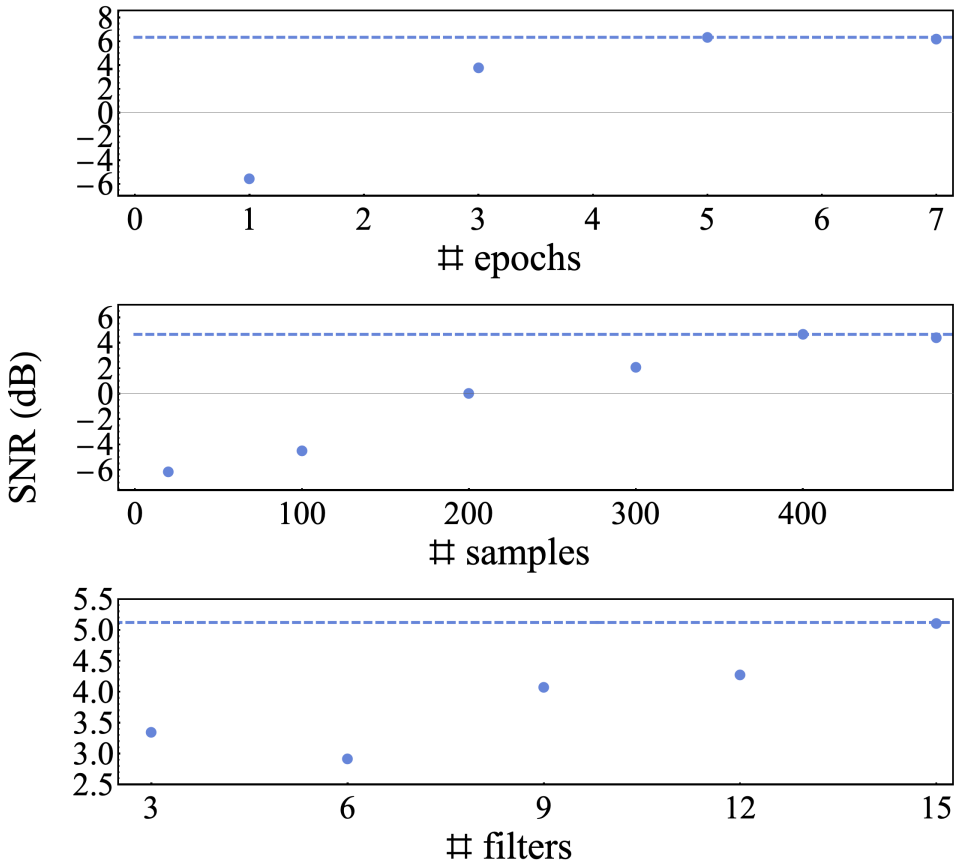


Figure 5.5: Signal-to-noise ratio for the Ising bilayer as a function of the number of epochs (upper panel), of the number of samples in the training set (middle panel) and of the number of convolutional filters (lower panel). The dashed lines guide the eye towards the highest attainable signal-to-noise ratio in each dataset. Figure adapted from Ref. [RDC+20].

for the trilayer Ising model.

5.2.4 Scaling properties and robustness of the approach

Our results show that with the network and learning parameters that we used we were able to obtain a phase diagram of quality high enough to visually identify different phases. In addition, in this Section we characterize our method by quantifying signal to noise ratio (SNR) and studying its behavior when essential parameters are changed. We define the SNR as

$$\mathrm{SNR} \equiv \log_{10}\left(\frac{\frac{1}{N}\sum_i (x_i - \nu)^2}{\nu^2}\right) , \tag{5.8}$$

x_i being the values of the $\nabla^2 u$ field of Eq. (5.3), the summation extending over a region containing N values, ν being the 'noise', i.e. the average value of $\nabla^2 u$ in a subset of the region far away from a phase transition. We evaluate the SNR for the Ising bilayer on a strip centered on $\beta K = 1.1$, exhibiting a sharp phase transition at $\beta J \approx 0.26$ as clear from Fig. 5.2. At first, we vary the number of training epochs, observing that the SNR is rapidly increasing before reaching a maximum value at around 5 epochs of training. This indicates that further training brings no benefit while providing a risk of overfitting, justifying our early-stopping approach. Secondly, we vary the number of samples in the

training set, showing a rapid increase in the SNR before reaching a plateau at about 400 samples, justifying our choice of using a slightly larger number (600) of samples in the training set. Lastly, we vary the number of convolutional filters in the CNN. Again, the general upwards trend shows that a larger number of convolutional filters helps in enhancing the quality of the reconstructed phase diagrams. However, we stress that in this latter case the SNR is quite high in the whole parameter region we consider. The lowest number of convolutional filters we consider (3) is already enough for achieving a good reconstruction of the phase diagram and a large SNR value. These analyses are shown in Fig. 5.5.

We have also analysed how the reconstructed phase transition is affected by the dimension of the stencil in Eq. (5.3). Using a 5-point, 9-point or 13-point stencil we have obtained SNR values of -1.36 dB, 0.38 dB and 3.88 dB, respectively. This confirms that the approach we are introducing takes indeed great advantage from the two-dimensional structure of the phase diagram, and information from second- and third-nearest neighbors is being used to sharply characterize the phase transition.

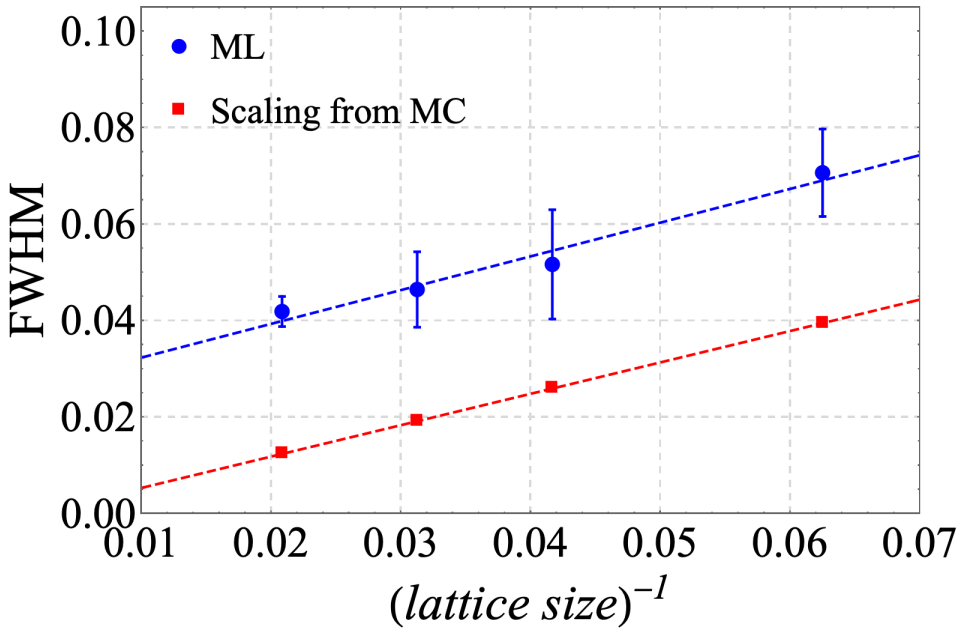


Figure 5.6: Full width at half maximum of the phase transition in the Ashkin-Teller model for $\beta K = 0.7$, as a function of the inverse lattice size, for four different lattice sizes ($L = 16, 24, 32, 48$). The red squares show the FWHM of the peak in magnetic susceptibility in the MC data, whereas the blue squares show the FWHM of the Laplacian peak obtain from our ML approach; errors are estimated by identifying the phase transition 10 times. The dashed lines guide the eye. Figure adapted from Ref. [RDC+20].

We also studied the width of the identified phase transitions. These phase transitions in the systems we consider have a certain 'natural' width, due to the finite size of the lattice in the underlying Monte Carlo simulations; moreover, we expect our approach to introduce an additional width when determining the transition point. In order to verify this assumption and to investigate the accuracy of our method, we analyzed the natural width associated to the phase transition in the Ashkin-Teller model for $\beta K = 0.7$, determining it by looking at the peak of magnetic susceptibility directly from Monte Carlo simulations, and determining its width through the full width at half maximum (FWHM). We compare it with the FWHM of the Laplacian peak we reconstruct from our machine

learning approach. The results are shown in Fig. 5.6; the FWHM of both the magnetic susceptibility (red squares, the red dotted line guides the eyes) and machine learning Laplacian (blue squares, the red blue line guides the eye) obey the same $\propto 1/L$ scaling with respect to the lattice size L. The constant offset between the two datasets can be understood, as anticipated, as the additional error introduced of our method, due to the discretization of the parameter space, and due to some intrinsic uncertainty associated to the machine learning procedure.

5.3 Characterization of phases in colloidal systems

5.3.1 Model and methods

Model

We used the SALR potential of the form previously studied in Refs.[SMZT04, STZ05, MLKB14, TSZ09, SPP17]. We already described the SALR potential in Section 1.4. The effective interparticle interaction is the sum of the standard Lennard-Jones and Yukawa potentials, which we just remind here for the convenience of the Reader:

$$V(r) = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] + \frac{A}{r} e^{-r\kappa}. \tag{5.9}$$

The parameter values were taken from [SPP17]: $A = 0.5$, $\kappa = 0.5$, $\varepsilon = 1.0$, $\sigma = 1.0$, where ε and σ are set to be the units of energy and length respectively. The plot of this potential was presented in Fig. 1.3. Accordingly, in reduced units temperature is $k_B T/\varepsilon$, where k_B is the Boltzmann constant, and time is given in units of $\tau = \sqrt{\frac{m\sigma^2}{\varepsilon}}$, where $m = 1$ is the mass scale. For this set of parameters, particle aggregation into clusters with preferred size occurs up to $k_B T/\varepsilon \approx 0.45$, even though the potential is found to be repulsion-dominated (positive second virial coefficient) for $k_B T/\varepsilon > 0.31$ [SPP17]. Apart from a detailed description of the low density behavior, for the chosen set of parameters the phase behavior of the system is not known precisely, but can be inferred by analogy to similar systems [ZZC16, Cia08, CPG13, DCDGF$^+$06].

Details of the simulation

In order to minimize size effects and to ensure commensurability between box size and the period of the formed structures the molecular dynamic simulations (MD) were performed in the isobaric-isothermal ensemble, NPT, where N is the number of particles, P is the pressure and T is the temperature. We used $N = 15625$ and $P = 0.5\varepsilon/\sigma^3$.

Equilibrium simulations were run using the HOOMD-blue package [GNA$^+$15, ALT08] with a time step $dt = 0.005\tau$. The runs consist of 10^8 MD steps with either linearly increasing or decreasing temperature in the range of $0.01 < k_B T/\varepsilon < 0.8$. For decreasing temperature runs the initial configuration was a simple cubic lattice, and the final configuration of this procedure was later used to initialize the linear heating protocol. For temperatures away from the phase transitions both heating and cooling protocols gave the same number densities of the system, thus we can infer that the equilibrium was reached.

The nonequilibrium MD simulations were performed in LAMMPS [Pli95] adjusted to use the SLLOD equation of motion in the NPT ensemble [TD17]. The shearing procedure was initialized from the equilibrium structures, then after the 10^6 steps needed to reach

the stationary state, production runs consisting of 10^6 steps were executed. The shear was imposed in the x direction, and the velocity gradient was in the y direction. Therefore, the shear rate is defined as $R = v_x/y$ and is measured in inverse time units, τ^{-1}.

Details of the analysis

Dynamical properties of the equilibrium structures were described by computing components of the average displacement vector, $D = (d_x, d_y, d_z)$, given by

$$d_k = \frac{1}{N}\sum_{i=1}^{N}[R_i(t_0 + \Delta t) - R_i(t_0)], \quad k = x, y, z, \tag{5.10}$$

where $R_i(t)$ is the i-th particle location at time t, and $\Delta t = 10^4 dt$. The calculations were performed after rotating the structure so that

$$d_\parallel = \frac{d_x + d_y}{2}, \quad d_\perp = d_y \tag{5.11}$$

describe the displacement in directions parallel and perpendicular to the plane of the slabs respectively. For the temperature dependence analysis, time averages of d_k were calculated over 10^2 consecutive configurations of the cooling simulation protocol.

The algorithms were fed with the same data, that is a 2d map of the 3d structure obtained by perpendicular projection of particle positions on a 100×100 grid located at the $x = 0$ plane, i.e. the plane perpendicular to the flow direction. The normalized map was then transformed to a binary array, M, by the ceiling function which turns every nonzero value to 1. As a result, the only information contained in each element of M is whether at a given time-step there is a particle with y and z coordinates within a grid node regardless of its x coordinate value. An example of how a lamellar structure is projected on 2d binary matrix is shown in Fig. 5.7.

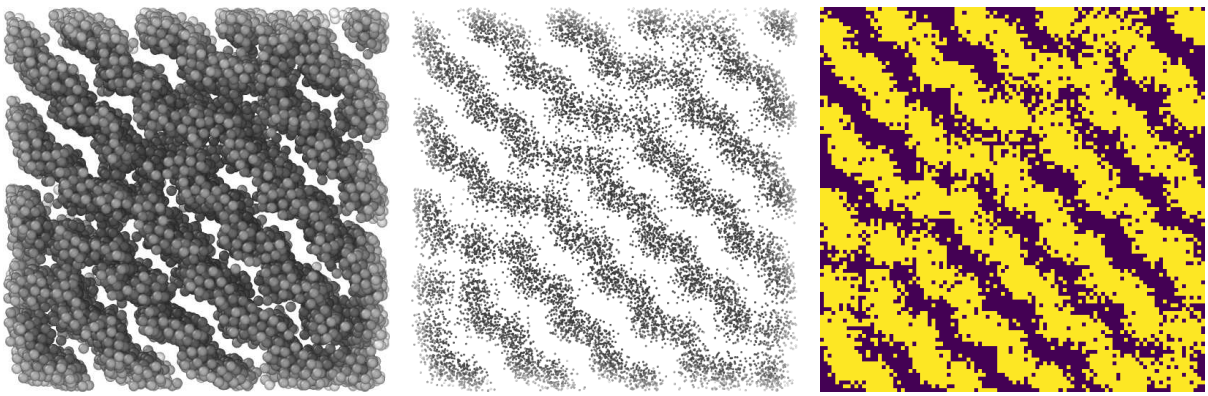


Figure 5.7: Stages of binary map production. Left panel: processed particle configuration seen from the top view of the (y,z) plane, perpendicular to the shear flow. Middle panel: spherical particles transformed to point particles. Right panel: (y,z) positions of the point particles were binned into 100x100 grid. The grid was then binarized depending on whether the bin was occupied (yellow color) or not not (dark color). Reproduced from Ref. [PRP20], with the permission of AIP Publishing.

Artificial neural networks [Dan13] are commonly used for data classification based on features that they were trained to distinguish. They are a powerful tool for encoding

functions that describe a set of features (e.g. Monte Carlo simulation, image data) into a set of classes (e.g. ordered and disordered phase). This encoding is achieved by decomposing the function into a set of small units, called neurons. The behavior of individual neurons depends on a set of trainable parameters that are gradually adjusted during the training phase. The training involves minimization of a cost function that quantifies the difference between the ground truth labels and current neural network output over the set of the neural network parameters. CNNs are an example of networks that are designed for pattern recognition based on the translationally invariant two-dimensional features of their inputs [GWK+18, DV19].

On the contrary, DNNs do not convolve the input with a set of filters; the input is instead densely connected to the first hidden layer of neurons. Hence, in that case the preparation of data for training includes flattening, a transformation from 2d matrices to 1d vectors. Thus in the case of DNN, the learning relies on finding correlations between artificially produced vectors. We have used both types of networks to categorize the binary maps, M. For both networks we used maps obtained at $k_BT/\varepsilon = 0.42$ with $R\tau = 0.01, 0.1, 30$ (low and high shear disordered structure), $R\tau = 6, 8, 10$ (ordered lamella) and $R\tau = 16, 20, 24$ (torn lamella). For each value of R we used 400 maps for training and the rest (100 maps) for validation. The DNN features a multilayer perceptron architecture with the input layer being followed by a 32-neuron hidden layer activated with ReLU activation [NH10]. The output layer is made of three softmax-axtivated neurons. The CNN consists of two convolutional layers followed by a softmax-activated output layer. The convolutional layers feature 32 (input) and 16 (hidden) filters of size 3x3. Both networks were trained using the Adam algorithm with the aim to minimize the cross entropy between the softmax output and ground truth labels. This algorithm was described already in Section 4.4.2. Despite the fairly simple and shallow architecture, for both sets we reached a classification accuracy of 1.0 and cross entropy loss function values of 0.0008 for the training set and 0.005 for the validation set.

The validity of the machine learning approach was verified by computing an order parameter, O_p, based on the structure factor, $S(k)$. To find the structure factor the following formula was used

$$S(k) = \frac{1}{N}\left[\sum_{i=1}^{N}\sin(k\cdot R_i)\right]^2 + \frac{1}{N}\left[\sum_{i=1}^{N}\cos(k\cdot R_i)\right]^2, \tag{5.12}$$

where $\cdot$ is the scalar product, and to describe the shear-induced lamellar ordering the wave vector, k, was set to point in the direction perpendicular to the lamellar slabs. Usually, the order parameter is the normalized height of the highest $S(k)$ peak that appears for $k > 2\pi/L$, where L is the size of the simulation box. Here, in order to fit O_p to the scale of λ, we transform $S(k_{max})$ linearly in the following way: $O_p = \frac{1}{2}(1 - S(k_{max})) + \lambda_{min}(R)$, where $\lambda_{min}(R)$ is the minimal value of $\lambda(R)$ at given temperature.

The snapshots were visualized using The Open Visualization Tool (OVITO) [Stu10].

5.3.2 Results

The Equilibrium Case

SALR particles form ordered microphases only for a specific range of model parameters which provides a proper balance between the attractive and repulsive interactions [ZC16b].

In the case of very narrow attraction wells, aggregates are formed, but global ordering does not occur [STZ05, TSZ09, RFGZ19]. Here, following [SPP17, Arc08, ZZC16] we choose a SALR potential with a wider attraction well. We have checked that the chosen parameters lead to self-assembly into globally ordered cluster, columnar and lamellar structures. Interestingly, the ordering occurs also at temperatures for which the SALR potential has a positive second virial coefficient [SPP17]. In Fig. 5.8 the lamellar structures obtained from equilibrium MD simulations in the NPT ensemble are presented. The phase transition between the lamellar structures with crystal-like slabs and fluid-like slabs is reflected in a drop of the average density. The transitions occur at different temperatures depending on the simulation protocols used, as described in the caption of Fig. 5.8. A similar but larger hysteresis is observed for the high-temperature melting transition. In both cases the observed density response suggests presence of a first order transition. The melting of the crystal-like ordering preserves lamellar structure, occurs only within slabs, and is a local change that leads to a relatively small hysteresis. On the other hand, in the high-temperature transition the melting turns a lamellar structure into a disordered fluid, and such global order-disorder transition results in a relatively large hysteresis. Transition between lamellar structures with different slab characteristics was not previously reported for 3d systems. However in 2d, similar multi-step melting of stripes was found [ORRB10].

The analogy with the 2d system is not only structural but also dynamical, since the displacements of the particles along different axes behave in a similar manner as in the 2d system (Fig. 5.9 and Fig. 8 in [ORRB10]). In particular, formation of the lamellar structure causes a significant drop in the average particle displacement along the axis perpendicular the slabs, $d_\perp$, and simultaneously a slight increase of $d_\parallel$. The difference between $d_\perp$ and $d_\parallel$ reflects the anisotropy of the system and the suppression of particle exchange between the slabs. Upon further decrease of temperature both the anisotropy and the displacements converge and the crystal lamellar structure is formed.

Nonequilibrium Simulations

In the case of a SALR potential that leads to periodically ordered structures, the effect of shear on the ordered microphases was studied in Ref. [SR18]. Thus, here we focus on the question of whether steady shear can induce ordering of disordered SALR fluids. We apply shear to the equilibrium structures at $P = 0.5\varepsilon/\sigma^3$ and for $k_BT/\varepsilon > 0.4$, that is, above the melting temperature of the lamellar phase. We find that the obtained stationary states show different kinds of ordering depending on the applied shear-rate and the temperature. Importantly, close to the equilibrium order-disorder phase transition, the steady state structures (Fig. 5.10) are highly similar to the equilibrium lamella with fluid-like slabs (Fig. 5.8). However, since many other states were found that were not as highly ordered, in what follows we describe and analyze structural transitions between different shear-induced morphologies by applying different machine learning methods to 2d binary maps of the obtained 3d structures, as described in Section 5.3.1.

Artificial Neural Networks

We applied two common supervised methods for pattern recognition: DNN and CNN. The aim of the training was to learn to classify the structures occuring at $k_BT/\varepsilon = 0.42$, where Principal Component Analysis (PCA), another machine learning method, suggests that in between critical shear rates required for ordering, a gradual structural transition from ordered to torn lamella is present. The results are shown in Fig. 5.12. In both cases

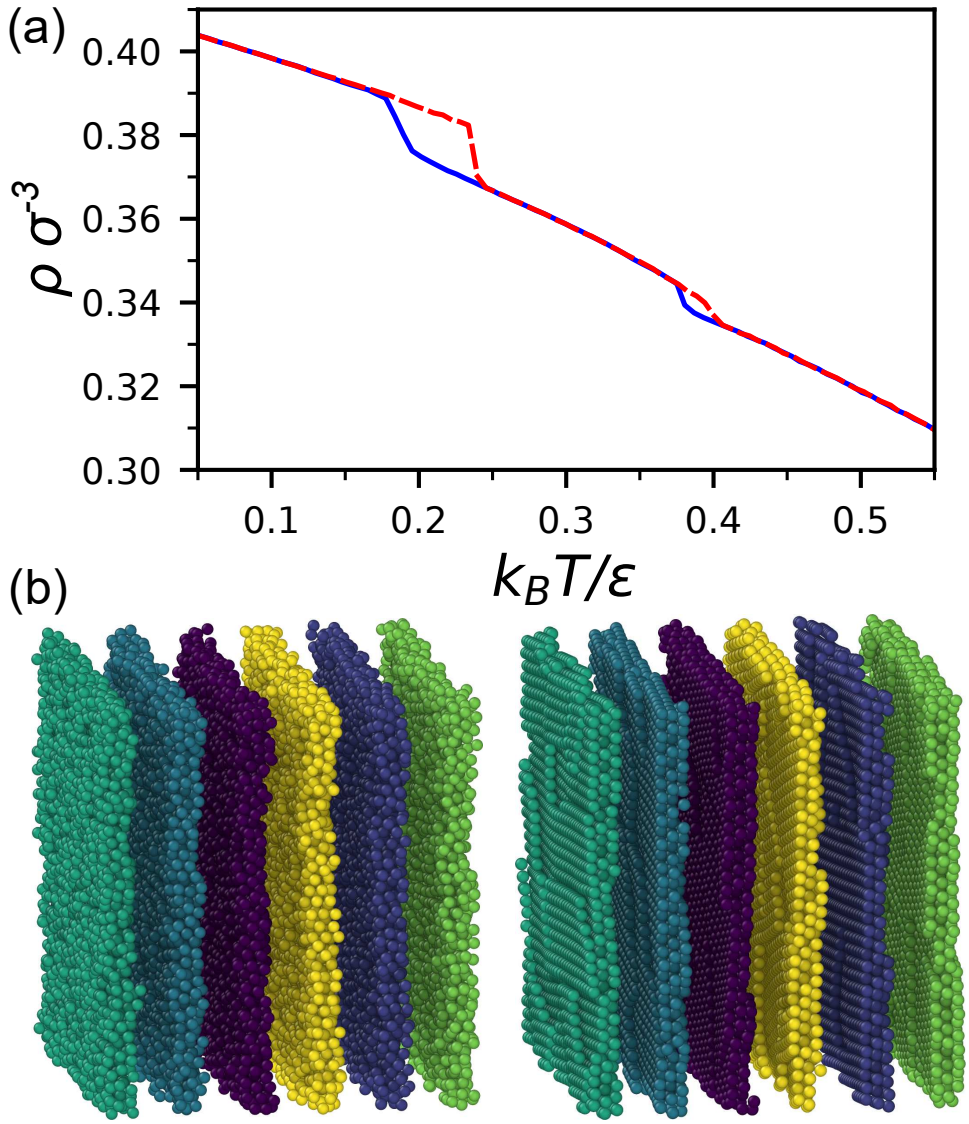


Figure 5.8: (a) Number density as a function of temperature at constant pressure $P = 0.5\varepsilon/\sigma^3$ for a sequence of simulations with linearly increasing (dashed red lined) or decreasing (solid blue line) temperature. The density drops appear at $k_BT/\varepsilon = 0.185$ and 0.377 for decreasing k_BT/ε, while at $k_BT/\varepsilon = 0.237$ and 0.40 for increasing. (b) Representative snapshots of lamellar structures with fluid-like (left) and crystal-like (right) ordering. Reproduced from Ref. [PRP20], with the permission of AIP Publishing.

the steady states for $R\tau \leq 0.5$ and $R\tau \geq 28$ are labeled by the networks as disordered, which agrees with the PCA predictions. In the intermediate range of the shear rate, both networks are able to distinguish the lamellar and the torn lamellar structures and predict that the transition occurs between $11 < R\tau < 20$. This suggests that, contrary to the order-disorder transitions, these steady states change gradually. Therefore, both neural networks' predictions seem to agree with the the PCA analysis. Interestingly, this is, however, not the case for $k_BT/\varepsilon = 0.55$ (not shown in this Thesis), where the thermal fluctuations are more prominent. The structures are far less ordered, and only CNN predictions to some extent agree with PCA whereas the DNN does not even capture the location of the transitions between disordered and the torn-lamellar structures. However, one should note that both networks were trained with data from $k_BT/\varepsilon = 0.42$.

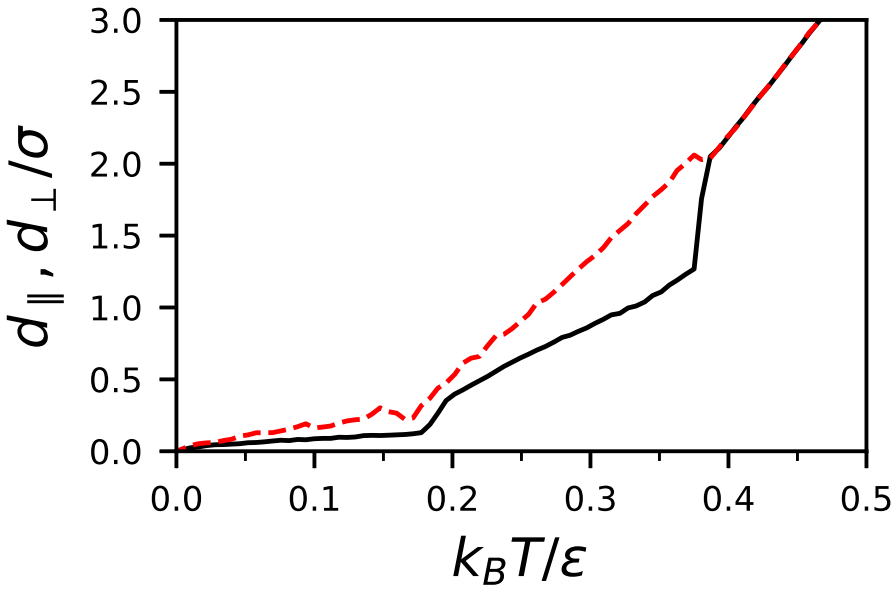


Figure 5.9: Average displacements versus temperature in directions parallel (dashed red) and perpendicular (solid black) to the lamellar slabs. Reproduced from Ref. [PRP20], with the permission of AIP publishing.

It is also interesting to learn how the neural network that we used makes the final decision on labelling a given structure. Some insight into that process can be found in the output, i.e. the last layer of the CNN. The number of neurons in the output layer is equal to to the number of structures that the network was trained to distinguish. The output values p_1, p_2 and p_3 correspond to the lamellar, torn lamellar and disordered structures, respectively. They assume values in the range $[0, 1]$. The final classification decision is made by choosing the structure corresponding to the largest-output softmax neuron of the output layer. In Fig. 5.11, the output value p_1 is shown for the shear rates at which the PCA predicts gradual transition between the lamellar and the torn lamellar structures. In this range of shear rates $p_2 = 1 - p_1$ and $p_3 = 0$. The ongoing structural transition is reflected in the high noisiness of the output. In particular, at the shear rate of $R\tau = 13$ where more then 90% of configurations were labelled as lamellar, the output plot shows that actually in many cases the $p_1 \approx 0.5$, meaning the output values of the lamellar and torn structures are close. This suggests that the network, although with high accuracy, made the decision with less certainty. An increase in the shear rate makes the predictions more ambiguous, so that for $R\tau = 15$, where about half of the configurations are labelled as lamellar and half as torn lamellar, the output resembles random noise. Increasing the shear rate to $R\tau = 17$ changes the situation dramatically. For most of the configurations, the classification decisions are made with high certainty. However, for some range of configurations, a kind of transition between the steady states can be observed. Further shear rate increases result in monostability, and the outputs become constants of value 0 or 1 (not shown).

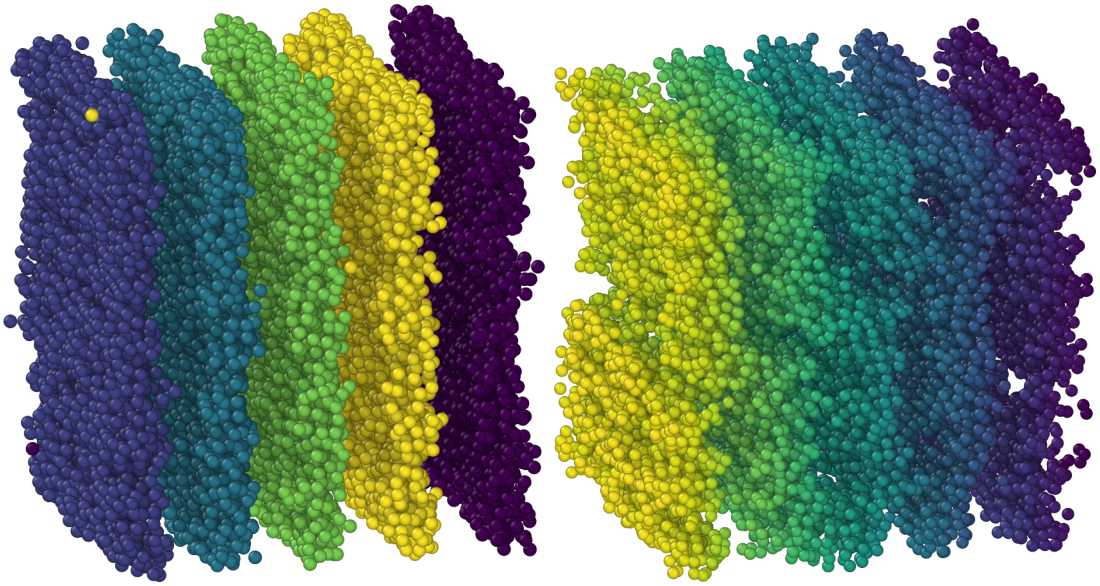


Figure 5.10: Representative steady states obtained after exposing disordered fluid to shear. Left panel: $(k_BT/\varepsilon, R\tau) = (0.42, 10)$, right panel: $(k_BT/\varepsilon, R\tau) = (0.65, 10)$. Reproduced from Ref. [PRP20], with the permission of AIP Publishing.

5.4 Chapter conclusions and outlook

In this Chapter, our work demonstrated that ML approaches are able to learn the order parameter driving a phase transition in layered models, also when this parameter is not immediately apparent from the snapshots without preprocessing. This is directly possible due to the convolutional filters which are, without any a priori knowledge, capable of learning even involved algebraic operations that uncover the order parameters from the data. This paves the way to the use of ML approaches to investigate the properties of systems of increasing complexity and to characterizing phases of matter described by multiple, possibly non-local order parameters. The universal approximation theorem [Hor91] ensures us that a neural network can, at least in principle, learn to recognize arbitrarily complex order parameters.

In particular, an interesting extension would be to study the multilayer Ising model with a number of layers increasing, the three-dimensional Ashkin-Teller and the trilayer Ahkin-Teller in two dimensions, which can be studied with the techniques introduced in the Chapter. Non-local couplings among the layers could be added, which would lead to non-local, more composite, operators. These results should be compared with the identification of hidden order done using non-ML techniques [MCL19].

Our approach may be used for other cases in which the identification of the order parameters is not straightforward [RSM$^+$15, LCH$^+$17, VZ06]. Even if our approach has been devised to deal with coupled spin models and can applied to different geometrical configurations, it is not clear *a priori* that it would succeed in other more complicated cases of coupled interacting systems, such as multilayer configurations of interacting bosons and fermions or bilayer quantum Hall systems. Of course, in order to study generically coupled models one needs to have an efficient algorithm to simulate the uncoupled systems. Nevertheless, we think the present work provides a methodological basis, highlighting the effect of interlayer coupling on the macroscopic properties and phases of coupled systems.

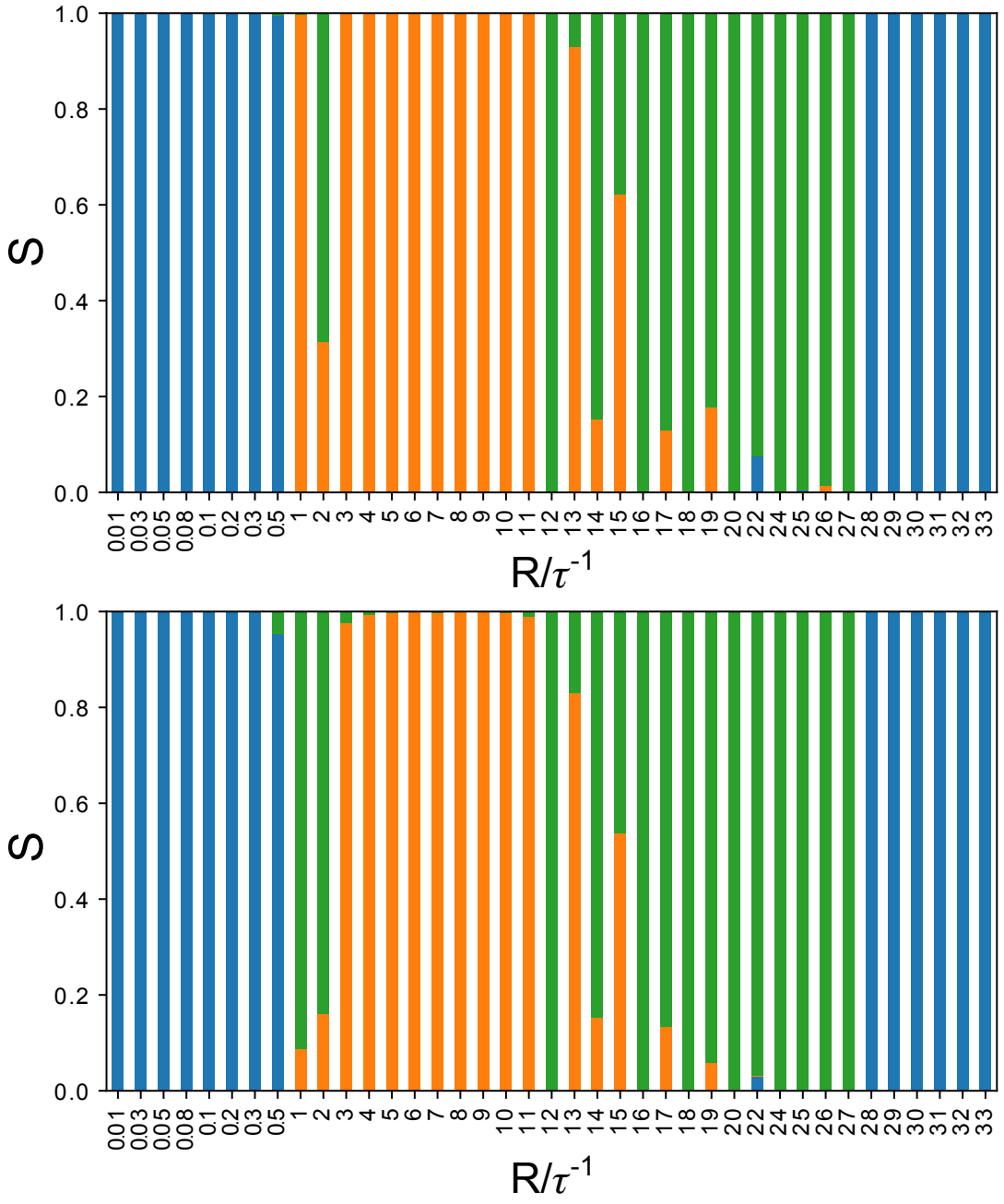


Figure 5.11: Fraction of configurations assigned to different structures by DNN (upper panel) and CNN (lower panel) as a function of the shear rate at $k_BT/\varepsilon = 0.42$. The neural networks were trained to distinguish between the disordered structures (blue), lamellar structure (orange) and torn lamellar structure (green). Reproduced from Ref. [PRP20], with the permission of AIP Publishing.

Naturally, the approach we introduced could also be extended in the future to characterize quantum models, or classical spin models with competition between short- and long-range interactions, or more involved spin models such as the XY model, discretising the continuous degrees of freedom[LRT17]. We expect that by an appropriate choice of the sizes and strides of the filter in the convolutional layer one could characterize antiferromagnetic order parameters, non-local order parameters and exotic order parameters, such as nematic and smectic phases. In this context, current experiments on

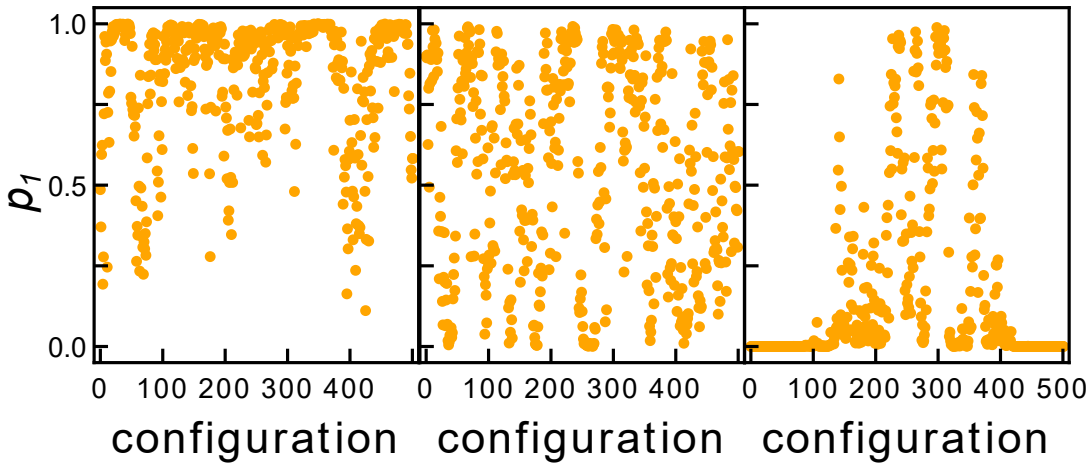


Figure 5.12: Output, p_1, of neuron corresponding to the lamellar structure in the last, softmax-activated layer of the convolutional neural network upon the gradual transition from the lamellar to torn lamellar structures at $k_BT/\varepsilon = 0.42$. From left to right: $R\tau = 13, 15, 17$. Reproduced from Ref. [PRP20], with the permission of AIP Publishing.

fermionic dipolar atoms [LBL12, PWZ15] promise to open a new window in the physics of competing long-range and short-range interactions [BPB+18], clearing the path for the comprehension of modulated phases in strongly interacting quantum systems.

The presence of spatially ordered structures is a *leitmotiv* for long-range and layerd systems such as ultra-thin magnetic films [AB92, KP93a, KP93b], iron-based superconductors and cuprates [PAdSN+10, TBT+16]. The pattern structure normally depends on several experimental conditions and it produces a particularly rich phase diagram. Most of the common features occurring in stripe forming systems and modulated phases remain obscure due to the challenges posed by the complicated order parameters, which occur in these cases [MCBS17, MCS12, BS09, BS11]. The ML technique introduced in the present Chapter may serve as an essential prove to finally uncover the complexity of such phases.

Our results pave the way for fully automated study of phase diagrams of more general and complicated spin systems. An exciting open problem lying in the realm of so-called explainable artificial intelligence (XAI) [DBH18] is whether machine learning techniques could not only learn to separate phases differing by a 'hidden' order parameter, but also identify that parameter. Another natural development of the present work is to use our fully-unsupervised technique to learn directly from experimental data [RKT+19, BCJ+19, ZMF+19]. Finally, it would be interesting to extend the results presented in this Chapter according to the variational procedure discussed in Ref. [KJR18].

In this Chapter we have also analyzed properties of lamellar structures formed by a SALR fluid in equilibrium and nonequilibrium conditions. We have found that the equilibrium lamellar phase can be formed with slabs that exhibit either fluid- or crystal-like ordering. The transition between the two is manifested by a jump in the density and by the average displacement becoming significantly anisotropic upon heating. Although the presence of crystal-like lamellar phase with hexagonally packed particles has been reported before [DCDGF+06], the phase transition between the two ordered lamellar phases has not been shown in previous attempts to describe phase behavior of the 3d SALR fluid [ZZC16, IR04, Arc08, DCDGF+06]. Interestingly, two distinct lamellar phases have been found in 2d lattice [APC14] and off-lattice [ORRB10] models, as well as in

the model for a 1d cross-section of the stripe phase [PCA13]. Importantly, the nature of the transition found here and in Ref. [ORRB10] is the same and shows a liquid-crystal transition within the slabs. On the other hand, the molten lamellar phase described in [APC14], consists of liquid-like stripes with randomly distributed topological defects. In the 3d equilibrium, no such defects were observed, but this can be model-dependent and requires further studies.

Out of equilibrium, the disordered SALR fluid was shown to form a number of anisotropic morphologies when exposed to steady shear. In particular, depending on temperature and applied shear rate, we have found that lamellar, torn lamellar, lamellar mixed with columnar, and hexagonally-ordered columnar structures can occur. The anisotropic structures with weak, global ordering were observed up to $k_BT/\varepsilon = 0.75$, with $k_BT/\varepsilon = 0.4$ being the temperature of the equilibrium order-disorder phase transition.

At equilibrium, one of the main obstacles that a system with competing interactions has to overcome to form an ordered state is finding the global minimum in a highly complex energy landscape. The barriers between the local minima that potentially can trap the system are steeper at lower temperatures. Thus, if one could make the system ordered above its melting temperature, it would place the system close to what becomes the global minimum upon cooling. Here, we have shown that a SALR system exposed to steady shear can order above its melting temperature, and thus shearing can be used to facilitate the formation of the lamellar structure.

The analysis of the shear-induced structures we observed was performed after reduction of data dimensionality. We have shown that in the case of such highly anisotropic structures as lamellae, the structural transitions that take place in 3d space can be successfully quantified using PCA applied to 2d binary maps of the structure.

CHAPTER 6

Conclusions

In this Thesis, we have studied several types of systems. Their nature and physical properties are different, however, all of them belong to the class of many-body systems. This was the fundamental source of difficulty in their investigations performed in this Thesis. Therefore, we resorted to approximate and numerical techniques to overcome these difficulties. Although we have used a wide spectrum of methods, all of them rely on minimization of a certain functional dependent on parameters introduced by a choice of a model. In the variational approach, we minimized the energy of the system with respect to the variational parameters of the ansatz, either analytically or numerically. In phase discrimination, we numerically minimized cost functions of the phase discrimination problem with respect to the weights and biases of the artificial neural networks.

The original results of the Thesis were presented in Chapter 3, Chapter 4, and Chapter 5. Each of these chapters was based on already published papers or a preprint submitted to the arXiv repository. In Chapter 3, based on Ref. [RL18], we derived a Hamiltonian for a spinful angulon in the presence of an external magnetic field. We applied variational technique to retrieve a physically intuitive picture of relevant properties of the system. In Chapter 4, based on Ref [RLM21], we developed a neural-network quantum state ansatz for non-additive systems. We exemplified this approach on the polaron Hamiltonian, a non-additive system with significant physical meaning. In Chapter 5, based on Refs. [RDC$^+$20, PRP20], we used artificial neural networks, in particular convolutional neural networks, to study classical spin lattice models as well as colloidal systems.

At the end of each Chapter, we presented conclusions and outlook resulting from the work described in each of them. However, these conclusions intersect each other and converge towards a bigger picture of the potential outlook, both for the systems studied and methods used in this Thesis. Let us present these broader conclusions in the rest of this Chapter.

The angulon theory, now 6 years from its theorization, has already a proven track record of explaining important phenomena in quantum physics. These developments were described already in Chapter 1 and continued in Chapter 3. However, there is still large room for further developments. To compare, the translational counterpart of angulon, the polaron, has been a subject of active research for almost ninety years (Ref. [Lan33] was published in 1933). The angulon, with its treatment of interactions between angular momenta that provides physical insight as opposed to black-box numerical approach,

provides hope for further studies explaining involved phenomena in atomic, molecular, and solid-state physics. Let us now provide a more detailed picture of potential outlook.

The angulon theory could be developed further to treat even more involved systems. A particular example here would be the hydroxyl radical (OH). This is a very natural extension – the original angulon theory covered only rotational angular momentum, while in Chapter 3 we extended the theory to cover also the spin angular momentum; the hydroxyl radical, in addition, has a non-zero orbital angular momentum. Such an extension is not only of interest purely from the theoretical perspetive – there exists already exciting experimental data, lacking detailed microscopic explanation, both on OH Λ-doubling [RLD13] and Zeeman effect [Dou] in the presence of the bath.

For this potential extension, one could take the following path. First, one would develop a theoretical model of Λ-doubling in the presence of the bath. Then, using experimental data, one could fix the OH-helium interaction interaction parameters (magnitudes and ranges) from this Λ-doubling model. In the second step, one would develop a model of Zeeman splitting in the presence of the bath. Then, finally, one could evaluate the consistence of both models by checking how well experimental Zeeman splitting is reproduced when using the parameters established from fitting the Λ-doubling.

Other possible extensions of the angular quasiparticle approach involves solid-state magnetism. In spite of intensive theoretical and experimental efforts, exhaustive microscopic description of phenomena related to solid-state magnetism remains to pose a challenge. One of the most important ingredients yet to be done for such a complete theoretical picture is the description of the microscopic mechanisms of angular momentum transfer between the electron spin, orbital motion and the lattice degrees of freedom [GC15b].

This is important not only from the fundamental point of view. These insights are crucial, for example, to reduce decoherence in solid-state qubits [DRS$^+$16]. This is one of the most promising setups for future large-scale quantum computing [Pre18] with potentially a wide range of real-life applications. However, directly studying angular momentum exchange in solid-state systems is usually extremely challenging. Therefore, one resorts to model, controllable systems. An example of such model system is the oxygen molecule, O_2, immersed in helium nanodroplets. The ground state of the oxygen molecule is $^3\Sigma_g^-$, – this denotes that the orbital angular momentum is $L = 0$, while spin is $S = 1$. Studying this configuration naturally complements the finished work of Chapter 3, where the case $S = 1/2$, $L = 0$ was studied and the already discussed potential extension to the hydroxyl radical, which tackles the case $S = 1/2$, $L = 1$.

Therefore, another potential extension is to explore the use of the angulon theory to explain the observed change in Stern-Gerlach effect when the oxygen molecule O_2 is immersed in helium nanodroplets, as compared to the same experiment in the gas phase. One expects that in Stern-Gerlach experiment, the three-fold multiplicity of the total angular momentum projection results in a three-fold splitting of the molecular beam when exposed to a magnetic field. However, experimentalists have already observed that only one of the three sub-beams is present. This leads us to formulate the following hypothesis: it is the interaction with the bath that splits the energy levels corresponding to different angular momentum projections. This would subsequently induce a decay to the lowest-lying state. This would be then a direct experimental evidence that the transfer of angular momentum – which one would describe using techniques derived from the theoretical description of the angulon quasiparticle – is indeed happening in the system.

As already said, the extension to oxygen molecule and hydroxyl radical will probably be started with analytic approach similar to the techniques presented in Chapter 3. This would give analytic insight into these problems. However, it is also possible that, after refinement and adaption, the artificial neural network quantum state approach, similar to the one described in Chapter 4, will also work for the hydroxyl radical and oxygen molecule. These approaches, when refined to cover the more involved angular momentum algebra, are expected to provide more accurate numerical results that might be directly comparable with experimental data.

Regarding machine learning, this Thesis has covered two different paradigms. In Chapter 4, we exploited neural-network quantum states, while in Chapter 5 we harnessed the power of convolutional neural networks to uncover the phase diagrams of many-body systems. However, the machine learning approaches in physics and, hence, the outlook of these Chapters, stretch way beyond these two paradigms. In what follows in this Chapter, we will cover particularly exciting subfields of machine learning that might shed new light on physics in the coming years.

A particular machine learning method that could also rise to prominence in (quantum) physics is reinforcement learning [SB18]. In reinforcement learning, one changes the paradigm – now, there is an *agent* performing *actions* in an *environment*. The actions to be taken in a given state of the environment are governed by the policy function, which assigns numerical scores to state-action pairs. This function is usually approximated by an artificial neural network and trained in a process which combines random exploration of the environment with the exploitation of the already partially trained policy. Compared to supervised and unsupervised learning, this method is known to be very challenging [DAMH19] – the challenges include extremely high dimensionality of the state-action space, only partial observation of the environment, or very limited training data. Thus, the reinforcement learning boom might came with a delay compared to developments in supervised and unsupervised learning, nevertheless, the perspectives are extremely exciting.

Reinforcement learning has already a track record of successful applications in physics. One particular field in which reinforcement learning has been successful, is quantum control [NBSN19, BDS$^+$18, ZWA$^+$19]. These advancements in quantum control might be useful for future quantum computing or for real-life experiments in other subfields of quantum physics. However, the already achieved progress is not limited to quantum computing – there have already been successful applications for improving the dynamics of sensors [SFB20], optimizing quantum error correction codes [NDD$^+$19], or optimizing quantum circuits [FNML21].

The so-called meta learning, or "learning to learn" is also an expanding field of machine learning. In meta learning, the machine learning techniques are not applied to the problem itself, but to metadata concerning the machine learning approach and experiments. The goal is to use machine learning itself to choose the most feasible machine learning technique to solve a given problem. Such approaches in computer science date back already to years 1980s and 1990s [Sch87, BBC90]. Their importance grew then with the deep learning revolution, which we mentioned already in Chapter 2. However, the meta learning methods are not yet widespread in physics applications of machine learning. One could expect the meta learning techniques to gradually start automating more and more tasks in machine learning approaches, including in physics, regardless if they belong to unsupervised, supervised, or reinforcement learning. This would fit into the long-standing aim [WB09]

to automate (at least part of) scientific discovery.

Obviously, conceptual advancements are the most important trigger of wider applications of machine learning approaches in physics. However, it is worth noting that with time, it also becomes increasingly easier to perform numerical experiments to test these conceptual advancements. The developments include both general-use software packages, such as already mentioned jax package [BFH+18, FJL18], but also specialized software packages for use in physics, such as NetKet (based on jax), whose third version has been recently announced [VHS+21]. Compared to earlier software, such as TensorFlow [ABC+16] or Pytorch [Ket17], they provide a more flexible way of implementing machine learning concepts in physics, especially novel architectures of neural-network quantum states. Yet, they keep high performance, enabling scaling to larger systems – for instance, more particles in the system or larger cap on bosonic excitations allowed per particle.

These modern software packages combine several advantages advantages for the researcher in a single, ready-to-use package. First, they enable just-in-time compilation [Ayc03] of Python code, thus greatly speeding up execution of programs, albeit at a cost of introducing small constraints to the programmer. Secondly, they involve automatic differentiation, a technique enabling efficient computation of derivatives of functions specified by a computer program. This is crucial for machine learning applications as they almost always incorporate calculation of optimization target gradients. Thirdly, the same programs can be executed without change on GPUs and even on more powerful TPUs (Tensor Processing Units) [JYPP18]. Compared to usual execution on a CPU, even a multi-core one, GPU and TPU enable much more massive parallelization, enabling scaling way beyond the CPU capabilities. Historically, code adaptation to these software accelerators used to be a considerable challenge requiring significant field knowledge. One can expect these developments in available software to continue, thus providing a further boost to the already rapidly expanding field of machine learning in physics.

Bibliography

[AB92] R Allenspach and A Bischof. Magnetization direction switching in Fe/Cu(100) epitaxial films: Temperature and thickness dependence. *Physical Review Letters*, 69(23):3385–3388, December 1992.

[ABC+16] Martin Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. Tensorflow: A system for large-scale machine learning. In *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pages 265–283, 2016.

[ABC+17] Francesco Ancilotto, Manuel Barranco, François Coppens, Jussi Eloranta, Nadine Halberstadt, Alberto Hernando, David Mateo, and Marti Pi. Density functional theory of doped superfluid liquid helium and nanodroplets. *Int. Rev. in Phys. Chem.*, 36(4):621, 2017.

[AFM+12] K. Aikawa, A. Frisch, M. Mark, S. Baier, A. Rietzler, R. Grimm, and F. Ferlaino. Bose-einstein condensation of erbium. *Physical Review Letters*, 108:210401, 2012.

[ALBvLM14] Louis-François Arsenault, Alejandro Lopez-Bezanilla, O. Anatole von Lilienfeld, and Andrew J. Millis. Machine learning for many-body physics: The case of the Anderson impurity model. *Physical Review B*, 90:155136, Oct 2014.

[ALT08] Joshua A Anderson, Chris D Lorenz, and Alex Travesset. General purpose molecular dynamics simulations fully implemented on graphics processing units. *Journal of computational physics*, 227(10):5342–5359, 2008.

[And61] P. W. Anderson. Localized magnetic states in metals. *Physical Review*, 124:41–53, Oct 1961.

[AP88] Bruce J Ackerson and Philip N Pusey. Shear-induced order in suspensions of hard spheres. *Physical Review Letters*, 61(8):1033, 1988.

[APC14] Noe G Almarza, Jakub Pękalski, and Alina Ciach. Two-dimensional lattice model for periodic ordering of clusters and stripes. ii. monte carlo simulations. *Journal of Chemical Physics*, 140:164708, 2014.

[APC16] Noe G Almarza, Jakub Pękalski, and Alina Ciach. Effects of confinement on pattern formation in two dimensional systems with competing interactions. *Soft Matter*, 12:7551–7563, 2016.

[Arc08] Andrew J Archer. Two-dimensional fluid with competing interactions exhibiting microphase separation: Theory for bulk and interfacial properties. *Physical Review E*, 78(3):031402, 2008.

[ART+00] RM Amos, JG Rarity, PR Tapster, TJ Shepherd, and SC Kitson. Fabrication of large-area face-centered-cubic hard-sphere colloidal crystals by shear alignment. *Physical Review E*, 61(3):2929, 2000.

[ASBn+18] Yuto Ashida, Tao Shi, Mari Carmen Bañuls, J. Ignacio Cirac, and Eugene Demler. Solving quantum impurity problems in and out of equilibrium with the variational approach. *Physical Review Letters*, 121:026805, Jul 2018.

[Ayc03] John Aycock. A brief history of just-in-time. *ACM Computing Surveys (CSUR)*, 35(2):97–113, 2003.

[BAFL04] Ronald Blaak, Stefan Auer, Daan Frenkel, and Hartmut Löwen. Crystal nucleation of colloidal suspensions under shear. *Physical Review Letters*, 93(6):068303, 2004.

[BAT17] Peter Broecker, Fakher F Assaad, and Simon Trebst. Quantum phase recognition via unsupervised machine learning. *arXiv preprint arXiv:1707.00663*, July 2017.

[Bax07] Rodney J Baxter. *Exactly Solved Models in Statistical Mechanics*. Dover books on physics. Dover Publications, 2007.

[BBC90] Yoshua Bengio, Samy Bengio, and Jocelyn Cloutier. *Learning a synaptic learning rule*. Citeseer, 1990.

[BC21] Lorenzo Buffoni and Filippo Caruso. New trends in quantum machine learning (a). *EPL (Europhysics Letters)*, 132(6):60004, 2021.

[BCJ+19] Annabelle Bohrdt, Christie S Chiu, Geoffrey Ji, Muqing Xu, Daniel Greif, Markus Greiner, Eugene Demler, Fabian Grusdt, and Michael Knap. Classifying snapshots of the doped Hubbard model with machine learning. *Nature Physics*, 15(9):921–924, July 2019.

[BCMT17] Peter Broecker, Juan Carrasquilla, Roger G Melko, and Simon Trebst. Machine learning quantum phases of matter beyond the fermion sign problem. *Scientific reports*, 7(1):1–10, 2017.

[BDN+19] Giacomo Bighin, Nicolò Defenu, István Nándori, Luca Salasnich, and Andrea Trombettoni. Berezinskii-Kosterlitz-Thouless Paired Phase in Coupled XY Models. *Physical Review Letters*, 123(10):164–6, September 2019.

[BDPZ12] Mikhail A Baranov, Marcello Dalmonte, Guido Pupillo, and Peter Zoller. Condensed matter theory of dipolar quantum gases. *Chemical Reviews*, 112(9):5012–5061, 2012.

[BDS+18] Marin Bukov, Alexandre GR Day, Dries Sels, Phillip Weinberg, Anatoli Polkovnikov, and Pankaj Mehta. Reinforcement learning in different phases of quantum control. *Physical Review X*, 8(3):031086, 2018.

[Ber05] Peter F Bernath. *Spectra of Atoms and Molecules*. Oxford University Press, UK, 2 edition, 2005.

[BFH+18] James Bradbury, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. JAX: composable transformations of Python+NumPy programs, 2018.

[BH03] Jean-Louis Barrat and Jean-Pierre Hansen. *Basic concepts for simple and complex liquids*. Cambridge University Press, 2003.

[BKG+13] Jonathan B Balewski, Alexander T Krupp, Anita Gaj, David Peter, Hans Peter Büchler, Robert Löw, Sebastian Hofferberth, and Tilman Pfau. Coupling a single electron to a Bose–Einstein condensate. *Nature*, 502(7473):664–667, October 2013.

[BL17] Giacomo Bighin and Mikhail Lemeshko. Diagrammatic approach to orbital quantum impurities interacting with a many-particle environment. *Physical Review B*, 96:085410, Aug 2017.

[BLLY21a] M. Brooks, M. Lemeshko, D. Lundholm, and E. Yakaboylu. Molecular impurities as a realization of anyons on the two-sphere. *Physical Review Letters*, 126:015301, Jan 2021.

[BLLY21b] Morris Brooks, Mikhail Lemeshko, Douglas Lundholm, and Enderalp Yakaboylu. Emergence of anyons on the two-sphere in molecular impurities. *arXiv preprint arXiv:2108.06966*, 2021.

[BOST95] Richard Brower, Kostas Orginos, Yue Shen, and Chung-I Tan. Monte Carlo study of the Yukawa coupled two-spin Ising model. *Physica A: Statistical Mechanics and its Applications*, 221(4):554 – 564, 1995.

[BP02] H.-P. Breuer and F. Petruccione. *The Theory of Open Quantum Systems*. Oxford University Press, Oxford, 2002.

[BPB+18] S Baier, D Petter, J H Becher, A Patscheider, G Natale, L Chomaz, M J Mark, and F Ferlaino. Realization of a Strongly Interacting Fermi Gas of Dipolar Atoms. *Physical Review Letters*, 121(9):093602, August 2018.

[BS09] Daniel G Barci and Daniel A Stariolo. Orientational order in two dimensions from competing interactions at different scales. *Physical Review B*, 79(7):85, February 2009.

[BS11] Daniel G Barci and Daniel A Stariolo. Microscopic approach to orientational order of domain walls. *Physical Review B*, 84(9):175, September 2011.

[BSA04] Egor Babaev, Asle Sudbø, and N W Ashcroft. A superconductor to superfluid phase transition in liquid metallic hydrogen. *Nature*, 431(7009):666–668, October 2004.

[Bur69] Donald Bures. An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite w*-algebras. *Transactions of the American Mathematical Society*, 135:199–212, 1969.

[BWP+17] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

[Car96] J. Cardy. *Scaling and renormalization in statistical physics*. 1996.

[Car20] Juan Carrasquilla. Machine learning for quantum matter. *Advances in Physics: X*, 5(1):1797528, 2020.

[CBC+19] Igor N Cherepanov, Giacomo Bighin, Lars Christiansen, Anders Vestergaard Jørgensen, Richard Schmidt, Henrik Stapelfeldt, and Mikhail Lemeshko. Far-from-equilibrium dynamics of angular momentum in a quantum many-particle system. *arXiv preprint arXiv:1906.12238*, 2019.

[CBS+21] Igor N Cherepanov, Giacomo Bighin, Constant A Schouder, Adam S Chatterley, Simon H Albrechtsen, Alberto Viñas Muñoz, Lars Christiansen, Henrik Stapelfeldt, and Mikhail Lemeshko. Excited rotational states of molecules in a superfluid. *arXiv preprint arXiv:2107.00468*, 2021.

[CCC+19] Giuseppe Carleo, Ignacio Cirac, Kyle Cranmer, Laurent Daudet, Maria Schuld, Naftali Tishby, Leslie Vogt-Maranto, and Lenka Zdeborová. Machine learning and the physical sciences. *Reviews of Modern Physics*, 91(4):045002, 2019.

[CCMK17] Kelvin Ch'ng, Juan Carrasquilla, Roger G. Melko, and Ehsan Khatami. Machine learning phases of strongly correlated fermions. *Physical Review X*, 7:031038, Aug 2017.

[CCS+20] Adam S Chatterley, Lars Christiansen, Constant A Schouder, Anders V Jørgensen, Benjamin Shepperson, Igor N Cherepanov, Giacomo Bighin, Robert E Zillich, Mikhail Lemeshko, and Henrik Stapelfeldt. Rotational coherence spectroscopy of molecules in helium nanodroplets: Reconciling the time and the frequency domains. *Physical Review Letters*, 125(1):013001, 2020.

[CG10] Alina Ciach and Wojciech T Góźdź. Mesoscopic description of network-forming clusters of weakly charged colloids. *Condensed Matter Physics*, 2010.

[CGG18] Stefanie Czischek, Martin Gärttner, and Thomas Gasenzer. Quenches near ising quantum criticality as a challenge for artificial neural networks. *Physical Review B*, 98:024311, Jul 2018.

[CGS05] E. M. Chudnovsky, D. A. Garanin, and R. Schilling. Universal mechanism of spin relaxation in solids. *Physical Review B*, 72:094426, 2005.

[Cia08] Alina Ciach. Universal sequence of ordered structures obtained from mesoscopic description of self-assembly. *Physical Review E*, 78(6):061505, 2008.

[CIG07] M. A. Cazalilla, A. Iucci, and T. Giamarchi. Competition between vortex unbinding and tunneling in an optical lattice. *Physical Review A*, 75:051603, May 2007.

[CL17] Igor N. Cherepanov and Mikhail Lemeshko. Fingerprints of angulon instabilities in the spectra of matrix-isolated molecules. *Physical Review Materials*, 1:035602, Aug 2017.

[CL18] Zi Cai and Jinguo Liu. Approximating quantum many-body wave functions using artificial neural networks. *Physical Review B*, 97(3):035116, 2018.

[CM17] Juan Carrasquilla and Roger G Melko. Machine learning phases of matter. *Nature Physics*, 13(5):431–434, 2017.

[CMB00] Alan A Catherall, John R Melrose, and Robin C Ball. Shear thickening and order–disorder effects in concentrated colloids at high shear rates. *Journal of Rheology*, 44(1):1–25, 2000.

[CMC20] Kenny Choo, Antonio Mezzacapo, and Giuseppe Carleo. Fermionic neural-network states for ab-initio electronic structure. *Nature communications*, 11(1):1–7, 2020.

[CPG13] Alina Ciach, Jakub Pękalski, and Wojciech T Góźdź. Origin of similarity of phase diagrams in amphiphilic and colloidal systems with competing interactions. *Soft Matter*, 9:6301, 2013.

[CT17] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017.

[DAMH19] Gabriel Dulac-Arnold, Daniel Mankowitz, and Todd Hester. Challenges of real-world reinforcement learning. *arXiv preprint arXiv:1904.12901*, 2019.

[Dan13] Graupe Daniel. *Principles of artificial neural networks*, volume 7. World Scientific, 2013.

[DBC+19] Emmanuel Gbenga Dada, Joseph Stephen Bassi, Haruna Chiroma, Shafi'i Muhammad Abdulhamid, Adebayo Olusola Adetunmbi, and Opeyemi Emmanuel Ajibuwa. Machine learning for email spam filtering: review, approaches and open research problems. *Heliyon*, 5(6):e01802, 2019.

[DBH18] Filip Karlo Došilović, Mario Brčić, and Nikica Hlupić. Explainable artificial intelligence: A survey. In *2018 41st International convention on information and communication technology, electronics and microelectronics (MIPRO)*, pages 0210–0215. IEEE, 2018.

[DCDGF+06] Antonio De Candia, Emanuela Del Gado, Annalisa Fierro, Nicolas Sator, Marco Tarzia, and Antonio Coniglio. Columnar and lamellar phases in attractive colloidal systems. *Physical Review E*, 74(1):010403, 2006.

[Dev16] JT Devreese. Fröhlich Polarons. Lecture course including detailed theoretical derivations. *arXiv preprint arXiv:1611.06122*, 2016.

[DG04] Gesualdo Delfino and Paolo Grinza. Universal ratios along a line of critical points. the Ashkin-Teller model. *Nuclear Physics B*, 682(3):521 – 550, 2004.

[Dic54] Robert H. Dicke. Coherence in spontaneous radiation processes. *Physical Review*, 93:99–110, Jan 1954.

[DLDS17] Dong-Ling Deng, Xiaopeng Li, and S. Das Sarma. Quantum entanglement in neural network states. *Physical Review X*, 7:021021, May 2017.

[DM98] Gesualdo Delfino and Giuseppe Mussardo. Non-integrable aspects of the multi-frequency sine-Gordon model. *Nuclear Physics B*, 516(3):675 – 703, 1998.

[Dou] Gary E. Douberly. Unpublished.

[DRS+16] F. Donati, S. Rusponi, S. Stepanow, C. Wäckerlin, A. Singha, L. Persichetti, R. Baltic, K. Diller, F. Patthey, E. Fernandes, J. Dreiser, Ž. Šljivančanin, K. Kummer, C. Nistor, P. Gambardella, and H. Brune. Magnetic remanence in single atoms. *Science*, 352(6283):318–321, 2016.

[DSH77] Peter J Domaille, Timothy C Steimle, and David O Harris. The rotational spectrum of the state of the SrF radical using laser microwave optical double resonance. *Journal of Molecular Spectroscopy*, 68(1):146–155, oct 1977.

[DV19] Anamika Dhillon and Gyanendra K. Verma. Convolutional neural network: a review of models, methodologies and applications to object detection. *Progress in artificial intelligence*, Dec 2019.

[Emi13] David Emin. *Polarons*. Cambridge University Press, 2013.

[Eul76] Leonhard Euler. Novi commentarii academiae scientiarum petropolitanae. *Nr*, 20:189–207, 1776.

[Fey55] Richard Phillips Feynman. Slow electrons in a polar crystal. *Physical Review*, 97(3):660, 1955.

[FGN00] M. Fabrizio, A.O. Gogolin, and A.A. Nersesyan. Critical properties of the double-frequency sine-gordon model with applications. *Nuclear Physics B*, 580(3):647 – 687, 2000.

[FH00] Bretislav Friedrich and Dudley Herschbach. Steric proficiency of polar 2Σ molecules in congruent electric and magnetic fields. *Phys. Chem. Chem. Phys.*, 2(4):419–428, 2000.

[FHT75] Robert W Field, David O Harris, and Takehiko Tanaka. Continuous wave dye laser excitation spectroscopy CaF A2Πr-X2Σ+. *Journal of Molecular Spectroscopy*, 57(1):107–117, 1975.

[FJL18] Roy Frostig, Matthew James Johnson, and Chris Leary. Compiling machine learning programs via high-level tracing. *Systems for Machine Learning*, 2018.

[FK21] J. Thorben Frank and Michael J. Kastoryano. Learning neural network quantum states with the linear method, 2021.

[FM19] Giammarco Fabiani and Johan H. Mentink. Investigating ultrafast quantum magnetism with machine learning. *SciPost Phys.*, 7:4, 2019.

[FNML21] Thomas Fösel, Murphy Yuezhen Niu, Florian Marquardt, and Li Li. Quantum circuit optimization with deep reinforcement learning. *arXiv preprint arXiv:2103.07585*, 2021.

[Frö54] Herbert Fröhlich. Electrons in lattice fields. *Advances in Physics*, 3(11):325–361, 1954.

[FVB20] Francesco Ferrari, Roser Valentí, and Federico Becca. Variational wave functions for the spin-peierls transition in the su-schrieffer-heeger model with quantum phonons. *Physical Review B*, 102:125149, Sep 2020.

[GBC16] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[GC15a] D. A. Garanin and E. M. Chudnovsky. Angular momentum in spin-phonon processes. *Physical Review B*, 92:024421, 2015.

[GC15b] D. A. Garanin and E. M. Chudnovsky. Angular momentum in spin-phonon processes. *Physical Review B*, 92:024421, Jul 2015.

[GDKG70] F. R. Gamble, F. J. DiSalvo, R. A. Klemm, and T. H. Geballe. Superconductivity in layered structure organometallic crystals. *Science*, 168(3931):568–570, 1970.

[GE97] F Ghezzi and JC Earnshaw. Formation of meso-structures in colloidal monolayers. *Journal of Physics: Condensed Matter*, 9(37):L517, 1997.

[Gla63] Roy J. Glauber. Coherent and incoherent states of the radiation field. *Physical Review*, 131:2766–2788, Sep 1963.

[GLL09] Alessandro Giuliani, Joel L Lebowitz, and Elliott H Lieb. Pattern formation in systems with competing interactions. In *AIP Conference Proceedings*, volume 1091, pages 44–54. American Institute of Physics, 2009.

[GMF+20] Chaoyu Guo, Xiangzhi Meng, Huixia Fu, Qin Wang, Huimin Wang, Ye Tian, Jinbo Peng, Runze Ma, Yuxiang Weng, Sheng Meng, Enge Wang, and Ying Jiang. Probing nonequilibrium dynamics of photoexcited polarons on a metal-oxide surface with atomic precision. *Physical Review Letters*, 124:206801, May 2020.

[GNA+15] Jens Glaser, Trung Dac Nguyen, Joshua A Anderson, Pak Lui, Filippo Spiga, Jaime A Millan, David C Morse, and Sharon C Glotzer. Strong scaling of general-purpose molecular dynamics simulations on gpus. *Computer Physics Communications*, 192:97–107, 2015.

[GS98] Christian Grosche and Frank Steiner. *Handbook of Feynman path integrals*. 1998.

[GS18] David J Griffiths and Darrell F Schroeter. *Introduction to quantum mechanics*. Cambridge University Press, 2018.

[GWK+18] Jiuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Gang Wang, Jianfei Cai, et al. Recent advances in convolutional neural networks. *Pattern Recognition*, 77:354–377, 2018.

[HAC+00] Christopher Harrison, Douglas H. Adamson, Zhengdong Cheng, John M. Sebastian, Srinivasan Sethuraman, David A. Huse, Richard A. Register, and P. M. Chaikin. Mechanisms of ordering in striped patterns. *Science*, 290(5496):1558–1560, 2000.

[HC19] Michael J. Hartmann and Giuseppe Carleo. Neural-network approach to dissipative quantum many-body dynamics. *Physical Review Letters*, 122:250502, Jun 2019.

[Heg05] D. C. Heggie. The Classical Gravitational N-Body Problem. *arXiv preprint astro-ph/0503600*, 2005.

[HFM+21] Damian Hofmann, Giammarco Fabiani, Johan H Mentink, Giuseppe Carleo, and Michael A Sentef. Role of stochastic noise and generalization error in the time propagation of neural-network quantum states. *arXiv preprint arXiv:2105.01054*, 2021.

[HGM+65] N. B. Hannay, T. H. Geballe, B. T. Matthias, K. Andres, P. Schmidt, and D. MacNair. Superconductivity in graphitic compounds. *Physical Review Letters*, 14:225–226, Feb 1965.

[HH79] K P Huber and G Herzberg. *Molecular spectra and molecular structure. IV. Constants of diatomic molecules*. Litton, 1979.

[HL73] Klaus Hepp and Elliott H Lieb. On the superradiant phase transition for molecules in a quantized radiation field: the dicke maser model. *Annals of Physics*, 76(2):360–404, 1973.

[HLIM93] Per Lyngs Hansen, Jesper Lemmich, John Hjort Ipsen, and Ole G. Mouritsen. Two coupled ising planes: Phase diagram and interplanar force. *Journal of Statistical Physics*, 73(3):723–749, Nov 1993.

[HLO+20] Jonathan Heek, Anselm Levskaya, Avital Oliver, Marvin Ritter, Bertrand Rondepierre, Andreas Steiner, and Marc van Zee. Flax: A neural network library and ecosystem for JAX, 2020.

[Hol59] Th Holstein. Studies of polaron motion: Part i. the molecular-crystal model. *Annals of physics*, 8(3):325–342, 1959.

[Hor91] Kurt Hornik. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4(2):251–257, January 1991.

[HSK+12] Geoffrey E Hinton, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, and Ruslan R Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580*, 2012.

[HSS17] Wenjian Hu, Rajiv R. P. Singh, and Richard T. Scalettar. Discovering phases, phase transitions, and crossovers through unsupervised machine learning: A critical examination. *Physical Review E*, 95:062122, Jun 2017.

[IFT12] Mauro Iazzi, Stefano Fantoni, and Andrea Trombettoni. Anisotropic ginzburg-landau and lawrence-doniach models for layered ultracold fermi gases. *EPL (Europhysics Letters)*, 100(3):36007, 2012.

[IR04] A Imperio and L Reatto. A bidimensional fluid system with competing interactions: spontaneous and induced pattern formation. *Journal of Physics: Condensed Matter*, 16(38):S3769, 2004.

[IRZ08] Alessandra Imperio, Luciano Reatto, and Stefano Zapperi. Rheology of colloidal microphases in a model with competing interactions. *Physical Review E*, 78(2):021402, 2008.

[IS20] Naoki Irikura and Hiroki Saito. Neural-network quantum states at finite temperature. *Physical Review Research*, 2:013284, Mar 2020.

[JMRW64] Joshua Jortner, Lothar Meyer, Stuart A. Rice, and E. G. Wilson. Energy transfer phenomena in liquid helium. *Physical Review Letters*, 12:415–416, Apr 1964.

[JWS+16] N. B. Jørgensen, L. Wacker, K. T. Skalmstang, M. M. Parish, J. Levinsen, R. S. Christensen, G. M. Bruun, and J. J. Arlt. Observation of attractive and repulsive polarons in a Bose-Einstein condensate. *Physical Review Letters*, 117:055302, 2016.

[JY12] Deborah S Jin and Jun Ye. Introduction to Ultracold Molecules: New Frontiers in Quantum and Chemical Physics. *Chemical reviews.*, 112:4801, 2012.

[JYPP18] Norman Jouppi, Cliff Young, Nishant Patil, and David Patterson. Motivation for and evaluation of the first tensor processing unit. *IEEE Micro*, 38(3):10–19, 2018.

[KA09] Robert M. Konik and Yury Adamov. Renormalization group for treating 2D coupled arrays of continuum 1D systems. *Physical Review Letters*, 102:097203, Mar 2009.

[KACE09] Markus Koch, Gerald Auböck, Carlo Callegari, and Wolfgang E. Ernst. Coherent Spin Manipulation and ESR on Superfluid Helium Nanodroplets. *Physical Review Letters*, 103:035302, Jul 2009.

[KB14] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[Ket17] Nikhil Ketkar. Introduction to pytorch. In *Deep learning with python*, pages 195–208. Springer, 2017.

[KJR18] Maciej Koch-Janusz and Zohar Ringel. Mutual information, neural networks and the renormalization group. *Nature Physics*, 14(6):578, 2018.

[KP93a] A Kashuba and V L Pokrovsky. Stripe domain structures in a thin ferromagnetic film. *Physical Review Letters*, 70(20):3155–3158, May 1993.

[KP93b] A B Kashuba and V L Pokrovsky. Stripe domain structures in a thin ferromagnetic film. *Physical Review B*, 48(14):10335–10344, October 1993.

[KRLdJ05] B. Koopmans, J. J. M. Ruigrok, F. Dalla Longa, and W. J. M. de Jonge. Unifying Ultrafast Magnetization Dynamics. *Physical Review Letters*, 95:267207, 2005.

[KSH12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.

[KTSB17] Seher Karakuzu, Luca F. Tocchio, Sandro Sorella, and Federico Becca. Superconductivity, charge-density waves, antiferromagnetism, and phase separation in the hubbard-holstein model. *Physical Review B*, 96:205145, Nov 2017.

[Lan33] L D Landau. Über die Bewegung der Elektronen in Kristalgitter. *Phys. Z. Sowjetunion*, 3:664, 1933.

[LAS+13] Adam D Law, Mélodie Auriol, Dean Smith, Tommy S Horozov, and D Martin A Buzza. Self-assembly of two-dimensional colloidal clusters by tuning the hydrophobicity, composition, and packing geometry. *Physical Review Letters*, 110(13):138301, 2013.

[LBF04] H Lefebvre-Brion and R W Field. *The Spectra and Dynamics of Diatomic Molecules*. Elsevier, New York, 2004.

[LBL12] Mingwu Lu, Nathaniel Q Burdick, and Benjamin L Lev. Quantum Degenerate Dipolar Fermi Gas. *Physical Review Letters*, 108(21):215301, May 2012.

[LBYL11] Mingwu Lu, Nathaniel Q. Burdick, Seo Ho Youn, and Benjamin L. Lev. Strongly dipolar bose-einstein condensate of dysprosium. *Physical Review Letters*, 107:190401, 2011.

[LCD+87] A. Leggett, S. Chakravarty, A. T. Dorsey, Matthew P. A. Fisher, A. Garg, and W. Zwerger. Dynamics of the dissipative two-state system. *Rev. Mod. Phys.*, 59:1, 1987.

[LCH+17] M H Lee, C P Chang, F T Huang, G Y Guo, B Gao, C H Chen, S W Cheong, and M W Chu. Hidden Antipolar Order Parameter and Entangled Néel-Type Charged Domain Walls in Hybrid Improper Ferroelectrics. *Physical Review Letters*, 119(15):157601–6, October 2017.

[LCND08] Roman M. Lutchyn, Łukasz Cywiński, Cody P. Nave, and S. Das Sarma. Quantum decoherence of a charge qubit in a spin-fermion model. *Physical Review B*, 78:024508, 2008.

[LdJ04] Liangbin Li and Wim H de Jeu. Shear-induced smectic ordering in the melt of isotactic polypropylene. *Physical Review Letters*, 92(7):075506, 2004.

[Leh99] K. K. Lehmann. Potential of a neutral impurity in a large ^{4}He cluster. *Mol. Phys.*, 97:645, 1999.

[Lem17] Mikhail Lemeshko. Quasiparticle approach to molecules interacting with quantum solvents. *Physical Review Letters*, 118:095301, 2017.

[LF09] Mikhail Lemeshko and Bretislav Friedrich. Rotational and rotationless states of weakly bound molecules. *Physical Review A*, 79, 2009.

[LKA+20] Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*, 2020.

[LL13] Lev Davidovich Landau and Evgenii Mikhailovich Lifshitz. *Quantum mechanics: non-relativistic theory*, volume 3. Elsevier, 2013.

[LLP53] TD Lee, FE Low, and Do Pines. The Motion of Slow Electrons in a Polar Crystal. *Physical Review*, 90(2):297, 1953.

[LP48] L D Landau and S I Pekar. Effective mass of a polaron. *Zh. Eksp. i Theor. Fiz.*, 18:419, 1948.

[LRT17] Cosimo Lupo and Federico Ricci-Tersenghi. Approximating the xy model on a random graph with a q-state clock model. *Physical Review B*, 95:054433, Feb 2017.

[LS17] Mikhail Lemeshko and Richard Schmidt. Molecular impurities interacting with a many-particle environment: from ultracold gases to helium nanodroplets. *Low Energy and Low Temperature Molecular Scattering*, 2017.

[LSVdJ08] Stijn Lievens, NI Stoilova, and Joris Van der Jeugt. A linear chain of interacting harmonic oscillators: solutions as a wigner quantum system. In *Journal of Physics: Conference Series*, volume 128, page 012028. IOP Publishing, 2008.

[LYB+20] Xiang Li, Enderalp Yakaboylu, Giacomo Bighin, Richard Schmidt, Mikhail Lemeshko, and Andreas Deuchert. Intermolecular forces and correlations mediated by a phonon bath. *The Journal of chemical physics*, 152(16):164302, 2020.

[Mah90] G D Mahan. *Many-particle physics*. Physics of solids and liquids. Plenum, New York, NY, 1990.

[MCBS17] Alejandro Mendoza-Coto, Daniel G Barci, and Daniel A Stariolo. Quantum and thermal melting of stripe forming systems with competing long-range interactions. *Physical Review B*, 95(14):175, April 2017.

[MCCC19] Roger G Melko, Giuseppe Carleo, Juan Carrasquilla, and J Ignacio Cirac. Restricted boltzmann machines in quantum physics. *Nature Physics*, 15(9):887–892, 2019.

[MCL19] Stefano Martiniani, Paul M Chaikin, and Dov Levine. Quantifying Hidden Order out of Equilibrium. *Physical Review X*, 9(1):011031–13, February 2019.

[MCS12] Alejandro Mendoza-Coto and Daniel A Stariolo. Coarse-grained models of stripe forming systems: Phase diagrams, anomalies, and scaling hypothesis. *Physical Review E*, 86(5):85, November 2012.

[MFKM10] Ryan McGorty, Jerome Fung, David Kaz, and Vinothan N Manoharan. Colloidal self-assembly at an interface. *Materials Today*, 13(6):34–42, 2010.

[MKL19] Johann H Mentink, MI Katsnelson, and Mikhail Lemeshko. Quantum many-body dynamics of the einstein–de haas effect. *Physical Review B*, 99(6):064428, 2019.

[MLKB14] Ethayaraja Mani, Wolfgang Lechner, Willem K. Kegel, and Peter G. Bolhuis. Equilibrium and non-equilibrium cluster phases in colloids with competing interactions. *Soft Matter*, 10:4479–4486, 2014.

[MLY20] Mikhail Maslov, Mikhail Lemeshko, and Enderalp Yakaboylu. Synthetic spin-orbit coupling mediated by a bosonic environment. *Physical Review B*, 101:184104, May 2020.

[Mon16] Guido Montúfar. Restricted Boltzmann Machines: Introduction and Review. In *Information Geometry and its Applications IV*, pages 75–115. Springer, 2016.

[MTSL16] Bikashkali Midya, Michał Tomza, Richard Schmidt, and Mikhail Lemeshko. Rotation of cold molecular ions inside a Bose-Einstein condensate. *Physical Review A*, 94:041601(R), 2016.

[Mus10] G. Mussardo. *Statistical field theory: an introduction to exactly solved models in statistical physics*. 2010.

[MZB14] Pietro Massignan, Matteo Zaccanti, and Georg M Bruun. Polarons, dressed molecules and itinerant ferromagnetism in ultracold Fermi gases. *Rep. Prog. Phys.*, 77:034401, 2014.

[NB99] M Newman and G Barkema. *Monte carlo methods in statistical physics chapter 1-4*. Oxford University Press: New York, USA, 1999.

[NBSN19] Murphy Yuezhen Niu, Sergio Boixo, Vadim N Smelyanskiy, and Hartmut Neven. Universal quantum control through deep reinforcement learning. *npj Quantum Information*, 5(1):1–8, 2019.

[NDD+19] Hendrik Poulsen Nautrup, Nicolas Delfosse, Vedran Dunjko, Hans J Briegel, and Nicolai Friis. Optimizing quantum error correction codes with reinforcement learning. *Quantum*, 3:215, 2019.

[NDM+14] Arash Nikoubashman, Raleigh L Davis, Brian T Michal, Paul M Chaikin, Richard A Register, and Athanassios Z Panagiotopoulos. Thin films of homopolymers and cylinder-forming diblock copolymers under shear. *ACS nano*, 8(8):8015–8026, 2014.

[NDYI17] Yusuke Nomura, Andrew S. Darmawan, Youhei Yamaji, and Masatoshi Imada. Restricted Boltzmann machine learning for solving strongly correlated quantum systems. *Physical Review B*, 96:205152, Nov 2017.

[NGM+04] Kostya S Novoselov, Andre K Geim, Sergei V Morozov, De-eng Jiang, Yanshui Zhang, Sergey V Dubonos, Irina V Grigorieva, and Alexandr A Firsov. Electric field effect in atomically thin carbon films. *science*, 306(5696):666–669, 2004.

[NH10] Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In *Icml*, 2010.

[Nom20] Yusuke Nomura. Machine learning quantum states—extensions to fermion–boson coupled systems and excited-state calculations. *Journal of the Physical Society of Japan*, 89(5):054706, 2020.

[NS19] Alexandra Nagy and Vincenzo Savona. Variational quantum monte carlo method with a neural-network ansatz for open quantum systems. *Physical Review Letters*, 122:250501, Jun 2019.

[OE75] J Oitmaa and I G Enting. Critical behaviour of a two-layer Ising system. *Journal of Physics A: Mathematical and General*, 8(7):1097–1114, jul 1975.

[OI14] Takahiro Ohgoe and Masatoshi Imada. Variational monte carlo method for electron-phonon coupled systems. *Physical Review B*, 89(19):195139, 2014.

[Ons44] Lars Onsager. Crystal statistics. i. a two-dimensional model with an order-disorder transition. *Physical Review*, 65:117–149, Feb 1944.

[ORRB10] C. J. Olson Reichhardt, C. Reichhardt, and A. R. Bishop. Structural transitions, melting, and intermediate phases for stripe- and clump-forming systems. *Physical Review E*, 82:041502, Oct 2010.

[Orú14] Román Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics*, 349:117–158, 2014.

[OS01] Manfred Opper and David Saad. *Advanced mean field methods: Theory and practice*. MIT press, 2001.

[PAdSN+10] Colin V Parker, Pegor Aynajian, Eduardo H da Silva Neto, Aakash Pushp, Shimpei Ono, Jinsheng Wen, Zhijun Xu, Genda Gu, and Ali Yazdani. Fluctuating stripes at the onset of the pseudogap in the high-Tc superconductor Bi2Sr2CaCu2O8+x. *Nature*, 468(7324):677–680, December 2010.

[PBC19] Jakub Pekalski, Eldar Bildanau, and Alina Ciach. Self-assembly of spiral patterns in confined systems with competing interactions. *Soft Matter*, 15:7715–7721, 2019.

[PCA13] Jakub Pekalski, Alina Ciach, and Noé G Almarza. Periodic ordering of clusters in a one-dimensional lattice model. *The Journal of chemical physics*, 138(14):144903, 2013.

[Pek46] Sl Pekar. Local quantum states of electrons in an ideal ion crystal. *Zhurnal Eksperimentalnoi I Teoreticheskoi Fiziki*, 16(4):341–348, 1946.

[Pli95] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1–19, 1995.

[PMSL95] Thomas Palberg, Wolfgang Mönch, Jürgen Schwarz, and Paul Leiderer. Grain size control in polycrystalline colloidal solids. *Journal of Chemical Physics*, 102(12):5082–5087, 1995.

[Pol64] Boris T Polyak. Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5):1–17, 1964.

[Pre18] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018.

[PRP20] Jakub Pękalski, Wojciech Rządkowski, and Athanassios Z Panagiotopoulos. Shear-induced ordering in systems with competing interactions: A machine learning study. *The Journal of chemical physics*, 152(20):204905, 2020.

[PS00] N V Prokofev and P C E Stamp. Theory of the spin bath. *Rep. Prog. Phys.*, 63:669, 2000.

[PS16] L P Pitaevskii and S Stringari. *Bose-Einstein Condensation and Superfluidity*. Oxford University Press, 2016.

[PWZ15] Jee Woo Park, Sebastian A Will, and Martin W Zwierlein. Ultracold Dipolar Gas of Fermionic Na23K40Molecules in Their Absolute Ground State. *Physical Review Letters*, 114(20):205302–5, May 2015.

[Qia99] Ning Qian. On the momentum term in gradient descent learning algorithms. *Neural networks*, 12(1):145–151, 1999.

[RBB80] S. T. Ruggiero, T. W. Barbee, and M. R. Beasley. Superconductivity in quasi-two-dimensional layered composites. *Physical Review Letters*, 45:1299–1302, Oct 1980.

[RCGLV16] R. Rodríguez-Cantano, T. González-Lezana, and P. Villarreal. Path integral Monte Carlo investigations on doped helium clusters. *Int. Rev. Phys. Chem.*, 35:37, 2016.

[RDC+20] Wojciech Rządkowski, Nicolo Defenu, Silvia Chiacchiera, Andrea Trombettoni, and Giacomo Bighin. Detecting composite orders in layered models via machine learning. *New Journal of Physics*, 22(9):093026, 2020.

[RFGZ19] José Ruiz-Franco, Nicoletta Gnan, and Emanuela Zaccarelli. Rheological investigation of gels formed by competing interactions: A numerical study. *Journal of Chemical Physics*, 150(2):024905, 2019.

[RKT+19] Benno S Rem, Niklas Käming, Matthias Tarnowski, Luca Asteria, Nick Fläschner, Christoph Becker, Klaus Sengstock, and Christof Weitenberg. Identifying quantum phase transitions using artificial neural networks on experimental data. *Nature Physics*, page 1, 2019.

[RL16] Elena S. Redchenko and Mikhail Lemeshko. Libration of strongly-oriented polar molecules inside a superfluid. *Chem. Phys. Chem.*, 17:3649, 2016.

[RL18] Wojciech Rzadkowski and Mikhail Lemeshko. Effect of a magnetic field on molecule–solvent angular momentum transfer. *The Journal of chemical physics*, 148(10):104307, 2018.

[RLD13] Paul L. Raston, Tao Liang, and Gary E. Douberly. Anomalous Λ-Doubling in the Infrared Spectrum of the Hydroxyl Radical in Helium Nanodroplets. *The Journal of Physical Chemistry A*, 117(34):8103–8110, 2013. PMID: 23373600.

[RLM21] Wojciech Rzadkowski, Mikhail Lemeshko, and Johan H Mentink. Artificial neural network states for non-additive systems. *arXiv preprint arXiv:2105.15193*, 2021.

[RM20] Jonas B. Rigo and Andrew K. Mitchell. Machine learning effective models for quantum systems. *Physical Review B*, 101:241105, Jun 2020.

[Roy18] C Patrick Royall. Hunting mermaids in real space: known knowns, known unknowns and unknown unknowns. *Soft Matter*, 14(20):4020–4028, 2018.

[RSM+15] Scott C Riggs, M C Shapiro, Akash V Maharaj, S Raghu, E D Bauer, R E Baumbach, P Giraldo-Gallo, Mark Wartenbe, and I R Fisher. Evidence for a nematic component to the hidden-order parameter in URu2Si2 from differential elastoresistance measurements. *Nature Communications*, 6(1):2727–6, March 2015.

[Sai17] Hiroki Saito. Solving the Bose–Hubbard model with machine learning. *Journal of the Physical Society of Japan*, 86(9):093001, 2017.

[SB16] Karl A H Sellin and Egor Babaev. First-order phase transition and tricritical point in multiband U(1)London superconductors. *Physical Review B*, 93(5):503–5, February 2016.

[SB18] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[SBP15] Boris V Svistunov, Egor S Babaev, and Nikolay V Prokof'ev. *Superfluid states of matter*. Crc Press, 2015.

[ŠC12] Anđela Šarić and Angelo Cacciuto. Fluid membranes can drive linear aggregation of adsorbed spherical nanoparticles. *Physical review letters*, 108(11):118101, 2012.

[ŠC13] Anđela Šarić and Angelo Cacciuto. Self-assembly of nanoparticles adsorbed on fluid and elastic membranes. *Soft Matter*, 9(29):6677–6695, 2013.

[Sch87] Jürgen Schmidhuber. *Evolutionary principles in self-referential learning, or on learning how to learn: the meta-meta-... hook*. PhD thesis, Technische Universität München, 1987.

[Sch10] Marco Schiró. Real-time dynamics in quantum impurity models with diagrammatic monte carlo. *Physical Review B*, 81:085126, Feb 2010.

[SCR07a] Sandro Sorella, Michele Casula, and Dario Rocca. Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods. *The Journal of Chemical Physics*, 127(1):014105, 2007.

[SCR07b] Sandro Sorella, Michele Casula, and Dario Rocca. Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods. *The Journal of chemical physics*, 127(1):014105, 2007.

[SFB20] Jonas Schuff, Lukas J Fiderer, and Daniel Braun. Improving the dynamics of quantum sensors with reinforcement learning. *New Journal of Physics*, 22(3):035001, 2020.

[SG99] R. P. Sear and W. M. Gelbart. Microphase separation versus the vapor-liquid transition in systems of spherical particles. *Journal of Chemical Physics*, 110:4582, 1999.

[SH20] Markus Schmitt and Markus Heyl. Quantum many-body dynamics in two dimensions with artificial neural networks. *Physical Review Letters*, 125:100503, Sep 2020.

[SHM+16] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

[SK17] Hiroki Saito and Masaya Kato. Machine learning technique to find quantum many-body ground states of bosons on a lattice. *Journal of the Physical Society of Japan*, 87(1):014001, 2017.

[SL06] Frank Stienkemeier and Kevin K Lehmann. Spectroscopy and dynamics in helium nanodroplets. *J. Phys. B*, 39:R127, 2006.

[SL15] Richard Schmidt and Mikhail Lemeshko. Rotation of Quantum Impurities in the Presence of a Many-Body Environment. *Physical Review Letters*, 114(20):203001, may 2015.

[SL16] Richard Schmidt and Mikhail Lemeshko. Deformation of a Quantum Many-Particle System by a Rotating Impurity. *Physical Review X*, 6(1):011012, feb 2016.

[SMB10] Dominik Scherer, Andreas Müller, and Sven Behnke. Evaluation of pooling operations in convolutional architectures for object recognition. In *International conference on artificial neural networks*, pages 92–101. Springer, 2010.

[SMDS11] John K Salmon, Mark A Moraes, Ron O Dror, and David E Shaw. Parallel random numbers: as easy as 1, 2, 3. In *Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 1–12, 2011.

[SMZT04] Francesco Sciortino, Stefano Mossa, Emanuela Zaccarelli, and Piero Tartaglia. Equilibrium cluster phases and low-density arrested disordered states: the role of short-range attraction and long-range repulsion. *Physical review letters*, 93(5):055701, 2004.

[SPP17] Andrew P Santos, Jakub Pekalski, and Athanassios Z Panagiotopoulos. Thermodynamic signatures and cluster properties of self-assembly in systems with competing interactions. *Soft Matter*, 13(44):8055–8063, 2017.

[SR18] Daniel Stopper and Roland Roth. Nonequilibrium phase transitions of sheared colloidal microphases: Results from dynamical density functional theory. *Physical Review E*, 97:062602, 2018.

[SS96] Jesús Salas and Alan D Sokal. Dynamic critical behavior of a Swendsen-Wang-type algorithm for the Ashkin-Teller model. *Journal of Statistical Physics*, 85(3-4):297–361, January 1996.

[SS98] T. Schneider and J.M. Singer. *Phase transition approach to high temperature superconductivity*. 1998.

[SS11] David Sholl and Janice A Steckel. *Density functional theory: a practical introduction*. John Wiley & Sons, 2011.

[SSBS05] J Smiseth, E Smørgrav, E Babaev, and A Sudbø. Field- and temperature-induced topological phase transitions in the three-dimensional N-component London superconductor. *Physical Review B*, 71(21):12–41, June 2005.

[SSC+04a] Anna Stradner, Helen Sedgwick, Frédéric Cardinaux, Wilson CK Poon, Stefan U Egelhaaf, and Peter Schurtenberger. Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature*, 432(7016):492–495, 2004.

[SSC+04b] Anna Stradner, Helen Sedgwick, Frédéric Cardinaux, Wilson CK Poon, Stefan U Egelhaaf, and Peter Schurtenberger. Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature*, 432:492, 2004.

[SSC+17] Benjamin Shepperson, Anders A. Søndergaard, Lars Christiansen, Jan Kaczmarczyk, Robert E. Zillich, Mikhail Lemeshko, and Henrik Stapelfeldt. Laser-induced alignment of iodine molecules in He-nanodroplets: Long-time coherence, revivals, and breaking-free. *Physical Review Letters*, 118:203203, 2017.

[SSH79] W Su, John R Schrieffer, and Alan J Heeger. Solitons in polyacetylene. *Physical review letters*, 42(25):1698, 1979.

[SSH80] W Su, John R Schrieffer, and Alan J Heeger. Soliton excitations in polyacetylene. *Physical Review B*, 22(4):2099, 1980.

[SSJ+21] Magnus G Skou, Thomas G Skov, Nils B Jørgensen, Kristian K Nielsen, Arturo Camacho-Guardian, Thomas Pohl, Georg M Bruun, and Jan J Arlt. Non-equilibrium quantum dynamics and formation of the bose polaron. *Nature Physics*, pages 1–5, 2021.

[Sta20] Felix Stahlberg. Neural machine translation: A review. *Journal of Artificial Intelligence Research*, 69:343–418, 2020.

[Stu10] Alexander Stukowski. Visualization and analysis of atomistic simulation data with ovito–the open visualization tool. *Modelling and Simulation in Materials Science and Engineering*, 18(1):015012, 2010.

[STZ05] F. Sciortino, P. Tartaglia, and E. Zaccarelli. One-dimensional cluster growth and branching gels in colloidal systems with short-range depletion attraction and screened electrostatic repulsion. *Journal of physical chemistry B*, 109:21942, 2005.

[SW87] R H Swendsen and J S Wang. Nonuniversal Critical-Dynamics in Monte-Carlo Simulations. *Physical Review Letters*, 58(2):86–88, 1987.

[Sza08a] Krzysztof Szalewicz. Interplay between theory and experiment in investigations of molecules embedded in superfluid helium nanodroplets. *International Reviews in Physical Chemistry*, 27(2):273–316, 2008.

[Sza08b] Krzysztof Szalewicz. Interplay between theory and experiment in investigations of molecules embedded in superfluid helium nanodroplets. *Int. Rev. Phys. Chem.*, 27:273, 2008.

[TAAS13] Sumesh P Thampi, Santosh Ansumali, R Adhikari, and Sauro Succi. Isotropic discrete Laplacian operators from lattice hydrodynamics. *Journal of Computational Physics*, 234:1–7, February 2013.

[TBT+16] Makariy A Tanatar, Anna E Böhmer, Erik I Timmons, M Schütt, G Drachuck, Valentin Taufour, K Kothapalli, Andreas Kreyssig, Sergey L Bud'ko, Paul C Canfield, Rafael M Fernandes, and Ruslan Prozorov. Origin of the Resistivity Anisotropy in the Nematic Phase of FeSe. *Physical Review Letters*, 117(12):127001, September 2016.

[TD17] Billy D Todd and Peter J Daivis. *Nonequilibrium molecular dynamics: theory, algorithms and applications*. Cambridge University Press, 2017.

[TIH+16] T. Tsatsoulis, C. Illg, M. Haag, B. Y. Mueller, L. Zhang, and M. Fähnle. Ultrafast demagnetization after femtosecond laser pulses: Transfer of angular momentum from the electronic system to magnetoelastic spin-phonon modes. *Physical Review B*, 93:134411, 2016.

[Tin96] M. Tinkham. *Introduction to Superconductivity*. 1996.

[TM02] Jian Tang and ARW McKellar. Infrared spectra of n 2 o–4 he, n 2 o–3 he, and ocs–3 he complexes. *The Journal of chemical physics*, 117(6):2586–2591, 2002.

[TM18] Giacomo Torlai and Roger G. Melko. Latent space purification via neural density operators. *Physical Review Letters*, 120:240503, Jun 2018.

[TP15] W. Töws and G. M. Pastor. Many-Body Theory of Ultrafast Demagnetization and Angular Momentum Transfer in Ferromagnetic Transition Metals. *Physical Review Letters*, 115(21):217204, nov 2015.

[Tsv07] Alexei M Tsvelik. *Quantum field theory in condensed matter physics*. Cambridge university press, 2007.

[TSZ09] Juan Carlos Fernandez Toledano, Francesco Sciortino, and Emanuela Zaccarelli. Colloidal systems with competing interactions: from an arrested repulsive cluster phase to a gel. *Soft Matter*, 5(12):2390–2398, 2009.

[TV04] J Peter Toennies and Andrey F Vilesov. Superfluid helium droplets: A uniquely cold nanomatrix for molecules and molecular complexes. *Ang. Chem. Int. Ed.*, 43(20):2622–2648, 2004.

[VBRC19] Filippo Vicentini, Alberto Biella, Nicolas Regnault, and Cristiano Ciuti. Variational neural-network ansatz for steady states in open quantum systems. *Physical Review Letters*, 122:250503, Jun 2019.

[VCN+19] Tom Vieijra, Corneel Casert, Jannes Nys, Wesley De Neve, Jutho Haegeman, Jan Ryckebusch, and Frank Verstraete. Restricted boltzmann machines for quantum states with nonabelian or anyonic symmetries. *arXiv preprint arXiv:1905.06034*, 2019.

[VHS+21] Filippo Vicentini, Damian Hofmann, Attila Szabó, Dian Wu, Christopher Roth, Clemens Giuliani, Gabriel Pescia, Jannes Nys, Vladimir Vargas-Calderon, Nikita Astrakhantsev, et al. Netket 3: Machine learning toolbox for many-body quantum systems. *arXiv preprint arXiv:2112.10526*, 2021.

[VHSBK18] Rodrigo A. Vargas-Hernández, John Sous, Mona Berciu, and Roman V. Krems. Extrapolating quantum observables with machine learning: Inferring multiple phase transitions from properties of a single phase. *Physical Review Letters*, 121:255702, Dec 2018.

[VMK88] Dmitri A. Varshalovich, A N Moskalev, and V K Khersonskii. *Quantum theory of angular momentum*. World Scientific, 1988.

[VTHS13] Romain Vasseur, Kien Trinh, Stephan Haas, and Hubert Saleur. Crossover physics in the nonequilibrium dynamics of quenched quantum impurity systems. *Physical Review Letters*, 110:240601, Jun 2013.

[VVG+16] Venema Liesbeth, Verberck Bart, Georgescu Iulia, Prando Giacomo, Coudert Elsa, Milana Silvia, Maragkou Maria, Persechini Lina, Pacchioni Giulia, and Fleet Luke. The quasiparticle zoo. *Nature Physics*, 12:1085, dec 2016.

[VZ06] C M Varma and Lijun Zhu. Helicity Order: Hidden Order Parameter in URu2Si2. *Physical Review Letters*, 96(3):1265–4, January 2006.

[Wan16] Lei Wang. Discovering phase transitions with unsupervised learning. *Physical Review B*, 94:195105, Nov 2016.

[WB09] David Waltz and Bruce G Buchanan. Automating science. *Science*, 324(5923):43–44, 2009.

[Wei12] U. Weiss. *Quantum Dissipative Systems*. World Scientific, 4 edition, 2012.

[WG02] George M Whitesides and Bartosz Grzybowski. Self-assembly at all scales. *Science*, 295(5564):2418–2421, 2002.

[Wol89] U Wolff. Collective Monte-Carlo Updating for Spin Systems. *Physical Review Letters*, 62(4):361–364, 1989.

[WS17] Sebastian J. Wetzel and Manuel Scherzer. Machine learning of explicit order parameters: From the ising model to su(2) lattice gauge theory. *Physical Review B*, 96:184410, Nov 2017.

[Wu82] F. Y. Wu. The Potts model. *Rev. Mod. Phys.*, 54:235–268, Jan 1982.

[YDL17] Enderalp Yakaboylu, Andreas Deuchert, and Mikhail Lemeshko. Emergence of non-abelian magnetic monopoles in a quantum impurity problem. *Physical Review Letters*, 119:235301, Dec 2017.

[YGL+20] Enderalp Yakaboylu, Areg Ghazaryan, Douglas Lundholm, Nicolas Rougerie, Mikhail Lemeshko, and Robert Seiringer. Quantum impurity model for anyons. *Physical Review B*, 102:144109, Oct 2020.

[YH19] Nobuyuki Yoshioka and Ryusuke Hamazaki. Constructing neural stationary states for open quantum many-body systems. *Physical Review B*, 99:214306, Jun 2019.

[YL17] Enderalp Yakaboylu and Mikhail Lemeshko. Anomalous Screening of Quantum Impurities by a Neutral Environment. *Physical Review Letters*, 118:085302, 2017.

[YL18] Enderalp Yakaboylu and Mikhail Lemeshko. Anyonic statistics of quantum impurities in two dimensions. *Physical Review B*, 98:045402, Jul 2018.

[YMD+18] Enderalp Yakaboylu, Bikashkali Midya, Andreas Deuchert, Nikolai Leopold, and Mikhail Lemeshko. Theory of the rotating polaron: Spectrum and self-localization. *Physical Review B*, 98:224506, Dec 2018.

[YSL18] Enderalp Yakaboylu, Mikhail Shkolnikov, and Mikhail Lemeshko. Quantum groups as hidden symmetries of quantum impurities. *Physical review letters*, 121(25):255302, 2018.

[ZC16a] Yuan Zhuang and Patrick Charbonneau. Equilibrium phase behavior of the square-well linear microphase-forming model. *Journal of physical chemistry B*, 120(26):6178–6188, 2016.

[ZC16b] Yuan Zhuang and Patrick Charbonneau. Recent advances in the theory and simulation of model colloidal microphase formers. *Journal of physical chemistry B*, 120(32):7775–7782, 2016.

[ZKW04] RE Zillich, Y Kwon, and KB Whaley. Roton-rotation coupling of acetylene in h 4 e. *Physical review letters*, 93(25):250401, 2004.

[ZMF+19] Yi Zhang, A Mesaros, Kazuhiro Fujita, SD Edkins, MH Hamidian, K Ch'ng, H Eisaki, S Uchida, JC Séamus Davis, Ehsan Khatami, and Eun-Ah Kim. Machine learning in electronic-quantum-matter imaging experiments. *Nature*, 570(7762):484–490, June 2019.

[ZN14] Lifa Zhang and Qian Niu. Angular Momentum of Phonons and the Einstein–de Haas Effect. *Physical Review Letters*, 112(8):085503, feb 2014.

[ZWA+19] Xiao-Ming Zhang, Zezhu Wei, Raza Asad, Xu-Chen Yang, and Xin Wang. When does reinforcement learning stand out in quantum control? a comparative study on state preparation. *npj Quantum Information*, 5(1):1–7, 2019.

[ZZC16] Yuan Zhuang, Kai Zhang, and Patrick Charbonneau. Equilibrium phase behavior of a continuous-space microphase former. *Physical Review Letters*, 116:098301, 2016.