Dissertation

submitted to the
Combined Faculties of the Natural Sciences and Mathematics
of the Ruperto-Carola-University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Put forward by

Sebastian Johann Wetzel

born in Lauingen a.d. Donau (Germany)

Oral Examination: 9.5.2018

Exploring Phase Diagrams with Functional Renormalization and Artificial Neural Networks: From the Hubbard Model to Lattice Gauge Theory

Referees: Prof. Dr. Christof Wetterich
Prof. Dr. Thomas Gasenzer

Abstract

This thesis is dedicated to provide physicists with new and improved techniques to examine phase diagrams and phase transitions.

One the one hand, an analysis of the effects of different regularization schemes in functional renormalization group calculations is provided. Building on this knowledge, an investigation of the phase diagram of the Hubbard-Model on the square lattice is performed using the functional renormalization group. The calculation reveals leading instabilities in the d-wave superconducting and different antiferromagnetic channels. In the symmetry broken phases there are a changing Fermi surface geometry, coexistence phases of d-wave superconductivity and antiferromagnetism as well as a mutual tendency of superconductivity and antiferromagnetism to repel each other.

On the other hand, a scheme to discover phase transitions using unsupervised artificial neural networks is developed. Further, a method to interpret artificial neural networks is introduced. These methods are applied to systems ranging from the two dimensional Ising Model to four dimensional SU(2) lattice gauge theory. They find the existence of different phases, calculate phase boundaries and derive the explicit formulas of the quantities by which the neural network distinguishes between phases. It turns out that these quantities are order parameters and other thermodynamic quantities.

Kurzfassung

Diese Arbeit stellt Physikern neue und verbesserte Techniken zur Untersuchung von Phasendiagrammen und Phasenübergängen zur Verfügung.

Einerseits wird eine Analyse der Effekte verschiedener Regularisierungsschemata in funktionalen Renormierungsgruppenberechnungen präsentiert. Aufbauend auf diesem Wissen wird eine Untersuchung des Phasendiagramms des Hubbard-Modells auf dem Quadratgitter unter Verwendung der funktionalen Renormierungsgruppe durchgeführt. Die Berechnung zeigt führende Instabilitäten in supraleitenden d-Wellen und verschiedenen antiferromagnetischen Kanälen. In den symmetriegebrochenen Phasen gibt es eine sich ändernde Fermiflächengeometrie, Koexistenzphasen von d-Wellen-Supraleitung und Antiferromagnetismus sowie eine gegenseitige Tendenz, dass sich Supraleitung und Antiferromagnetismus abstoßen.

Andererseits wird ein Schema zur Entdeckung von Phasenübergängen unter Verwendung von unüberwachten künstlichen neuronalen Netzen entwickelt. Ferner wird ein Verfahren zur Interpretation von künstlichen neuronalen Netzen eingeführt. Diese Methoden werden auf Systeme angewendet, die vom zweidimensionalen Ising-Modell bis zur vierdimensionalen SU(2) Gittereichtheorie reichen. Sie finden die Existenz verschiedener Phasen, berechnen Phasengrenzen und leiten die expliziten Formeln der Größen ab, durch die das neuronale Netz zwischen Phasen unterscheidet. Es stellt sich heraus, dass diese Größen die Ordnungsparameter und andere thermodynamische Größen sind.

Contents

1 Publications

This thesis contains the following publications:

[1] **Physics and the choice of regulators in functional renormalisation group flows**

Pawlowski, Jan M.; Scherer, Michael M.; Schmidt, Richard; Wetzel, Sebastian J.
Annals of Physics,Volume 384, Pages 165-197, September 2017
doi: 10.1016/j.aop.2017.06.017

[2] **Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders**

Wetzel, Sebastian J.
Phys. Rev. E 96, 022140,18 August 2017
doi: 10.1103/PhysRevE.96.022140

[3] **Machine learning of explicit order parameters: From the Ising model to SU(2) lattice gauge theory**

Wetzel, Sebastian J.; Scherzer Manuel
Phys. Rev. B 96, 184410, 8 November 2017
doi: 10.1103/PhysRevB.96.184410

[4] **Exploring the Hubbard Model on the Square Lattice at Zero Temperature with a Bosonized Functional Renormalization Approach**

Wetzel, Sebastian J.
eprint arXiv:1712.04297

2 Introduction

2.1 Phase Diagrams

Phase diagrams can be understood as a comprehensible collection of macroscopic knowledge about physical systems. They describe under which conditions, internal structure or external influence a given physical system exhibits certain macroscopic properties. A phase diagram contains regions of certain phases of a physical system. They are divided from each other by either phase transitions, where the system undergoes a drastic macroscopic change, or crossovers, where the system changes continuously. Phase diagrams are used to describe the macroscopic properties in a wide range of physical disciplines. A first phase diagram many people learn in school is the phase diagram of water. This is based on the fact that the Celsius temperature scale is aligned at phase transitions from ice to liquid water and liquid water to steam. Furthermore, phase diagrams are used to describe phases of magnets, quantum chromodynamics (QCD), or superconductivity. In order to draw phase diagrams, one needs to determine the nature of the occurring phases and the location and the type of the transition. While the nature of the phases is unique for each system, phase transitions can be grouped into certain classes. Ehrenfest categorized phase transitions by a jump in the n-th derivative of the free energy as a phase transition of n-th order [5]. Landau further explained that almost all phases can be attributed to symmetries of the physical system [6]. The symmetry is manifested on one side of the phase transition, while it is broken on the other side. In the vicinity of phase transitions of second or higher order, the behavior of the free energy can be described by so-called critical exponents. They only depend on the dimensionality of the examined system and on the symmetry which is broken at the phase transition, but not on the microscopic details. This phenomenon is called universality. Beyond these classifications, lines of phase transitions can cross to exhibit multi-critical phenomena or end in critical end points.

In physics very few phase diagrams can be calculated analytically. The most prominent model which is analytically solvable and possesses a phase transition is the two-dimensional Ising Model, which describes spins on a square lattice [7]. It

was solved[1] by Onsager in 1944 [8].

Recently new types of phase transitions have been discovered which could not be classified by either of the former schemes. For example, the Berezinskii-Kosterlitz-Thouless transition (BKT transition) is a phase transition in the two-dimensional XY model [9, 10]. It is a transition from bound vortex-antivortex pairs at low temperatures to unpaired vortices and anti-vortices above a critical temperature, or in other words the binding-unbinding transition of topological defects. In this system conventional phase transitions are forbidden by the Mermin-Wagner-Hohenberg theorem [11, 12]. However, the theorem does not apply in this special case, since the BKT transition is a phase transition of infinite order which does not break any symmetries. Another class of non-conventional phase transitions are topological phase transitions. They are zero temperature transitions between different topological orders. A prototype of a topological phase is the fractional quantum hall state [13, 14]. It is the effect that at low temperatures, the Hall conductance in a two-dimensional electron system in a magnetic field is quantized on plateaus at fractional values.

The understanding of phase diagrams is a central task for experimental and theoretical researchers. The phase diagrams of many systems like QCD or high temperature superconductors are still not fully explored. With the discovery of new types of materials and phases of matter, we face even more complicated phases. Hence, physics is in need of new techniques to calculate phase diagrams where conventional methods have not succeeded. In order to achieve this task for more complicated models we discuss two techniques in this thesis. We employ, examine and improve the functional renormalization group. Further, we develop new methods based on artificial neural networks, to automatically reveal phases with only little human supervision.

2.2 The Renormalization Group

The renormalization group is a method in theoretical physics which translates parameters of physical models between different energy scales. It can be utilized to calculate thermodynamic potentials like the free energy and thus obtain the partition function.

[1] Solving a model means calculating the partition function or an equivalent thermodynamic potential.

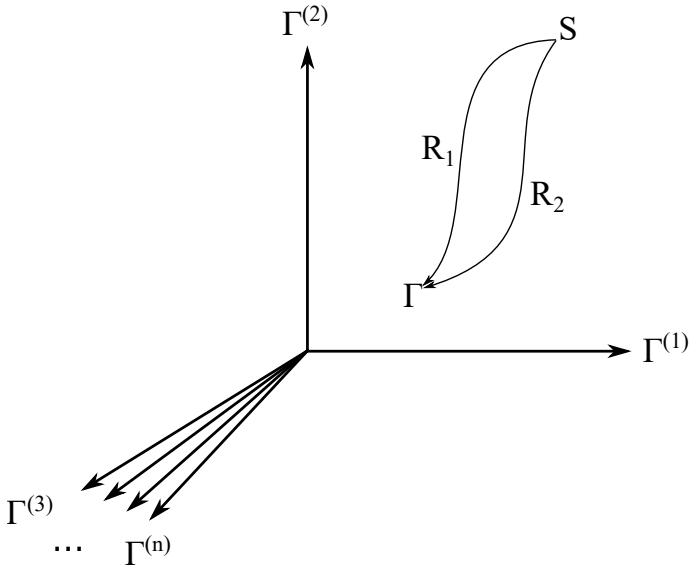

Figure 2.1: Renormalization group flow of the effective average action Γ_k in theory space. In the microscopic limit $k = \Lambda$ the initial conditions are the microscopic action S. In the macroscopic limit $k = 0$ the effective average action Γ_k approaches the full quantum effective action Γ. Different regulators R_1, R_2 induce different flow trajectories which have identical start and end points.

2.2.1 History of the Renormalization Group

Renormalization was first mentioned in 1953 [15] by Stueckelberg and Petermann. They discovered the existence of a renormalization transformation which transfers quantities in quantum field theory from the bare terms to the counter terms. In 1954 Gell-Mann and Low introduced the concept of a renormalization group transformation being related to the energy scale in quantum electrodynamics (QED) [16]. The modern version of such an energy scale dependent renormalization group transformation was introduced by Callan and Symanzik in 1970 [17]. In 1966 Kadanoff gave an explanation of the physical meaning of the renormalization process, and was able to incorporate physics on separate length scales via the "block-spin" renormalization group [18]. Wilson's ideas in 1971 [19, 20] were responsible for putting the renormalization group in a computational framework. With this development it was able to solve problems like the Kondo-Problem and could be employed to examine phase transitions and critical phenomena in detail. It was later reformulated in terms of exact equations by Wegner and Houghton [21] as well as Polchinski [22]. In this thesis we employ the exact renormalization flow equation for the effective average action which was devised by Wetterich in 1993 [23] and shortly afterwards by Morris [24].

2.2.2 The Functional Renormalization Group

The functional renormalization group equation, also known as Wetterich equation [23]

$$\partial_k \Gamma_k = \frac{1}{2} \mathrm{STr} \left[\left(\Gamma^{(2)} + R_k \right)^{-1} \partial_k R_k \right] , \tag{2.1}$$

is a functional differential equation for the effective average action Γ_k. This flow equation interpolates between the microscopic action S and the full quantum effective action Γ. The Wetterich equation generates this interpolation by integrating out quantum fluctuations on momentum scales larger than k. By lowering k continuously, the effective average action follows a flow trajectory in theory space, see Fig. 2.1. The regularization scheme is formulated in terms of the regulator R_k. The Wetterich equation is independent of the choice of the regulator as long as it obeys certain conditions. However, truncating the effective average action destroys this property. The magnitude of this effect is discussed in our paper *Physics and the Choice of Regulators in Functional Renormalisation Group Flows* [1]. In this publication we also propose how to optimize the renormalization group flow in order to obtain reliable physical results.

2.3 Artificial Neural Networks

Artificial neural networks, see Fig. 2.2, are computer algorithms that are based on statistical learning inspired by the human brain. They belong to the class of machine learning algorithms: without being explicitly programmed for a certain task, the computer is able to learn from data and generalize their knowledge to previously unseen examples.

2.3.1 History of Artificial Neural Networks

The elementary constituents of artificial neural networks have already been proposed in 1943 by McCulloch and Pitts [25] as a simple mathematical model for neurons. The first learning procedure for these neurons has been developed shortly after by Hebb [26]. The application of such neurons as a machine learning algorithm started in 1957 [27] when Rosenblatt employed the perceptron algorithm, built upon the McCulloch-Pitts neuron, for character recognition. By that time a large scale application of artificial neural networks was not possible due to limited computational

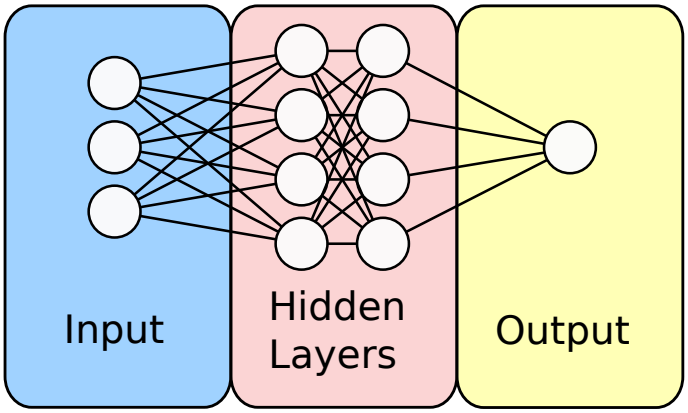

Figure 2.2: An artificial neural network is a weighted directed graph. In this picture we see a feed-forward neural network, in which neurons are connected with other neurons in the previous layer and the next layer. The input is processed through several hidden layers, after which an output is generated.

resources and the inefficiency of the training. Research on artificial neural networks stagnated until 1986 when it was revived by Rumelhart, Hinton and Williams [28]. They rediscovered the backpropagation algorithm, originally discovered by Werbos in 1974 [29], an efficient learning algorithm which is still employed today to train state-of-the-art artificial neural networks. By now our computational resources have grown so far that artificial neural networks are used to classify images [30], to recognize language [31] or to beat humans in complex games [32].

2.3.2 Artificial Neural Networks in Physics

Artificial neural networks have been applied to physics in its traditional form. This means that these algorithms have been trained to perform a certain task on some training data and applied to execute the same task on previously unseen data (supervised learning). For example, these algorithms have been applied to classify jet images from CERN [33]. In other cases physicists have applied machine learning algorithms to find structure in data without supervision (unsupervised learning). For example, unsupervised learning was employed to group different molecular systems [34], where the number of parameters is too large to for conventional approaches.

Since 2016 machine learning in physics has undergone major change. Instead of the conventional machine learning tasks, one is now able to simulate physical systems [35–45]. Furthermore, physicists have been able to answer specific questions about physical systems, up to the point that one can now extract complete phase diagrams. This is achieved by a combination of Monte-Carlo simulations and dif-

ferent machine learning algorithms, without the need for any physical deduction by humans. It is possible to calculate phase boundaries in the context of supervised learning [46], where it turned out that artificial neural networks outperform other machine learning algorithms on complicated models. Within one year artificial neural networks have successfully calculated accurate transition temperatures in systems with topological phases [47], in strongly correlated electron systems [48] and in lattice gauge theory [3]. Information about the existence of different phases in a physical system can be found by unsupervised learning algorithms [2, 49–51]. Furthermore, by interpreting the machine learning algorithms, one can extract explicit expressions of the order parameters, corresponding to the nature of the ordered phases [3, 49]. The importance of the interpretation is twofold: on the one hand, one can infer the dominant physical properties of the ordered phase. On the other hand, one is able to trust the results of machine learning calculations, if they base their results on physical quantities. This thesis includes the introduction of artificial neural network-based autoencoders to detect phases in an unsupervised manner, see *Unsupervised Learning of Phase Transitions: From Principal Component Analysis to Variational Autoencoders* [2]. More about how to find phase transitions in lattice gauge theory and how to interpret neural networks in such cases can be found in *Machine Learning of Explicit Order Parameters: From the Ising Model to SU(2) Lattice Gauge Theory* [3].

2.4 High-Temperature Superconductivity

In condensed matter physics, scientists examine emergent phenomena from the electromagnetic interaction between the building blocks of matter. One of the biggest challenges of condensed matter physics in the last three decades was to explain high-temperature superconductivity.

2.4.1 History of Superconductivity

The defining property of superconductivity is the vanishing of the electrical resistance below a critical temperature T_c. Conventional superconductivity was already found in 1911 by Kamerlingh Onnes in mercury below a critical temperature of $T_c = 4.2K$ [53]. In 1957 the underlying mechanism was explained by Bardeen, Cooper and Schrieffer in the so-called BCS theory [54]. At low temperatures, lattice vibrations introduce an attractive interaction between electrons, such that they form

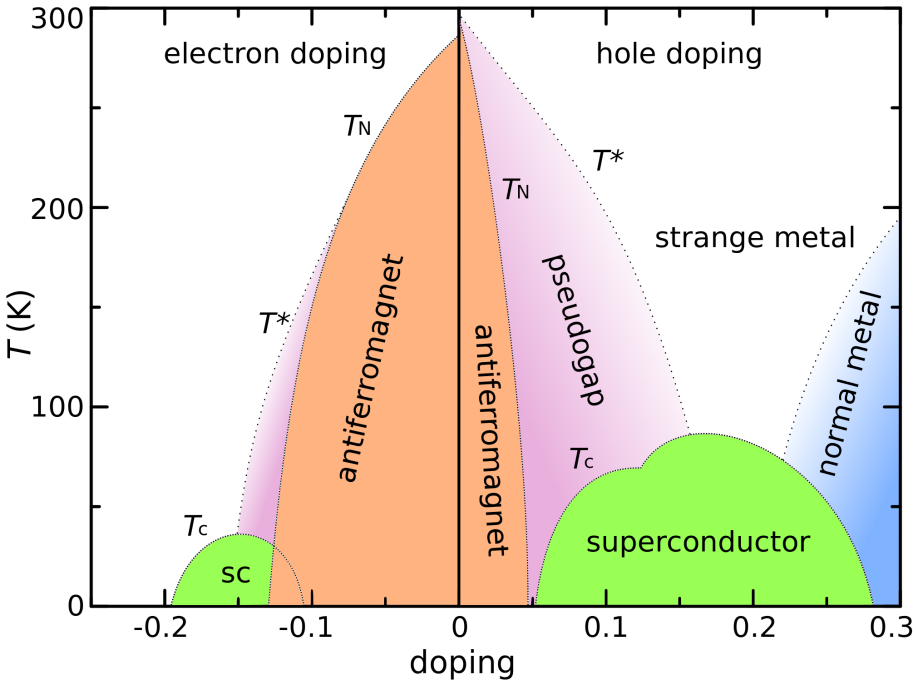

Figure 2.3: Schematic phase diagram of a cuprate ceramic superconductor. At low temperatures and zero doping the ground state is antiferromagnetic. However, certain regions of doping allow for d-wave superconductivity [52].

so-called Cooper pairs. Below the critical temperature T_c, the Cooper pairs condense and undergo a phase transition to a superfluid exhibiting s-wave symmetry. A different kind of superconductivity was discovered in 1986 in doped cuprate ceramics by Bednorz and Müller. While these materials are insulating at room temperature, the resistance drops to zero at a temperature of $35K$ [55]. In these superconductors the charge carriers are still Cooper pairs, though the formation of these pairs is no longer induced by lattice vibrations. Since this kind of superconductivity is based on a different mechanism and allows for much higher transition temperatures, it is called high-temperature superconductivity. In the last three decades the examination of the phase diagram of high-temperature superconductors, see Fig. 2.3, and the underlying mechanism of the Cooper pair formation have been subject of a huge research effort. Despite much progress, the phase diagram is still not fully understood. In recent years there have been discoveries of new types of superconductors like iron-based superconductors [56, 57] or hydrogen sulfide at high pressure [58, 59]. The highest known transition temperatures are at $T_c = 134K$ for cuprate ceramics, $T_c = 56K$ in iron-based superconductors and $T_c = 204K$ for pressurized hydrogen sulfide.

2.4.2 Hubbard Model

While many features of high-temperature superconductivity in cuprate ceramics are still not understood, it is believed that many properties can already be addressed by examining the electrons moving in the two-dimensional copper oxide layers. The most promising model to describe the electrons in these layers is the Hubbard Model on the square lattice [60–62]. It is defined by the Hamiltonian

$$H = -\sum_{ij,\sigma} t_{ij} c^{\dagger}_{i,\sigma} c_{j,\sigma} + h.c. + U \sum_{i} (c^{\dagger}_{i,\downarrow} c_{i,\downarrow})(c^{\dagger}_{i,\uparrow} c_{i,\uparrow}) \ . \tag{2.2}$$

Here $c^{\dagger}_{i,\sigma}$ and $c_{j,\sigma}$ are the creation and annihilation operators of the electrons. The first term describes the creation of electrons on one site while simultaneously removing an electron on another site. Hence, this expression represents the hopping of electrons from one fixed lattice site to another lattice site described by hopping parameters t_{ij}. In cuprates only transitions to neighboring lattice sites are allowed. Thus, a common convention is $t_{ij} = t$ for nearest neighbors, $t_{ij} = t'$ for next-to-nearest neighbors and $t_{ij} = 0$ otherwise. The last term counts electrons with opposing spins, up or down, on the same lattice site. According to the Pauli principle only one electron of each spin can occupy the same site. However, such a configuration is penalized by a Coulomb repulsion U. Even tough this model is relatively simple, it contains enough physics to explain most of the important features of cuprate superconductors, see Fig. 2.3. At low temperatures and zero doping the model correctly exhibits an insulating antiferromagnetic ground state. Doping induces a transition to a d-wave superconductor and even further doping can lead to a ferromagnetic ground state. Furthermore, there are many other phases, like pseudogap phases and a strange metal phase which are not completely understood. Until now the renormalization group has been employed to give insight into the physics of the Hubbard Model. Purely fermionic renormalization calculations revealed the leading instabilities of the Hubbard model in the antiferromagnetic and in the d-wave channel [63–78]. They were unsuccessful in incorporating the interplay between antiferromagentism and d-wave superconductivity. In order to analyze this interaction, a combination of different schemes like renormalization group and mean field calculations were employed [79–84]. The approach used in this thesis builds upon a bosonized renormalization group analysis of the Hubbard model [85–92]. In this scheme the truncation explicitly allows for emerging bosonic composite fields, corresponding to superconductivity and antiferromagnetism. Thus, the interplay between the different orders can be analyzed naturally within one consistent framework. In this thesis

we contribute to the understanding of the phase diagram of the Hubbard Model at zero temperature, for more details see *Exploring the Hubbard Model on the Square Lattice at Zero Temperature with a Bosonized Functional Renormalization Approach* [4]. While our results are consistent with purely fermionic zero temperature calculations [78, 83, 84], we have easier access to some parts of the phase diagram. We are able to resolve a transition between two different kinds of antiferromagnetism, commensurate and incommensurate. Furthermore, our renormalization flow can be continued into the symmetry broken phases, where we find a coexistence between antiferromagnetism and superconductivity, and their tendency to repel each other. Lastly, our flow equations are able to follow the deformation of the Fermi surface by antiferromagnetism.

2.5 Strong Force

While condensed matter physics deals with phenomena related to the electromagnetic force, quantum chromodynamics (QCD) is the theory of the strong interaction between quarks and gluons. They are constituents that make up the atomic nucleus on length scales far below what is considered in condensed matter physics. A prominent example where both research areas touch is nuclear fusion. Here, the short ranged strong interaction needs to overcome the Coulomb repulsion of the protons to form stable atomic nuclei under the release of energy.

2.5.1 History of Quantum Chromodynamics

QCD is the theory of the strong interaction. It was developed to describe many new particles which were discovered in the 1950s. These particles were later in 1962 given the name hadrons, from greek: “large”, “massive”. The first to succeed in finding a mathematical description of the new particles were Gell-Mann and Ne’eman [93]. At that point it was possible to group the hadrons into classes with similar properties and masses. In 1963 Gell-Mann and Zweig [94, 95] proposed that the structure of these groups could be explained if hadrons were composite particles consisting of three flavors of smaller particles - the quarks. Experimentalists tried to find evidence for these quarks, but failed, since they could not isolate quarks from the hadrons. Gell-Mann’s mathematical model worked too well to dismiss it, but at that point quarks were merely considered as convenient mathematical constructs, but not real particles. It later turned out that quarks are real particles, but are

confined to hadrons at low energies[2]. In 1973 Fritzsch, Leutwyler and Gell-Mann introduced the concept of quarks possessing a color charge which acts as the source of the strong force [96]. This force is mediated by gluons, which themselves are charged particles and thus interact with each other[3]. Their mathematical description is based on the general field theory developed in 1954 by Yang and Mills [97]. Gross, Politzer and Wilczek discovered in 1973 that the strong interaction becomes asymptotically weaker as the energy scale, or temperature scale, increases [98, 99] . This phenomenon, called asymptotic freedom, allowed physicists to explain many results of high-energy experiments by employing the quantum field theory technique of perturbation theory. Finally, in 1979, evidence of gluons was discovered in three-jet events at DESY. After the experiments got more precise, physicists where successful in verifying perturbative QCD to a very high accuracy at CERN.

2.5.2 Quantum Chromodynamics

QCD is based on the gauge group SU(3) for the color degree of freedom. It is defined by its Lagrangian density

$$\mathcal{L}_{QCD} = \underbrace{-\frac{1}{4} G^a_{\mu\nu} G^{\mu\nu,a}}_{\mathcal{L}_{gluon}} + \underbrace{\sum_f \bar{\psi}^b_f (i\gamma^\mu D_\mu - m_f)^{bc} \psi^c_f}_{\mathcal{L}_{quarks}} , \tag{2.3}$$

with the gauge covariant derivative $D_\mu = \partial_\mu - igA_\mu$ contracted with the Dirac matrices γ_μ. The first term $\mathcal{L}_{gluon}$ describes the gluons in a Yang-Mills theory. The gluon field strength tensor is the commutator of the covariant derivatives $G_{\mu\nu} = 1/ig[D_\mu, D_\nu]$. In contrast to electromagnetism, which is described by a U(1) gauge theory, SU(3) is a non-Abelian gauge group, this means the fields A_μ do not commute themselves $[A_\mu, A_\nu] \neq 0$. The second part $\mathcal{L}_{quarks}$ describes the dynamics of the quarks ψ_f, each with a different mass m_f. The gauge covariant derivative couples the gluons to the quarks.

The high temperature limit of QCD exhibits a quark-gluon plasma of asymptotically free interacting quarks and gluons. Since the interaction strength grows weak at high energies, it can be accessed by perturbation theory. QCD at low energies is inaccessible by conventional methods. In this regime the interaction strength grows up to a point where QCD can only be examined by non-perturbative methods. The most prominent features in this region of the phase diagram are the confinement of

[2]This observation is consistent with lattice QCD computations, but is not mathematically proven.
[3]This is in contrast to electromagnetism where the force is mediated by photons which do not carry any charge.

quarks and gluons and chiral symmetry breaking. Confinement is the observation that quarks and gluons cannot be found isolated in ordinary matter. Chiral symmetry is an independent symmetry associated with each chirality, like left or right handed. The classical Lagrangian for QCD couples left and right handed quark fields only through mass terms which causes the explicit breaking of this symmetry. Even without explicit mass terms, chiral symmetry is broken by the quark masses of the QCD vacuum. Chiral symmetry breaking is responsible for the mass generation of nucleons from elementary light quarks, accounting for approximately 99% of the mass of visible matter.

A clean definition of these both properties as phase indicators is only valid in certain limits. For vanishing explicit quark masses the chiral symmetry gets restored above a critical temperature. In the limit of infinitely heavy quarks a phase transition from a confining phase to a deconfined phase takes place. In full QCD with physical quarks none of these assumptions are appropriate. In quenched calculations where fermion fluctuations are neglected the chiral restoration and confinement-deconfinement temperature coincide within reasonable accuracy [100]. The high energy region, the quark-gluon plasma, and low energy regime, a gas of hadrons, are depicted in the schematic phase diagram Fig. 2.4.

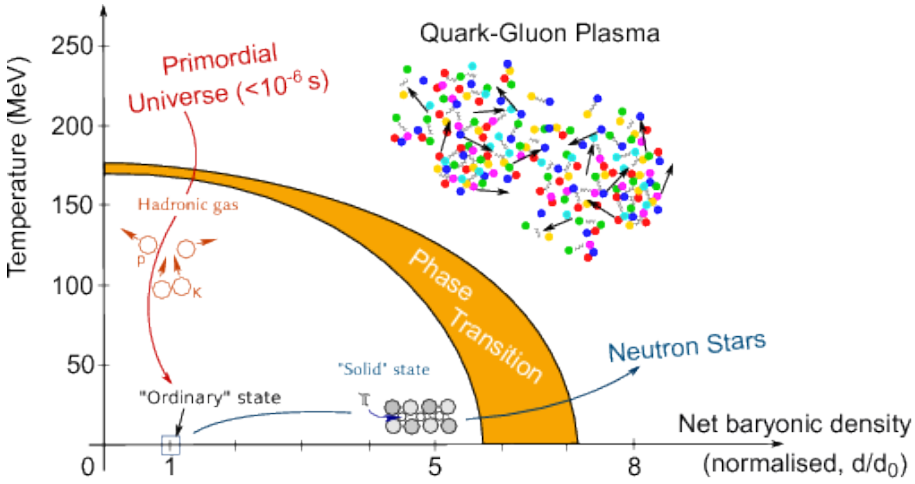

Figure 2.4: Phase diagram of QCD matter, ranging from regular nuclear matter to quark-gluon plasma. Regular matter is confined while the quark-gluon plasma is not [101].

2.5.3 Lattice Gauge Theory

The most common approach to QCD is lattice QCD, which amounts to a discretization of QCD on a spacetime lattice. One can link it to a continuous spacetime when making the lattice infinitely large while its sites move infinitesimally close to each other. Some of the most prominent features of QCD are already present without considering dynamical quarks. This assumption reduces lattice QCD to lattice gauge theory without fermions [102]. It can be employed to study pure gauge theory on a discrete spacetime lattice. It is the simplest model which shares one of the most important features of QCD, color confinement. Confinement is the phenomenon that gluons, or quarks, cannot be observed isolated. Lattice gauge theory is an approach which allows us to study properties of QCD in a non-perturbative manner and is thus one possible technique to explore the low energy limit of the phase diagram, see Fig. 2.4. The discretization of spacetime allows for an efficient computation of observables by Monte-Carlo simulations. The confinement-deconfinement phase transition can be quantified by the Polyakov loop order parameter, which is defined as the average of timelike loops on lattice configurations. It is zero in the confining phase and finite in the deconfining phase.

This parameter plays a crucial role in our investigation of the phase diagram of SU(2) lattice gauge theory. As it turns out, the Polyakov loop is the dominant quantity, that an artificial neural networks learns automatically, when tasked to find the phase diagram of lattice gauge theory. It is worth noting that the Polyakov loop is an order parameter which is non-linear and non-local and thus an excellent demonstration for the power of our neural networks. For more details on how to explore the confinement-deconfinement transition with an artificial neural network, see *Machine Learning of Explicit Order Parameters: From the Ising Model to SU(2) Lattice Gauge Theory* [3].

3 Physics and the Choice of Regulators in Functional Renormalisation Group Flows

Pawlowski, Jan M.; Scherer, Michael M.; Schmidt, Richard; Wetzel, Sebastian J.
Annals of Physics,Volume 384, September 2017, Pages 165-197 [1]

Annals of Physics 384 (2017) 165–197

Contents lists available at ScienceDirect

Annals of Physics

journal homepage: www.elsevier.com/locate/aop

Physics and the choice of regulators in functional renormalisation group flows

Jan M. Pawlowski [a,b], Michael M. Scherer [a], Richard Schmidt [c,d], Sebastian J. Wetzel [a,*]

[a] Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany
[b] ExtreMe Matter Institute EMMI, GSI Helmholtzzentrum für Schwerionenforschung mbH, Planckstr. 1, D-64291 Darmstadt, Germany
[c] ITAMP, Harvard–Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[d] Physics Department, Harvard University, 17 Oxford Street, Cambridge, MA 02138, USA

HIGHLIGHTS

- We show the regulator dependence of functional renormalisation (FRG) group equations.
- Solving FRG equations along closed loops provides a distance between two regulators.
- We suggest a practical procedure for devising optimal regulators.
- A universality class crossover can be induced by changing relative cutoff scales.
- We examine the regulator dependence explicitly in the Fermi polaron problem.

ARTICLE INFO

Article history:
Received 30 January 2017
Accepted 27 June 2017
Available online 16 July 2017

Keywords:
Renormalisation group
Functional renormalisation group

ABSTRACT

The Renormalisation Group is a versatile tool for the study of many systems where scale-dependent behaviour is important. Its functional formulation can be cast into the form of an exact flow equation for the scale-dependent effective action in the presence of an infrared regularisation. The functional RG flow for the scale-dependent effective action depends explicitly on the choice of regulator, while the physics does not. In this work, we systematically investigate three key aspects of how the regulator choice affects RG flows: (i) We study flow trajectories along closed loops in the space of action functionals varying both, the regulator scale and shape function. Such a flow does not vanish in the presence of truncations. Based on a definition of the length of an RG trajectory, we suggest a constructive procedure for devising optimised

* Corresponding author.
E-mail address: wetzel_s@thphys.uni-heidelberg.de (S.J. Wetzel).

http://dx.doi.org/10.1016/j.aop.2017.06.017

regularisation schemes within a truncation. (ii) In systems with various field variables, a choice of relative cutoff scales is required. At the example of relativistic bosonic two-field models, we study the impact of this choice as well as its truncation dependence. We show that a crossover between different universality classes can be induced and conclude that the relative cutoff scale has to be chosen carefully for a reliable description of a physical system. (iii) Non-relativistic continuum models of coupled fermionic and bosonic fields exhibit also dependencies on relative cutoff scales and regulator shapes. At the example of the Fermi polaron problem in three spatial dimensions, we illustrate such dependencies and show how they can be interpreted in physical terms.

1. Introduction

In the past twenty years, the functional renormalisation group (FRG) approach [1] has been established as a versatile method allowing to describe many aspects of different physical systems in the framework of quantum field theory and statistical physics. Applications range from quantum dots and wires, statistical models, condensed matter systems in solid state physics and cold atoms over quantum chromodynamics to the standard model of particle physics and even quantum gravity. For reviews on the various aspects of the functional RG see [2–19].

The functional renormalisation group approach can be set-up in terms of an exact flow equation for the effective action of the given theory or model [1]. The choice of the initial condition at some large ultraviolet cutoff scale, typically a high momentum or energy scale, together with that of the regulator function determines both, the physics situation under investigation as well as the regularisation scheme. The functional RG flow for the scale-dependent effective action depends explicitly on the choice of regulator, while the physics does not. The latter is extracted from the full quantum effective action at vanishing cutoff scale implying a vanishing regulator. Hence, at this point no dependence on the choice of regulator is left, only the implicit choice of the regularisation scheme remains.

Typically, for the solution of the functional flow equation for the effective action one has to resort to approximations to the effective action as well as to the flow. Such a truncation of the full flow usually destroys the regulator independence of the full quantum effective action at vanishing cutoff. Therefore, devising suitable expansion schemes and regulators is essential for reliable results. In the context of perturbation theory and beyond, the principle of minimal sensitivity (PMS) [20] suggests to use extrema in the regulator dependence of specific observables as the most accurate result, see, e.g., [21–24]. Here, we strive for a constructive approach for finding an optimised regulator. Moreover, related considerations also allow for a discussion of the systematic error within a given truncation scheme. This has been examined in detail for the computation of critical exponents in models with a single scalar field in three dimensions within the lowest order of the local potential approximation (LPA), [7,25–33]: an optimisation procedure, [7,23,27] suggests a particular regulator choice – the flat regulator – which is also shown to yield the best results for the critical exponents.

The optimisation framework in [7] has been extended to general expansion schemes in a functional optimisation procedure including fully momentum-dependent approximation schemes. An application to momentum-dependent correlation functions in Yang–Mills theory and ultracold atoms can be found in [34–36].

Still, for more elaborate truncations, in particular higher orders of the derivative expansion in the LPA, including, e.g., momentum dependencies or higher-order derivative terms, little has been done when it comes to a practical implementation of constructive optimisation criteria. Also, more complex physical models with different symmetries such as, e.g., non-relativistic systems, or models with several different fields, for example mixed boson–fermion systems, demand for a thorough study of their regulator dependence in order to extract the best physical results from a given truncation.

In this paper we study the impact of different regulator choices on truncated functional renormalisation group flows in various models and further develop the functional optimisation procedure set-up in [7]. In its present form it allows for a practical and simple comparison of the quality of different regulators and for the construction of an optimised one.

The rest of the paper is organised as follows: In Section 2 we shortly introduce the FRG approach and explain how the choice of a specific regulator influences a truncated FRG flow. This is captured in terms of an integrability condition for closed loops in theory space upon a change of the regulator and RG scale, cf. Section 2.4. In Section 3 we then devise a road towards a practical constructive optimisation procedure. We discuss the length of an RG trajectory which has to be minimal for an optimised regulator, cf. Section 3.2. This procedure is then applied to a simply scalar model (Section 3.4). A more heuristic approach to models with various degrees of freedom such as two-scalar models and non-relativistic boson–fermion systems, is presented in Sections 4 and 5, respectively. To this end, we introduce a shift between the regulator scales of the different fields and show how this affects the results allowing for a change of the underlying physics upon varying the regulator. This, again, clearly demands for carefully choosing a regularisation scheme which could be performed by a optimisation procedure as suggested in this work.

2. Functional RG flows

The functional renormalisation group is based on the Wilsonian idea of integrating out degrees of freedom. In the continuum, this idea can be implemented by suppressing the fluctuations in the theory below an infrared cutoff scale k. An infinitesimal change of k is then described in terms of a differential equation for the generating functional of the theory at hand — Wetterich's flow equation [1]. The infrared suppression can be achieved by adding a momentum-dependent mass term to the classical action,

$$S[\varphi] \to S[\varphi] + \frac{1}{2}\int_p \varphi(p) R_k(p) \varphi(-p), \tag{1}$$

with $\int_p = \int \frac{d^d p}{(2\pi)^d}$. The regulator $R_k(p)$ tends towards a mass term for low momenta and vanishes sufficiently fast in the ultraviolet, see (3) and (4).

With the cutoff scale dependent action (1) also the one-particle-irreducible (1PI) effective action or free energy, $\Gamma_k[\phi]$, acquires a scale dependence. The nth field derivatives of the effective action, $\Gamma_k^{(n)}[\phi]$, are the 1PI parts of the n-point correlation functions in a general background $\phi = \langle\varphi\rangle$. The flow of Γ_k is given by

$$\partial_t \Gamma_k = \frac{1}{2}\mathrm{Tr}\, G_k[\phi]\, \partial_t R_k \text{ with } G_k[\phi] = \frac{1}{\Gamma_k^{(2)}[\phi] + R_k}, \tag{2}$$

where we have introduced the renormalisation time $t = \ln k/\Lambda$. Here, Λ is some reference scale, usually the ultraviolet scale, where the flow is initiated. The trace sums over all occurring indices, including the loop integration over momenta. The regulator is conveniently written as

$$R_k(p) = Z_k p^2\, r(p^2/k^2), \tag{3}$$

with the dimensionless shape function $r(y)$ that only depends on the dimensionless ratio $y = p^2/k^2$. The regulator functions fulfil the infrared and ultraviolet conditions

$$\lim_{y\to 0} y\, r(y) > 0\,, \qquad \lim_{y\to\infty,\epsilon>0} y^{d/2+\epsilon}\, r(y) = 0. \tag{4}$$

The first limit in (4) implements the infrared suppression of low momentum modes as the propagator G_k acquires an additional infrared mass due to R_k. The second limit guarantees that the ultraviolet is unchanged. The regulator $R_k(p)$ has to decay with higher powers as p^d in d dimensions in order to have a well-defined flow equation without the need of an ultraviolet renormalisation. Approximating the effective action, for example using a derivative expansion, can add additional constraints to the regulator. With (4) the flow equation (2) is ultraviolet finite due to the sufficiently fast decay of the

regulator in the ultraviolet. Here, we presented the relativistic case for simplicity. The arguments can be extended to the non-relativistic case, e.g. [37]. We discuss one specific example for such a non-relativistic system in Section 5.

2.1. Ultraviolet limit and regulator dependence

In the limit $k \to \infty$, the cutoff term in (1) suppresses all momentum fluctuations. To discuss this limit, we consider the RG running of the scale-k-dependent couplings $g_n(k)$ parametrising the theory in terms of a suitable basis of field monomials, i.e., $\Gamma_k = \sum_n g_n(k)\,\mathcal{O}_n(\partial,\phi)$. We classify the $g_n(k)$ according to their UV scaling dimension d_n that follows from the running of the couplings towards the ultraviolet (UV) with the flow equation (2). The UV scaling dimension d_n is the full quantum dimension, i.e., canonical plus anomalous dimension,

$$g_n(k) \sim k^{d_n}. \tag{5}$$

Terms in the effective action Γ_k whose couplings $g_n(k)$ have semi-positive UV scaling dimension, $d_n \geq 0$, dominate the UV behaviour. In turn, terms with couplings $g_n(k)$ with $d_n < 0$ are sub-leading or suppressed.

Let us demonstrate this at the example of the relativistic φ^4 field theory in $d = 3$ dimensions. This theory is super-renormalisable and the only parameter with a positive UV scaling dimension is the mass parameter $m_k^2 = \Gamma_k^{(2)}(p = 0, \phi = 0)$. The flow of the mass is derived from (2) with a second order field derivative evaluated at vanishing fields and momenta, to wit

$$\partial_t m_k^2 = -\frac{1}{2}\int_q G_k(q)\,\partial_t R_k(q)\,G_k(q)\,\Gamma_k^{(4)}(q,q,0,0), \tag{6}$$

where we have used that the three-point function vanishes due to the symmetry of the theory under $\phi \to -\phi$, i.e., $\Gamma_k^{(3)}[\phi = 0] = 0$. In the UV limit the flow (6) simplifies considerably. The four-point function tends towards a local scale-independent vertex, i.e., $\Gamma^{(4)}_{k\to\infty}(p_1,\ldots,p_4) \to \lambda_{\rm UV}$. For $k \to \infty$, the propagators in (6) are simply given by

$$G_k(q) = \frac{1}{q^2[1 + r(q^2/k^2)] + m_k^2}. \tag{7}$$

Here, we have also used that the wave function renormalisation $Z_k \to 1 + O(1/k)$ in the limit $k \to \infty$. This can be proven analogously to the following determination of the asymptotic scaling of the mass. For the mass we are hence led to the asymptotic UV flow

$$\partial_t m_k^2 = -k\,\lambda_{\rm UV}\int_{\bar q}\frac{\bar q^4\,\partial_{\bar q^2} r(\bar q^2)}{(\bar q^2[1 + r(\bar q^2)] + \bar m_k^2)^2}, \tag{8}$$

where quantities denoted with a bar are scaled with appropriated powers of k in order to make them dimensionless, i.e., $\bar q^2 = q^2/k^2$ and $\bar m_k^2 = m_k^2/k^2$. The flow (8) is further simplified if we reduce it to the leading UV scaling. To that end, we notice that the flow scales with k for $\bar m_k = 0$. Hence, for $k \to \infty$ we have $m_k^2 \propto \lambda_{\rm UV}\,k$ and $\bar m_k^2 \propto \lambda_{\rm UV}/k \to 0$. Accordingly, we have

$$m_k^2 = \mu(r)\lambda_{\rm UV}\,k + O(k^0),$$

with the dimensionless factor

$$\mu(r) = -\int_{\bar q}\frac{\partial_{\bar q^2} r(\bar q^2)}{[1 + r(\bar q^2)]^2}. \tag{9}$$

We conclude that the mass parameter m_k^2 diverges linearly with k. We also note that the constant $\mu(r)$ is non-universal and depends on the chosen regulator r. Additionally, the above simple example nicely reflects the regularisation and renormalisation scheme dependence in the present modern functional RG setting: UV divergences in standard perturbation theory are reflected in UV relevant terms such as $\mu(r)\lambda_{\rm UV}\,k$, that diverge for $k \to \infty$. The subtractions or renormalisation in perturbation theory

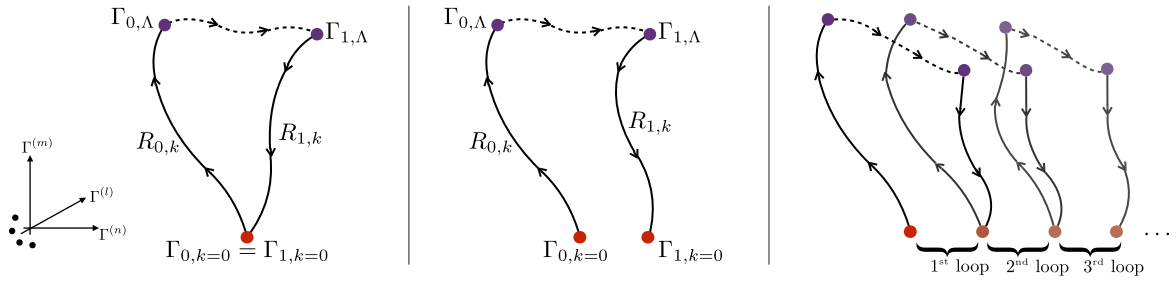

Fig. 1. Left panel: Pictorial representation of the integrability condition (12) in the theory space of action functionals. By means of (11), we can map the two actions at the initial scale Λ onto each other. Integrating out momentum fluctuations from $k = \Lambda$ towards $k = 0$ then yields the full quantum effective action for both regulators, i.e., $\Gamma_{0,k=0} = \Gamma_{1,k=0}$. Middle panel: Approximations lead to a violation of the integrability condition, cf. (15). Right panel: Pictorial representation of several successively performed loops according to (15) within an approximation for the full propagator G_k. This sketch exhibits how the result for the full effective action $\Gamma_{k=0}$ moves further and further away from its initial position in theory space.

are reflected in the consistent choice of the initial condition that makes the full effective action $\Gamma_{k=0}$ independent of the initial scale $k = \Lambda$. Accordingly, the initial mass m^2_Λ has to satisfy the flow equation (6) which again leads to (9) for m^2_Λ. In other words, the Λ-dependence of the flow is annihilated by that of the initial conditions. This accounts for a BPHZ-type renormalisation, for detailed discussions see e.g. [7,13,38] and references therein. Consequently, a part of the standard renormalisation scheme dependence is carried by the regulator dependence of $\mu(r)$.

Finally, the physics is entirely carried by the finite part of the UV limit, that is the $O(k^0)$ term in (9). Since this finite UV part has first to be mapped to $k \to 0$ via the flow, it also carries a renormalisation scheme dependence. In summary, the latter is given by a combination of the shape dependence and the finite part of the initial condition. This simple distinction can be used to rewrite the effective action in terms of renormalised fields and parameters for obtaining a finite UV limit, see e.g. [7,13].

2.2. Initial actions and integrability condition

The above discussion already highlights the regulator- or r-dependence of the flow and the scale-dependent effective action. However, despite this r-dependence of the flow, the final effective action $\Gamma_{k=0}$, when expanded in the basis of n-point functions, is unique up to RG transformations, see Appendix A for a discussion of this issue. This is illustrated in theory space in the left panel of Fig. 1. The initial effective actions at the UV scale Λ differ due to the different shape functions r. Nonetheless, we can map the initial effective actions onto each other by the following flow equation

$$\partial_s \Gamma_{s,\Lambda} = \frac{1}{2} \mathrm{Tr}\, G_\Lambda[\phi]\, \partial_s R_{s,\Lambda}, \tag{10}$$

with $R_{s,k} = p^2 r_s$ and where we use a one parameter family of shape functions $r_s(y)$ with $r_0(y) = r^{\mathrm{A}}(y)$ and $r_1(y) = r^{\mathrm{B}}(y)$ which analytically transforms r^{A} into r^{B}. Then, (10) is easily derived similarly to the cutoff flow (2). Consequently, the two initial actions are mapped onto each other by the s-integration of (10),

$$\Gamma_{1,\Lambda} = \Gamma_{0,\Lambda} + \frac{1}{2} \int_0^1 ds\, \mathrm{Tr}\, G_\Lambda[\phi]\, \partial_s R_{s,\Lambda}, \tag{11}$$

and the full quantum effective actions agree trivially, $\Gamma_{0,k=0} = \Gamma_{1,k=0}$, see also left panel of Fig. 1. This statement can be reformulated as an integrability condition

$$\int_0^1 \partial_s \Gamma_{s,\Lambda} + \int_\Lambda^0 \frac{dk}{k} \partial_t \Gamma_{1,k} + \int_0^\Lambda \frac{dk}{k} \partial_t \Gamma_{0,k} = 0, \tag{12}$$

which defines a closed loop in theory space.

2.3. Integrability condition and approximations

In general, the integrability condition (12) is violated in approximations to the full effective action, and the flow ceases to describe a total derivative with respect to t and s. The relation between an approximation of the effective action or rather of the propagator in the flow equation and the derivative operators ∂_t and ∂_s is more clearly seen in the general flow equation for a complete set of composite operators $\{I_k\}$ derived in [7],

$$\partial_t I_k[\phi] = \left(-\frac{1}{2} \mathrm{Tr}\, G_k[\phi]\, \partial_t R_k\, G_k[\phi] \frac{\delta^2}{\delta \phi^2} \right) I_k. \tag{13}$$

The set of general n-point correlation functions $\langle \phi(x_1) \cdots \phi(x_n) \rangle$ with connected and disconnected parts forms such a complete set, and is part of the set $\{I_k\}$ that satisfies (13). For a concise recent derivation of (13) see [39]. A further relevant example is $\delta \Gamma_k / \delta \phi$. For $\delta \Gamma_k / \delta \phi$ one easily sees that the ϕ-derivative of the flow (2) gives (13) with $I_k = \Gamma_k^{(1)}$. Note, however, that the effective action Γ_k does not satisfy (13). Another simple test is given by the full two-point function $G_k(p, q) + \phi(p)\phi(q)$. Importantly, as the set of composite operators that satisfy (13) includes all correlation functions it is complete. We conclude that the total t-derivative has the representation

$$\partial_t = \left(-\frac{1}{2} \mathrm{Tr}\, G_k[\phi]\, \partial_t R_k\, G_k[\phi] \frac{\delta^2}{\delta \phi^2} \right), \tag{14}$$

on the – complete – set of composite operators $\{I_k\}$. Eq. (14) makes explicit the consequences of approximations to the effective action for the total t-derivative: they imply approximations for the full propagators G_k and hence approximations to the representation (14) of the total t-derivative ∂_t. Consequently, an integration along a closed loop in regulator space does not necessarily vanish within an approximation to the effective action. Note that the notation ∂_t for the total t-derivative is common in the FRG community and we keep it for the sake of comparability.

For our discussion of flows that change regulators as well as cutoff scales we extend the notation with the parameter s to general one-parameter flows in theory space. Such a flow includes changes of the cutoff scale k with $k(s)$, changes of the shape of the regulator r_s and reparametrisations of the theory, see Appendix B for a detailed discussion. Within this unified approach a closed loop such as the global one in (12) has the simple representation

$$\oint_{\mathcal{C}} d\, \Gamma[\phi, R] = \int_0^1 ds\, \frac{d}{ds} \Gamma_s[\phi, R_s] = 0. \tag{15}$$

here, s parameterises the closed curve $\mathcal{C}$, and R_s describes a loop in regulator space with $R_0 = R_1$. In general, approximations to the effective action Γ_k lead to

$$\int_0^1 ds\, \frac{d}{ds} \Gamma_s[\phi, R_s] \neq 0, \tag{16}$$

for closed loops, see Fig. 2 for a pictorial representation. This also means that if such a loop is taken many periods eventually the result will be dominated completely by the errors introduced by the approximation of the representation to the total t-derivative. In particular, the global loop shown in the left panel of Fig. 1 does not close.

A violation (16) of the integrability condition (15) is a measure for the self-consistency of the approximation at hand. In the following we will use it in our quest for optimal regulators as well as a systematic error estimate. In Appendix C we discuss under which circumstances (15) is violated and when it is satisfied. A systematic error estimate of a given approximation can be set up as follows: Consider general closed loops in theory space initiated from a given regulator R_k^{A}. Then, we change the regulator at a fixed initial scale as in (10), and subsequently flow to vanishing cutoff scale. For sensible regulator choices, the spreading of the results for $\Gamma[\phi, R = 0]$ provides an error estimate, see Fig. 3 for a pictorial representation. This will be explicitly explored in the following Section 2.4.

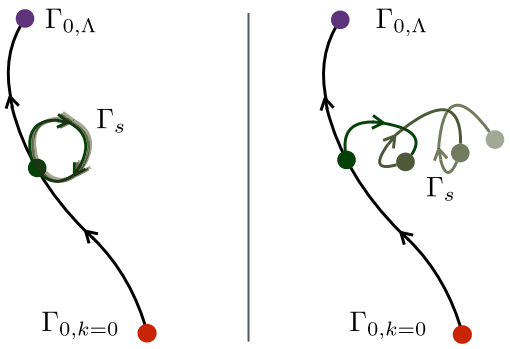

Fig. 2. Generalised one-parameter flows. Left panel: The exact flow equation provides the integrability condition, (15). Right panel: In general, the integrability condition is violated when approximations to G_k are employed, cf. (16).

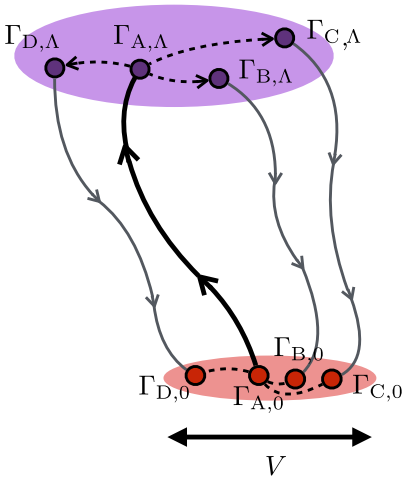

Fig. 3. A change of regulator at the UV scale Λ, Eq. (11), maps the action $\Gamma_{\mathrm{A},\Lambda}$ onto one of the others, $\Gamma_{\mathrm{B},\Lambda}$, $\Gamma_{\mathrm{C},\Lambda}$, $\Gamma_{\mathrm{D},\Lambda}$, corresponding to different regulators. Then, integrating the RG flow in an approximation towards $k = 0$ results in different approximations to the IR quantum effective action, i.e., $\Gamma_{\mathrm{i},0}$ with $i \in \{\mathrm{A,B,C,D}\}$, cf. Eq. (15). The spreading of these results from a large class of general regulators can be used as an error estimate for an approximation.

2.4. Scalar model

Generally, it is not possible to exactly solve the flow equation (2) for the flowing action Γ_k. Therefore, we have to devise suitable truncation schemes for the functional Γ_k. A simple scheme is given by the derivative expansion which assumes a small impact of momentum fluctuations on the correlation functions. In this section, we will investigate a three-dimensional $O(1)$ (also known as Z_2) symmetric scalar model to explore the effects of regulator choices on functional RG results. Our ansatz is given by the local potential approximation (LPA)

$$\Gamma_k = \int_x \Big[\frac{1}{2}(\partial_\mu \phi)^2 + U_k(\phi)\Big], \quad U_k = \sum_i \frac{\bar{\lambda}_i}{i!}(\bar{\rho} - \bar{\kappa})^i\,, \tag{17}$$

with the real scalar field ϕ, the scale-dependent effective potential U_k and the field invariant $\bar{\rho} = \frac{1}{2}\phi^2$, $\bar{\kappa}$ denotes the minimum of the effective potential.

The parameters of the effective potential are scale dependent quantities $\bar{\lambda}_n = \bar{\lambda}_n(k)$ and $\bar{\kappa} = \bar{\kappa}(k)$, however for brevity, we will not indicate this in the following. Further, we have set the wave function renormalisation to unity, dropping all non-trivial momentum-dependences. For calculations, we introduce the dimensionless effective potential and couplings $u(\rho) = U_k k^{-d}$, $\kappa = k^{2-d}\bar{\kappa}$, $\rho = k^{2-d}\bar{\rho}$ and $\lambda_i = \bar{\lambda}_i k^{-d+i(d-2)}$. Then, we can write the flow equation for the effective potential as

$$\partial_t u = -d\,u + (d-2)\rho\,u' + I(u' + 2\rho\,u''), \tag{18}$$

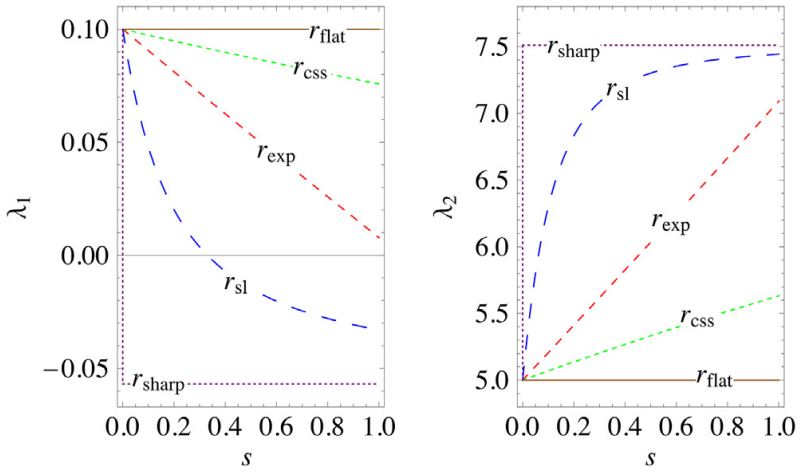

Fig. 4. Change of the initial conditions for different choices of linear interpolations from $r_{\rm flat}$ to r_i. The parameter-dependent regulator $r_{\rm css}$ was chosen with $c = 0.5$, $b = 1$, $h = 0.5$. The shape function $r_{\rm sl}$ with $c = 10$.

where primes denote derivatives with respect to ρ and we have defined the threshold function

$$I(w) = v_d \int_0^\infty y^{\frac{d}{2}+1} dy \, \frac{-2\,\partial_y r(y)}{y(1+r(y))+w}, \tag{19}$$

with $y = p^2/k^2$ and $v_d^{-1} = 2^{d+1}\pi^{\frac{d}{2}}\Gamma(d/2)$. Using the series expansion of the effective potential, (17), also for its dimensionless version, we extract the flow equations for the individual couplings by projections, see Appendix E for explicit expressions.

2.4.1. Switching regulators at fixed RG scale

As discussed in Section 2.2 a change in the regulator triggers a flow in the space of action functionals. In particular this means that switching from one regulator to another induces a change in the initial conditions as exhibited in (11). To visualise this change explicitly, we employ superpositions of two regulators at a fixed scale k, with $s \in [0, 1]$,

$$r_s(y) = s\, r^{\rm A}(y) + (1-s)\, r^{\rm B}(y)\,. \tag{20}$$

The flow equation with respect to the variable s is then,

$$\begin{aligned} \partial_s u &= J(u' + 2\rho\, u'')\,, \\ J(w) &= v_d \int_0^\infty y^{\frac{d}{2}} dy \frac{\partial_s r_s(y)}{y(1+r_s(y))+w}\,, \end{aligned} \tag{21}$$

where $\partial_s r_s(y) = r^{\rm A}(y) - r^{\rm B}(y)$. More generally, we do not require a linear superposition as specified in (20), but we can switch regulators on an arbitrary smooth trajectory while keeping the scale k fixed. The change of initial conditions from switching between different regulators is then given by the solution of the flow equation (21).

In Fig. 4, we show this solution for a collection of representative regulator shape functions listed in Table 1 which we switch by a linear interpolation. Here, we have integrated flow equations within an LPA expanded to fourth order in ϕ in the symmetric regime, concentrating on the four-scalar coupling and the only relevant coupling, the mass parameter. In Fig. 4 we follow the regulator-dependence of the two dimensionless couplings $\lambda_1 = m^2$ and λ_2, starting at $s = 0$ with the flat regulator $r_{\rm flat}$ and the initial conditions $\lambda_1^{(\rm in)} = 0.1$ and $\lambda_2^{(\rm in)} = 5.0$. Fig. 4 clearly exhibits the change in the initial conditions upon variations of the regulator shape function at fixed RG scale k. Interestingly, the largest difference in initial conditions starting from $r_{\rm flat}$ is given by switching to the sharp regulator $r_{\rm sharp}$.

Table 1
Step-like, compactly supported smooth [40,41], and exponential interpolating classes of regulator shape functions r, and respective parameter choice that provide exponential, flat and sharp regulators in the respective class.

Regulator type	Representation	Limits
Exponential	$r_{\rm exp}(y) = (\exp(y) - 1)^{-1}$	–
Flat (Litim)	$r_{\rm flat}(y) = (\frac{1}{y} - 1)\theta(1-y)$	–
Step-like	$r_{\rm sl}(y) = \frac{c}{y}\theta(1-y)$	$\lim_{c\to\infty} r_{\rm sl}(y) \to r_{\rm sharp}(y)$
Compactly supported smooth	$r_{\rm css}(y) = \frac{c\,\theta(1-hy^b)}{\exp(\frac{cy^b}{1-hy^b})-1}$	(i) $\lim_{b\to\infty} r_{\rm css}(y) \to r_{\rm sharp}(y)$, for $c>0, h=0$ (ii) $\lim_{c\to 0} r_{\rm css}(y) \to r_{\rm flat}(y)$, for $b=1, h=1$ (iii) $r_{\rm css}(y) = r_{\rm exp}(y)$, for $b=c=1, h=0$
Exponential interpolating	$r_{\rm int}(y) = \frac{(a-by)y^{n-1}}{\exp(y^n)-1}$	(i) $\lim_{a,n\to\infty} r_{\rm int}(y) \to r_{\rm sharp}(y)$, for $b=0$ (ii) $\lim_{n\to\infty} r_{\rm int}(y) \to r_{\rm flat}(y)$, for $a=1, b=1$ (iii) $r_{\rm int}(y) = r_{\rm exp}(y)$, for $a=1, b=0, n=1$

2.4.2. Loops in $k - R_k$ space

In addition to the change of the regulator shape from $R_k^{\rm A}$ to $R_k^{\rm B}$ at a fixed RG scale, we now allow for a dependence of the RG scale k on the loop variable s, i.e., $k \to k(s)$. Then, we can perform integrations along closed loops in theory space, cf. Fig. 2, and study the violation of the integrability condition, (15), explicitly. Such a combined change of regulator and RG scale can be incorporated in a linear interpolation between two scale dependent regulators

$$R_{s,k(s)} = a(s) R^{\rm A}_{k(s)} + \big(1 - a(s)\big) R^{\rm B}_{k(s)}, \tag{22}$$

where $a(s) \in [0,1]$ parametrises the switching from one regulator function to another. In order to solve the flow equations along a loop in $k - R_k$ space we add to (21) the terms which include solving the flow equations in k direction,

$$\frac{d}{ds} u = J(u' + 2\rho\, u'') + \frac{\partial_s k(s)}{k(s)} \big(-d\, u + (d-2)\rho\, u'\big)\,. \tag{23}$$

The s-derivative of the regulator in the threshold function $J(\omega)$ defined in (21) can be decomposed into two contributions

$$\frac{d}{ds} r_s(y) = \underbrace{\partial_s r_s(y)\big|}_{\text{change of shape}} - \underbrace{2\,\frac{\partial_s k(s)}{k(s)}\, y\, \partial_y r_s(y)}_{\text{explicit change of scale}}, \tag{24}$$

where the first term keeps track of the change of the shape of the regulator function $r_s(y)$, while the second term tracks the change of the cutoff scale $k(s)$. Evidently this is just a convenient splitting as the change of $k(s)$ can also be easily described by a change of r_s. This is seemingly a trivial remark but it hints at the fact that a change of the shape of the regulator may very well imply a change of the physical cutoff scale. This discussion will be detailed further in Section 3. Explicitly, the involved derivatives are given by

$$\begin{aligned} \partial_s r_s(y) &= \big(r^{\rm A}(y) - r^{\rm B}(y)\big)\partial_s a(s)\,, \\ \partial_y r_s(y) &= a(s)\partial_y r^{\rm A}(y) + (1-a(s))\partial_y r^{\rm B}(y)\,. \end{aligned} \tag{25}$$

In the following, we again employ a linear superposition between two regulators, e.g., $r_{\rm exp}$ and $r_{\rm sharp}$, and solve the flow of the scalar model in LPA to order ϕ^4 along a closed loop. Each closed loop in $k - R_k$ space along a rectangle contour then consists of four steps, see Fig. 5 for a representative contour:

1. The flow equation is solved from k_1 to k_2.
2. Switch the regulator continuously from $r^{\rm A}$ to $r^{\rm B}$.
3. Reverse the flow from k_2 to k_1.
4. Switch back from regulator $r^{\rm B}$ to $r^{\rm A}$.

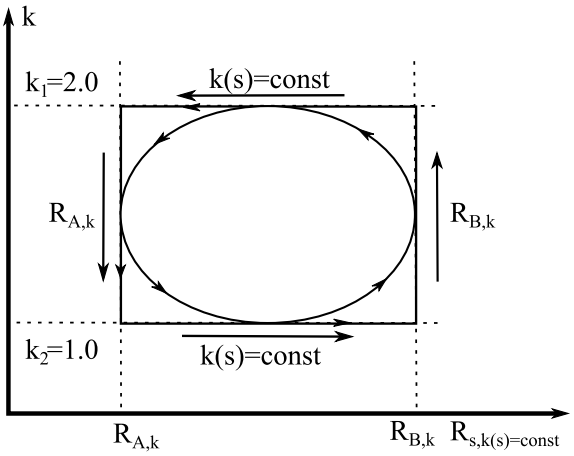

Fig. 5. Representative contours of a closed loop in $k - R_k$ space including changes of regulator shape as well as RG scale. The rectangle contour separates solving the flow equations and switching the regulators. The ellipse contour however, allows for simultaneous changes in the scale and the regulator shape.

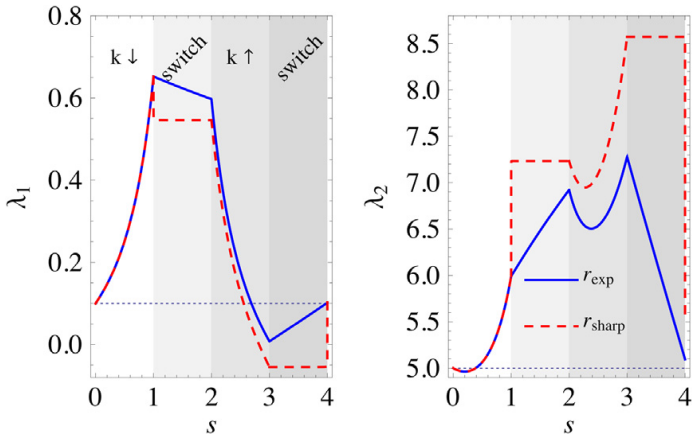

Fig. 6. Flow of the couplings $\lambda_{1,2}$ along one closed-loop rectangle shaped contour as defined in Fig. 5. The red dashed curve shows the switching from r_{flat} to r_{sharp} and the blue solid curve the switch from r_{flat} to r_{exp}. At the end of the closed-loop integration the values of the couplings $\lambda_{1,2}$ deviate slightly from their initial values.

The flow of the couplings λ_1 and λ_2 along one such closed loop is depicted in Fig. 6. For these calculations, we again use the initial values $\lambda_1^{(\text{in})} = 0.1$ and $\lambda_2^{(\text{in})} = 5.0$. We switch from r_{flat} to r_{sharp} (red dashed) or to r_{exp} (blue solid), respectively. The change from r_{flat} to r_{exp} is smooth as both regulators are finite for all momenta. In contrast, the change from r_{flat} to r_{sharp} shows a discontinuous peak in the flow of $\lambda_{1,2}$: the transition from the flat to the sharp regulator instantly lends an infinite infrared mass to the propagator for momenta lower than the cutoff scale of the sharp regulator. In either case the integration along one of our chosen closed-loop contours shows slight deviations from the initial values of λ_1 and λ_2, see Fig. 6.

The deviations from the initial values add up when the procedure of integrating along a closed-loop contour in $k - R_k$ space is repeated. This is shown in the upper panel of Fig. 7 for a consecutive integration along four of the closed loops as defined in Fig. 5. In fact, after these four closed loops the values of the coupling constants $\lambda_{1,2}$ strongly deviate from their initial values. For comparison, in Fig. 7, we also show an integration along an alternative closed-loop, defined by an ellipse contour, cf. Fig. 5. This integration can be performed in a completely analytical way for a transition from r_{flat} to r_{sharp} as shown in Appendix F.

Our study clearly demonstrates the violation of the integrability condition, (15), for truncated renormalisation group flows. The severeness of this violation depends on the chosen regulators,

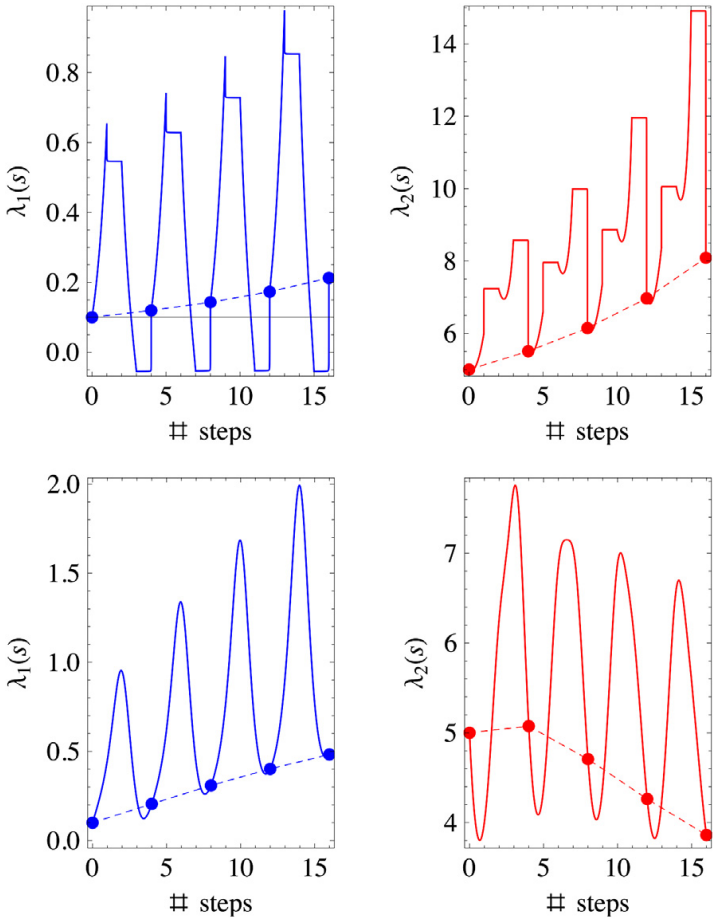

Fig. 7. Upper panels: Flow of the couplings $\lambda_{1,2}$ along four rectangular closed loops switching from $r_{\rm flat}$ to $r_{\rm sharp}$. Lower panels: Flow in $k - R_k$ space with $r_{\rm flat}$ and $r_{\rm sharp}$ along an ellipse contour (counter clockwise), with an interpolation starting with scale $k^{(\rm in)} = 2.0$ and $r = r_{\rm flat}(y) + r_{\rm sharp}(y/2)$, cf. Appendix F.

cf. Fig. 6 and indicates the necessity of an educated choice of the regularisation scheme in renormalisation group investigations to establish and improve the reliability of physical results. The following section is dedicated to devising such an educated choice in terms of an optimisation procedure.

3. Optimisation

In order to obtain the best possible results from the functional renormalisation group approach within a given truncation we would like to single out the optimal regulator scheme for the underlying systematic expansion. Here we follow the setup of functional optimisation put forward in [7]. The discussion of systematic error estimates related to optimisation requires a norm on the space of theories (at $k = 0$) in order to measure the severeness of the deviations. Here, we are not after a formal definition but rather a practical choice of such a norm.

We illustrate complications with the definition of such a norm by means of a simple example: we restrict ourselves to the local potential approximation (LPA), or LPA′ where in the latter we take into account constant wave function renormalisations Z_k. Then, a seemingly natural choice is the cartesian norm on theory space spanned by the constant vertices $\lambda_n = \Gamma^{(n)}[\phi_{\rm EoM}]$ evaluated, e.g., at the equation of motion $\phi_{\rm EoM}$. However, this falls short of the task as it weights a deviation in higher correlations or vertices λ_n in the same way as that of the lower ones, despite the fact that the lower ones are typically more important. Additionally, the $\Gamma^{(n)}$ are neither renormalisation group invariant nor do they scale identically, see (A.2).

If we extend the above setting to a general vertex expansion scheme, the coordinates in theory space are related to $\Gamma^{(n)}[\phi_{\rm EoM}](p_1, \ldots, p_n)$. These quantities are operators and the definition of the

related nth axis of the coordinates system requires a suitably chosen operator norm, for a more detailed discussion see [7]. Even though this general case can be set-up, for most practical purposes it is sufficient to rely on a simple definition of a norm adapted to the approximation at hand.

Let us assume that we found a norm that allows to define the length $L[\mathcal{C}]$ of a given flow along a trajectory $\mathcal{C}$ in theory space parametrised with $s \in [0, 1]$, flowing from some regulator $R_{s=0}$ to $R_{s=1}$. For example, we can consider the global flow with a given regulator from $k = \Lambda$ to $k = 0$, i.e., the flow trajectory does not necessarily have to be a closed loop. The discussions in the previous section suggest that, in a given approximation, a deviation from the full effective action gets amplified during the renormalisation flow. We thus accumulate an error while solving the flow equation. In turn we should try to minimise the flow length in order to minimise the systematic error. Accordingly, we have to compare the lengths of different trajectories $L[\mathcal{C}]$. This heuristic argument can be made more precise, [7]: without approximation the final effective action $\Gamma[\phi, R = 0]$ does not depend on the trajectory, in other words

$$\frac{\delta}{\delta R_s(p)} \int_0^1 ds\, \partial_s \Gamma[\phi, R] = 0. \tag{26}$$

Note however, that this discussion bears an intricacy, as it implies the comparison of the length of flow trajectories of physically equivalent effective actions $\Gamma[\phi, R^{\rm A}]$ and $\Gamma[\phi, R^{\rm B}]$ towards $\Gamma[\phi, 0]$. Therefore, we should compare trajectories that always start at physically equivalent effective actions at a large physical cutoff scale.

3.1. Physical cutoff scale

The cutoff parameter k is usually identified with the physical cutoff scale, but such an identification falls short in the general case. To understand this, let us re-evaluate the example of the flows with $r^{\rm A}$ and $r^{\rm B}$ leading to the circular flow (12). In the spirit of the discussion above it seems to be natural to compare the two flows from $k = \Lambda$ to $k = 0$ with the regulators $r_{s=0} = r^{\rm A}$ and $r_{s=1} = r^{\rm B}$, respectively, while the s-flow in this example simply switches the regulator at a fixed scale $k = \Lambda$. This picture fails trivially for

$$r_1(x) = r_0(x/\lambda)\,, \quad \text{with} \quad R_{1,k} = R_{0,\lambda k}, \tag{27}$$

where the change of regulators simply amounts to changing the scale. As trivial as this example is, it highlights a key question:

What is the physical cutoff scale for a given regulator?

In Ref. [7] it has been argued that within practical applications it is suggestive to use the physical gap of the theory as the practical definition of the physical infrared or cutoff scale. Strictly speaking this asks for the evaluation of the poles and cuts of the theory in a real time formulation. For the scalar and Yukawa-type theories explicitly discussed in the present work it has been shown in [42] that the real-time pole masses and the imaginary time curvature masses are very similar in advanced approximations. For the present purpose these subtleties are not relevant and we introduce a definition of the inverse gap as the maximum of the imaginary time propagator,

$$\frac{1}{k^2_{\rm phys}} = \max_{p,\phi} G[\phi, R]. \tag{28}$$

For the sake of simplicity we have restricted ourselves to constant backgrounds ϕ. In the general case, (28) picks out the maximal spectral value of the propagator G [7]. Note also that in theories with several fields one has to monitor the gaps of all the fields involved. In the present work this is important within the example theories studied: the relativistic $O(M)\oplus O(N)$ models as a simple model theory, as well as a non-relativistic Yukawa model for impurities in a Fermi gas. A further intricacy originates in different dispersion relations of the fields involved such as relativistic scalar field with p^2 and fermionic fields with $\not{p}$. Then, the relative physical cutoff scale may involve a nontrivial factor in comparison to the gap. The latter subtlety will be discussed elsewhere.

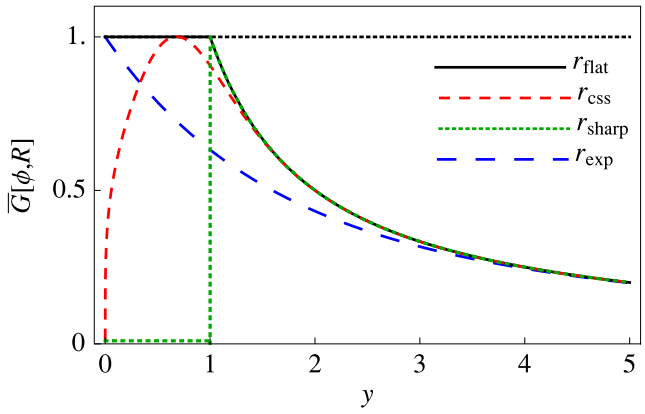

Fig. 8. Comparison of different normalised and regularised propagators $(y(1 + r_i(y)))^{-1}$ with shape functions $r_i(y)$ as listed in Table 1. Here, $r_{\rm css}$ is plotted with $b = 1.25$, $c = 3$ and $h = 0.12$.

Note also, that in (28) a fixed identical RG scheme for all regulators is required, as the propagator is not invariant under RG transformations $\partial_s G[\phi, R] = -2\gamma_\phi G[\phi, R]$, cf. Appendix A. Such a fixed RG scheme can be defined by first selecting one specific flow from $k_{\rm phys} = \Lambda$ to $k_{\rm phys}(k = 0)$, the latter being the physical gap of the theory at $k = 0$. Then regulator changing flows such as defined in (10) at the fixed physical UV scale $k_{\rm phys} = \Lambda$ lead to initial conditions within the same RG scheme defined at $k = 0$. This leads to closed flows without taking into account a further RG transformation at $k = 0$. Hence, (28) implies that the normalised dimensionless propagator satisfies a renormalisation group invariant bound,

$$\bar{G}[\phi, R] \leq 1 \,\forall\, p, \phi\,, \quad \text{with} \quad \bar{G}[\phi, R] = k^2_{\rm phys} G[\phi, R]. \tag{29}$$

In summary, we call theories in the presence of a regulator physically equivalent, if the gaps $k_{\rm phys}$ of all fields agree. In Fig. 8 we present some examples for this criterion for classical propagators in a theory with $V''(\phi_{\rm min}) = 0$. These examples are relevant for the LPA approximation which we predominantly use in the present work.

3.2. Optimisation and length of an RG trajectory

Now we are in the position to define the length of a flow trajectory $\mathcal{C}$. Keeping in mind the discussion of the coordinate system in theory space at the beginning of this Section, we reduce the task by using the effective action itself, evaluated on fields close to the solution of the quantum equations of motion $\phi_{\rm min}$ with

$$\left.\frac{\delta\Gamma[\phi, R]}{\delta\phi}\right|_{\phi=\phi_{\rm min}} = 0. \tag{30}$$

The value of the effective action has no physics interpretation and depends on the renormalisation procedure, i.e., the chosen regulator and initial condition. Therefore, we resort to the second derivative $\Gamma^{(2)}[\phi, R]$ rather than to $\Gamma[\phi, R]$ itself. Indeed, the natural choice is the connected two-point correlation function or rather the normalised dimensionless two-point function $\bar{G}[\phi, R] = k^2_{\rm phys}\langle\phi(p)\phi(-p)\rangle_c$, cf. (29). Here, the subscript c refers to the connected part. This is motivated by the fact that the master equation (2) only depends on the propagator, as do the operator representations for the total t- and s-derivatives, (14) and (B.2).

Measuring the length of the flow of the dimensionless propagator $\bar{G}[\phi, R]$ requires a coordinate system in theory space where the axes are, e.g., expansion coefficients of the propagator, $\bar{G}^{(n)}$ or the spectral values of the propagators linked to an expansion in the eigenbasis of $\bar{G}$

$$\bar{\lambda}_{\rm max} = 1\,, \quad \text{for} \quad \bar{\lambda} \in \operatorname{spec} \bar{G}[\phi, R(k_{\rm phys})]. \tag{31}$$

To summarise, the procedure for finding an optimal regulator for a given truncation of the effective action consists of: choosing an operator norm $\|.\|$ for $\bar{G}$ as well as for $\partial_s \bar{G}$, then, defining the length of a flow trajectory $\mathcal{C}$ with that of the length of $\|\partial_s \bar{G}\|$. Given a set of regulators, the one with the shortest flow length is singled out.

3.3. Functional optimisation criterion

Before we come to integrated flows, let us evaluate the consequences of the discussion above. Firstly, we assert that monotonous flows are shorter than non-monotonous ones. Assuming already a restriction to monotonous flows for G and hence $\bar{G}$, we find a simple optimisation criterion in terms of the dimensionless propagator: $\bar{G}$ is bounded from above by unity, see (29). Moreover, for optimal regulators the propagator is already as close as possible to this bound due to its monotonous dependence on t. This leaves us with

$$\|\bar{p}^2(\bar{G}[\phi, R_{\rm opt}] - \bar{G}[\phi, 0])\| = \min_{R\in R(k_{\rm phys})} \|\bar{p}^2(\bar{G}[\phi, R] - \bar{G}[\phi, 0])\|, \tag{32}$$

where $\bar{p}^2 = p^2/k^2_{\rm phys}$. The prefactor $\bar{p}^2$ has been introduced for convenience to easily accommodate also for massless modes at vanishing cutoff scale. The criterion (32) has been derived in [7], where it also has been shown that for optimised regulators local integrability is restored.

With Eq. (32), for a given background ϕ, we have reduced theory space to a one-dimensional subspace with a simple cartesian norm. Still, the space of regulators is infinite-dimensional and the length of a given flow curve parametrised by s is related to the size of the flow operator equations (14) and (B.2) for t-flows or s-flows, respectively. The flow operators involve second-order ϕ derivatives as well as kernel of the flow operator,

$$\mathcal{K}[\phi, R] = G[\phi, R]\, \partial_s R\, G[\phi, R]. \tag{33}$$

The ϕ derivatives act on the complete set of observables and their action is general. Therefore, we simply have to integrate the size of $\mathcal{K}[\phi, R]$ along the flow for computing a relevant length. For constant backgrounds ϕ we integrate over all spectral values of the operator

$$\|\mathcal{K}[\phi, R]\| = \int_0^\infty dp^2\, |\mathcal{K}[\phi, R]|, \tag{34}$$

giving a dimensionless quantity. This spectral definition can be extended to general backgrounds. Moreover, it can be extended to more general norms that, e.g., take into account the importance of smooth regulators for the derivative expansion [7]. The norm in (34) diverges for $\mathcal{K}[\phi, R]$ showing a infrared singularity with more than $1/p^2$. This can be amended with additional powers of $\bar{p}^2$.

In summary this leads us to the final expression for the length of a trajectory at a given value of $V''(\phi)$,

$$L[V'', R] = \int_0^1 ds \sqrt{1 + \|\mathcal{K}[\phi(V''), R]\|^2}, \tag{35}$$

where $\phi(V'')$ is chosen such that $V''(\phi(\omega))$ is fixed. Then, (35) is the length of the trajectory for G,

$$\int_0^1 \sqrt{ds^2 + \|dG[\phi(V''), R]\|^2} = L[V'', R], \tag{36}$$

where we have used that $\partial_s V'' \equiv 0$. With (36) the optimisation criterion (32) now can be recast into

$$L[V'', R_{\rm opt}] = \min_{R\in R(\Lambda_{\rm phys})} L[V'', R], \tag{37}$$

where $\Lambda_{\rm phys}$ indicates that all flows start at the same physical scale. Note that identical physical UV scales are typically easily identified. Hence, for global flows from the ultraviolet to the infrared we have trajectories with $\Gamma_\Lambda = \Gamma[\phi, R_\Lambda]$ with $R(s = 0) = R_\Lambda$, and $\Gamma_{k=0} = \Gamma[\phi, 0]$ with $R(s = 1) \equiv 0$. The optimal regulator should minimise the length of the flow $L[V'', R]$ for all ϕ. A comparison of the length for different regulators will be presented in the following Section 3.4.

We close with the remark that both criteria, (32) and (37), implement the functional optimisation criterion from [7], and hence are identical. In practical applications the one or the other may be more easily accessible.

3.4. Practical implementation

Let us exemplify the above construction at the example of the LPA approximation for one real scalar field. Its propagator for a given gap $k_{\rm phys}$ reads

$$G[\phi_{\rm min}, R] = \frac{1}{p^2 + \omega_{\rm min} + R}, \tag{38}$$

where it is understood that the cutoff scales in the regulator R is adjusted such that the maximum of the propagator is $1/k^2_{\rm phys}$, and $\omega_{\rm min} = V''(\phi_{\rm min})$ stands for the curvature at the minimum of the effective potential. Now, we use that an optimised regulator minimises infinitesimal flows as well as the rest of the flow towards $k = 0$. This statement holds for correlation functions and, in particular, for the propagator entailing that the difference between the optimal propagator for a given physical cutoff scale $k_{\rm phys}$ and the propagator at $k = 0$ is minimal.

Let us assume for the moment that $\omega_{\rm min}$ is already at the value it acquires at $k = 0$. Then, we are left with the condition to minimise

$$\begin{aligned} &|G[\phi_{\rm min}, R] - G[\phi_{\rm min}, 0]| \\ &\qquad = \left| \frac{R}{(p^2 + \omega_{\rm min} + R)(p^2 + \omega_{\rm min})} \right|, \end{aligned} \tag{39}$$

for all momenta with the constraint (28). We now make a further simplification and set $\omega_{\rm min} = 0$. Then, we are left with minimising

$$\left| \frac{r}{(p^2 + R)} \right|, \tag{40}$$

for all momenta. For momenta $p^2 \geq k^2_{\rm phys}$ we immediately arrive at $r_{\rm opt} = 0$. For $p^2 < k^2_{\rm phys}$ the regulator has to be positive in order to account for the gap condition (28). If this condition is saturated, (39) is minimised, leading to $p^2 + p^2 r_{\rm opt} = k^2_{\rm phys}$, and hence $r_{\rm opt} = k^2_{\rm phys}/p^2 - 1$ for the momenta $p^2 < k^2_{\rm phys}$. In combination with the vanishing for $p^2 \geq k^2_{\rm phys}$ this leads to the unique optimised regulator in LPA,

$$r_{\rm opt} = \left(\frac{k^2_{\rm phys}}{p^2} - 1 \right) \theta(k^2_{\rm phys} - p^2), \tag{41}$$

the flat or Litim regulator [27–29]. Note that there it has been introduced as one of a set of optimised regulators, being distinct by its analytic properties. It has been singled out as the unique solution of the functional optimisation in Ref. [7]. Indeed, the critical exponents in $O(N)$ theories truncated in a local potential approximation with this regulator are closest to the physical ones. The above simplified derivation can be upgraded to also take into account a given fixed $\omega_{\rm min}$. Again this leads to (41) with $k^2_{\rm phys} \to k^2$ where k^2 runs from $k^2_{\rm phys} - \omega_{\rm min}$ to zero.

Let us now also compare the lengths of the trajectories as defined in Section 3.2. The definition was adjusted such that it does not require the knowledge about $\Gamma^{(2)}[\phi, R]$ along the flow leading to simple practical computations. A more elaborate version of the present case is straightforwardly implemented by relaxing $\partial_s V'' \neq 0$.

In Fig. 9 we first compare the norms, (34), of the flow operator for different values of $\omega = V''/k^2$. We show the deviation of the ratios $\|\mathcal{K}[\omega, R_{{\rm exp},n}]\|/\|\mathcal{K}[\omega, R_{\rm flat}]\| - 1 \geq 0$ from one for all values of ω. Here, $R_{{\rm exp},n}$ are the exponential regulators with the corresponding shape function

$$r_{{\rm exp},n}(y) = \frac{y^{n-1}}{e^{y^n} - 1}, \tag{42}$$

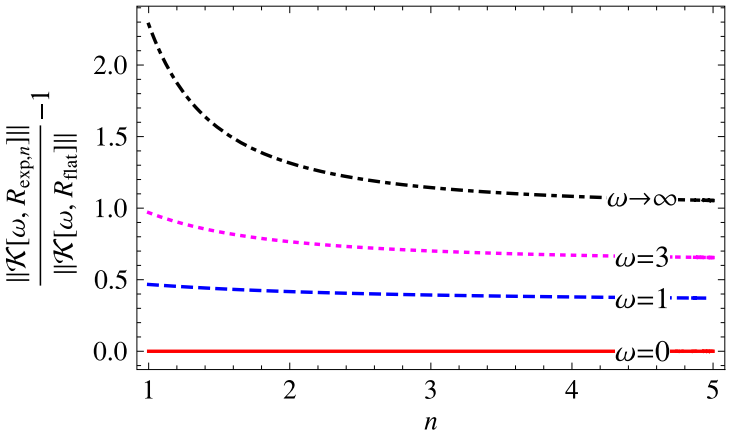

Fig. 9. Deviation of the ratio of $\|\mathcal{K}[\omega, R_{\mathrm{exp},n}]\|/\|\mathcal{K}[\omega, R_{\mathrm{flat}}]\| - 1$, as a function of the power n in the exponential, for $\omega \in [0, \infty)$. For $\omega \to \infty$ the bound saturates. For $\omega = 0$ the norm is a total derivative and is independent of the choice of regulator, hence the ratio is unity: $\omega = 0$: red straight line, $\omega = 1$: blue dashed line, $\omega = 3$: magenta dotted line, $\omega = \infty$: black dashed–dotted line.

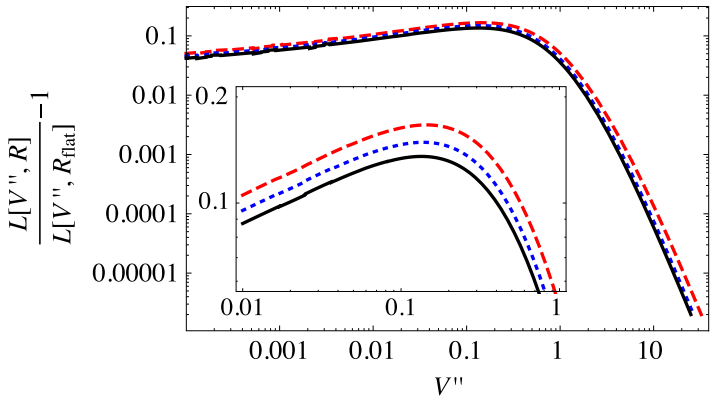

Fig. 10. Length $L[V'', R]/L[V'', R_{\mathrm{flat}}] - 1$ for different exponential regulators in comparison to the flat one: $n = 1$: red dashed line, $n = 2$: blue dotted line, $n = 4$: black straight line. The length is minimised for the flat regulator.

which is a specific subclass of the exponential interpolating regulator in Table 1 with $a = 1, b = 0$. The deviation is always bigger than zero, which singles out the flat regulator as the optimal one in LPA, see Fig. 9. Note in this context that Fig. 9 gives us the full information of the relative size of the integrands in the length of the flow in (35): for a given $V'' > 0$ the related ω diverges with $1/k^2_{\mathrm{phys}}$. This is easily confirmed with the explicit computation of the length, summarised in Fig. 10, where the global length is shown for given V''. One can observe that the flat regulator minimises the flow length $L[V'', R]$ which supports the optimisation criterion developed in Section 3.2.

Note also that the norms are defined such that the information about the physical scales $k^2_{\mathrm{phys}}(s = 1)/k^2_{\mathrm{phys}}(s = 0)$ is only encoded in $\omega(s = 1)/\omega(s = 0)$. Hence, for large V'' in comparison to the physical cutoff scales the difference between the different flows is large. However, in this regime the absolute size of the flow is small.

This concludes the discussion of the constructive criterion for finding an optimised regulator. Note that we have concentrated on the LPA for the sake of simplicity. In LPA we have the peculiar feature of a universal solution for the optimal regulator (41), neither depending on the dimension of the theory nor on the details of the interaction. Partly this property also holds beyond LPA due to the functional nature of the optimisation. Still, in general the optimal regulator depends on the full two-point function $\Gamma_k^{(2)}(p)$ or parts of it, for more details see [7]. This takes into account the details of the theory in a functional way. Moreover, an additional dimensional dependence originates in the operator norms used within the optimisation procedure. While highly interesting, a related discussion is beyond the scope of the present work.

4. Critical behaviour of multi-field models

Many interesting systems include a collection of different field degrees of freedom. In this situation the choice of suitable combinations of regulators is not straightforward and we have already mentioned the relativistic Yukawa models with structurally different dispersion for scalars and fermions. Here we study this case within a simple situation, a bosonic $O(M)\oplus O(N)$ model in $d=3$ dimensions. We show that the choice of relative cutoff scales generally has a crucial impact on the obtained results for the critical physics: The $O(M)\oplus O(N)$ model has two competing order parameter fields and the competing order makes it particularly sensitive to small effective changes of the relevant parameters. The model is studied in the lowest order of the derivative expansion, in LPA. It is well known that such a truncation already captures well the critical physics of scalar models despite the lack of non-trivial momentum dependences of propagators and vertices. The latter encode the anomalous dimensions of the system which are quantitatively small, here, and hence can be neglected.

However, the momentum dependences are also important for taking into account the momentum transfer present in the diagrams on the right hand side of the flow equation. For identical physical cutoff scales this momentum transfer is minimised. In turn, for shifted relative physical cutoff scales of different fields the diagrams have a sizeable momentum transfer. In such a case, physics that is well incorporated in the LPA with identical physical cutoff scales, is lost if the difference between the physical cutoff scales grows large. If one goes beyond LPA within systematic momentum-dependent approximation schemes this relative cutoff scale dependence eventually disappears. The discussion also emphasises the necessity of identical physical cutoff scales within a given approximation in the sense of an optimisation of approximation schemes.

In the present section we highlight the physics changes that are triggered by the change of relative cutoff scales in LPA. As discussed above, due to the missing momentum dependences of LPA, different relative cutoff scales effectively lead to different actions at $k=0$, see also Fig. 3. In LPA, the bosonic, d-dimensional $O(M)\oplus O(N)$ model has the following effective action, [43–47],

$$\Gamma_k=\int_x\left[\frac{1}{2}(\partial_\mu\phi)^2+\frac{1}{2}(\partial_\mu\chi)^2+U_k(\phi,\chi)\right], \tag{43}$$

where ϕ and χ are N- and M-component fields, respectively. The effective potential $U_k(\bar\rho_\phi,\bar\rho_\chi)$ only depends on the field invariants $\bar\rho_\phi=\phi^2/2$ and $\bar\rho_\chi=\chi^2/2$. The scale-dependent dimensionless effective potential is given by

$$u=u(\rho_\phi,\rho_\chi)=k^{-d}U_k(\bar\rho_\phi,\bar\rho_\chi)\quad\text{with}\quad\rho_i=k^{2-d}\bar\rho_i, \tag{44}$$

and $i\in\{\phi,\chi\}$. We further introduce the shape functions r_ϕ and r_χ to regularise the ϕ and χ field modes, respectively. The flow equation for the dimensionless effective potential (44) reads

$$\begin{aligned}\partial_t u=&-d\,u+(d-2)\rho_\phi u^{(1,0)}+(d-2)\rho_\chi u^{(0,1)}\\&+I_{R,\phi}(\omega_\phi,\omega_\chi,\omega_{\phi\chi})+(N-1)I_{G,\phi}(u^{(1,0)})\\&+I_{R,\chi}(\omega_\phi,\omega_\chi,\omega_{\phi\chi})+(M-1)I_{G,\chi}(u^{(0,1)}),\end{aligned} \tag{45}$$

where we have introduced suitable threshold functions $I_{i,j}(x,y,z)$, $i\in\{R,G\}$, $j\in\{\phi,\chi\}$ to separate the loop integration over the radial and Goldstone modes for the two fields. The explicit expressions for these threshold functions are listed in Appendix D. The arguments of the threshold functions are given by $\omega_\phi=u^{(1,0)}+2\rho_\phi u^{(2,0)}$, $\omega_\chi=u^{(0,1)}+2\rho_\chi u^{(0,2)}$ and $\omega_{\phi\chi}=4\rho_\phi\rho_\chi\left(u^{(1,1)}\right)^2$. For calculations, we expand the effective potential about the flowing minimum $(\kappa_\phi,\kappa_\chi)$, to wit

$$u(\rho_\phi,\rho_\chi)=\sum_{1\le i+j\le\mathrm{ord}}\frac{\lambda_{ij}}{i!j!}(\rho_\phi-\kappa_\phi)^i(\rho_\chi-\kappa_\chi)^j. \tag{46}$$

In this truncation we follow the flow of the couplings κ_i and λ_{ij} which are given in Appendix E. These $O(N)\oplus O(M)$ models possess a rich variety of fixed points exhibiting different types of multi-critical behaviour relevant to a number of physical systems [48–51]. For our further investigations, we list the properties of the three most important fixed points:

Table 2
List of values for the largest critical exponent y_1 in $d = 3$ dimensions for selected fixed points in LPA to ascending order in the truncation.

Fixed point	ρ^2	ρ^3	ρ^4	ρ^5	ρ^6
BFP in $O(2) \oplus O(2)$	1.78	1.61	1.78	1.77	1.75
BFP in $O(1) \oplus O(2)$	2.16	1.42	1.65	1.63	1.62
WF in O(2)	1.61	1.32	1.40	1.42	1.41
WF in O(1)	2.00	1.37	1.54	1.55	1.54

(i) The *decoupled fixed point* (DFP) is characterised by a decomposition into two disjoint $O(N)$ and $O(M)$ models where all mixed interactions vanish, e.g., $\lambda_{11} = 0$. It inherits all of the properties of the Wilson–Fisher (WF) fixed points of the separate sub-sectors.

(ii) The *isotropic fixed point* (IFP) features a symmetry enhancement where at each order in the fields the couplings are degenerate, e.g., $\lambda_{20} = \lambda_{02} = \lambda_{11}$. Therefore, the fixed point coordinates agree with the ones from an $O(N+M)$ symmetric model and it inherits all of its critical exponents.

(iii) The *biconical fixed point* (BFP) is a non-trivial fixed point with interactions in both sectors that does not provide a symmetry enhancement. This fact makes it interesting for our further analysis, because it can be easily distinguished by means of the critical exponents of a single field model.

We have listed values for the largest critical exponent $y_1 = 1/\nu_1$ for the models and fixed points which are important to this work in Table 2, showing results for different levels of the truncation. Generally, we sort the critical exponents according to the definition $y_1 > y_2 > y_3 > \cdots$.

4.1. Fixed points and relative cutoff scales

In this section, we examine the effects that occur when dealing with models whose sub-sectors are defined on separate cutoff scales. To that end, we investigate the $O(M) \oplus O(N)$ model using flat regulators in the two sectors, however, with separated cutoff scales

$$r_\phi(y) = \left(\frac{1}{y} - 1\right)\theta(1-y)\,,$$
$$r_\chi(y) = r_\phi\left(\frac{y}{c}\right) = \left(\frac{c}{y} - 1\right)\theta\left(1 - \frac{y}{c}\right)\,. \tag{47}$$

Consequently, for a generic $c \neq 1$ the fluctuations of one field are integrated out earlier than the fluctuations of the other: The second regulator has a built-in shift of all scales $k^2 \to c\,k^2$, $\Lambda^2 \to c\,\Lambda^2$. For $c < 1$ this leads to a suppression of the χ sector and the RG flow does not experience any χ fluctuations. Inversely, $c > 1$ suppresses the ϕ sector in a similar way. Only for $c = 1$ the physical cutoff scales are identical. Note that the statement about identical physical cutoff scales $k_{\rm phys}(\phi) = k_{\rm phys}(\chi)$ is only trivial in the present case where ϕ and χ are both scalar fields with the same dispersion and interactions. In the general case it is non-trivial to identify the relative cutoff scales and where the different representations of the optimisation may pay-off in particular.

The threshold functions for this choice of regulators with separate cutoff scales can be found in Appendix D. Now, we discuss the dependence of critical exponents of different fixed points on changes of the relative cutoff scale.

DFP critical exponents. At the DFP the fields ϕ and χ decouple. Introducing a scale-shifted regulator in a single $O(N)$-model, does not induce a difference in the critical exponents, since every momentum can simply be rescaled. The results of our investigation for the DFP in the $O(2) \oplus O(1)$ model are displayed in the left panel of Fig. 11, confirming the previous statement. The critical exponents for all values of c can be identified with the Wilson–Fisher critical exponents Table 2 at the corresponding order of the truncation.

IFP critical exponents. Introducing $c \neq 1$ revokes the symmetry between ϕ and χ, destroying the key property of this fixed point. Therefore, we will not further investigate this fixed point in the present context. However, we note that this is an illustrative example for how a unphysical regulator choice can lead to artificial regulator dependencies of FRG predictions.

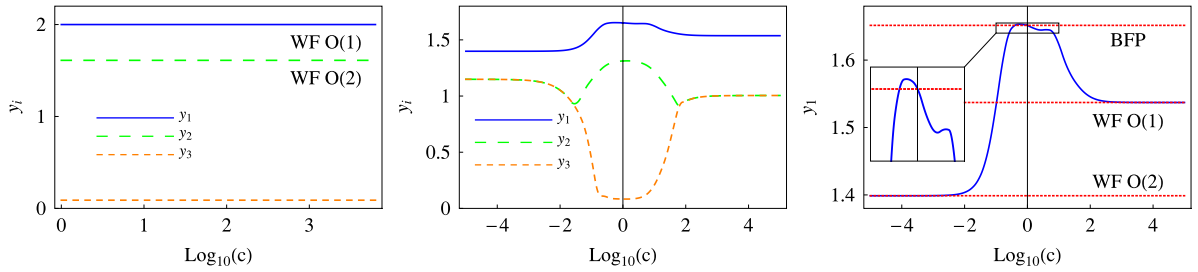

Fig. 11. Left panel: DFP critical exponents of the $O(2) \oplus O(1)$ model in LPA to order ρ^2 using separated cutoff scales. Due to the complete decoupling of both sectors, no dependence on the relative cutoff scale c can be observed. Middle panel: BFP critical exponents of the same model in LPA to order ρ^4. Here, the two sectors are coupled showing a severe dependence on c. Right panel: Close-up of the largest BFP critical exponent y_1 of the $O(2) \oplus O(1)$-model in LPA to order ρ^4 exhibiting the limiting cases $c \ll 1$ and $c \gg 1$ where one of the sectors is effectively suppressed and the critical behaviour is described by the Wilson–Fisher FP of the other sector.

BFP critical exponents. New insights can be gained by looking at the c dependence of the BFP where both sub-sectors are explicitly coupled. The critical exponents of the BFP clearly exhibit a severe dependence on the relative scale factor c, see middle and right panel of Fig. 11 for the example of the $O(2) \oplus O(1)$ model. In fact, the largest critical exponent, y_1, of the BFP tends to the WF critical exponents of one of the single sub-sectors as the other one is suppressed by a large relative cutoff scale. Explicitly, for $c \ll 1$, the critical exponent y_1 of the $O(2) \oplus O(1)$ BFP approaches the value of the WF fixed point of the $O(2)$ model. Analogously, for $c \gg 1$, y_1 approaches the WF fixed point of $O(1)$ model.

We assert that in systems with various field degrees of freedom, the choice of their relative cutoff scales has a severe impact on the described physics in LPA due to the missing momentum dependences. The change of the ratio c of the cutoff scales induces a change of universality classes. In the present simple $O(M) \oplus O(N)$ models the generic choice is $c = 1$ as the fields involved have identical dispersions and interactions. Identical dispersion relations minimise the momentum transfer, as discussed at the beginning of this section. We also emphasise again, that in more complicated systems with different sectors, and in particular fermion–boson systems, there is no clear *a priori* criterion for a suitable choice of regulators and their relative cutoff scales, see also the following section. In the inset of the right panel of Fig. 11, we further show that the value of the critical exponent y_1 at $c = 1$ is not singled out as a local extremum of the critical exponent $y_1(c)$. We note that, in contrast, the principle of minimal sensitivity [20] would single out one of the various extrema in the vicinity of $c = 1$ as a numerical prediction of the BFP critical exponent. We suggest that a control of this issue in multi-field models can be gained by the constructive optimisation procedure presented in Section 3 which, however, is beyond the scope of the present work.

4.2. Truncation dependence

The integrability condition (12) fails as we truncate the effective action. In turn, this suggests that the regulator dependence of the results becomes weaker when the level of truncation is increased. In this section we examine the dependence of the BFP critical exponent y_1 on the relative cutoff scale factor c as a function of the level of truncation. We focus on the $O(2) \oplus O(2)$ model and compare different orders of the LPA, i.e., to the orders between ρ^2 up to ρ^6.

Fig. 12 shows the deviation of y_1 at a given $c > 1$ from the value at $c = 1$ weighted by its difference to the limiting case of the corresponding $O(2)$ critical exponent

$$|\Delta y_1| = \left| \frac{y_1(c) - y_1(1)}{y_1(1) - y_{1,\mathrm{O}(2)}} \right| . \tag{48}$$

We see that for increasing order in the LPA from ρ^2 up to ρ^6, the dependence of y_1 on the relative cutoff scale c becomes weaker and weaker as suggested by the consecutive flattening of the curves. We conclude that a better truncation is more robust to uneducated regulator choices or in other words a low-level truncation requires a more sophisticated choice of the regulator scheme. This is expected since the untruncated effective action yields the same results for every choice of the regulator.

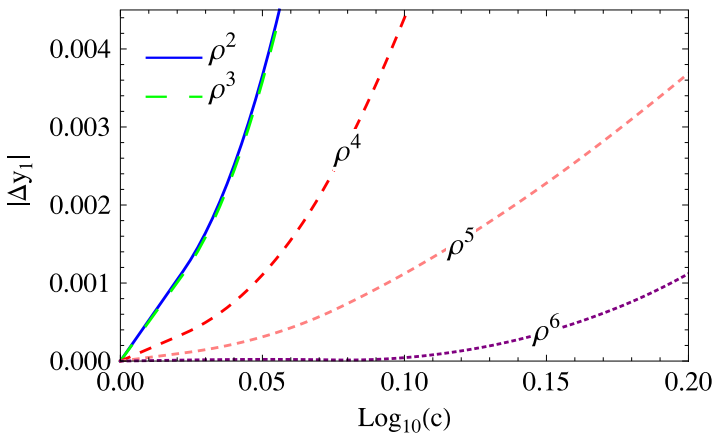

Fig. 12. Deviation from the critical exponent y_1 at the BFP in the $O(2) \oplus O(2)$ model in LPA from order ρ^2 to order ρ^6 as a function of the relative cutoff scale c. We clearly observe a flattening of the curves for higher orders of the truncation.

5. Non-relativistic Fermion–Boson models

In the last Section 4, we have discussed the question of relative cutoff scales in a simple scalar model with identical dispersions and interactions for the different fields. In the present Section we discuss relative cutoff scales and the impact of the shape dependence of regulators in a more complicated situation of a non-relativistic Yukawa system describing fermionic atoms and molecules. In contradistinction to the LPA approximation used in the last Section we also take into account momentum and frequency dependences of the propagators. Naturally, this does not fully cure the lack of momentum dependences of the approximation and we expect a modest regulator dependence of the corresponding results. Further aspects in non-relativistic systems are N-body hierarchies which lead to complete resummation schemes for three-, four- and N-body systems, which has been worked out for fermionic three-, and four-body cases for various systems, [52–57]. This gives us further access for an assessment of the regulator dependence of our results.

The FRG has found multiple applications in the study of non-relativistic systems, ranging from few- [52,53,58] to many-body problems [59–65]. In a prototypical scenario in condensed matter physics, fermionic and bosonic degrees of freedom interact with each other. Such situations occur for instance in models which describe the formation of molecules from atoms. Other examples include the interaction of electrons with collective excitations, such as phonons or magnons. In addition to the question of equivalent cutoff scales of fermions and bosons, similarly to the coupled $O(N)$ models discussed in Section 4, the exact fermionic N-body hierarchies present an additional challenge for the mixed non-relativistic bosonic–fermionic system within the evaluation of FRG flows. Optimal approximations have to take into account these exact hierarchies in addition to taking care of the momentum transfer. If both properties cannot be rescued in a given approximation, it is a priori not clear, in which order the various fields have to be integrated out for optimal results. Hence, in this situation, the question of the optimal ratio of cutoff scales is even more complicated as in the bosonic example treated in Section 4.

In the present Section we do not aim at a full resolution of this intricate question, but rather highlight the ensuing difficulties. We provide an analysis of the regulator dependences arising in a system consisting of a single impurity immersed in a non-relativistic Fermi sea of atoms at zero temperature. In this so-called Fermi polaron problem [66–69] the interaction of the impurity $\psi_\downarrow$ with the fermions $\psi_\uparrow$ in the Fermi sea is determined by the exchange of a molecular field ϕ which represents a bound state of the $\downarrow$-impurity with one of the medium $\uparrow$-atoms. The system is described by the action

$$S = \int_{\mathbf{x},\tau} \Big\{ \sum_{\sigma=\uparrow,\downarrow} \psi_\sigma^*[\partial_\tau - \Delta - \mu_\sigma]\psi_\sigma + \phi^*[\partial_\tau - \Delta/2 + \nu_\phi]\phi + h(\psi_\uparrow^*\psi_\downarrow^*\phi + h.c.) \Big\}, \tag{49}$$

where $\int_{\mathbf{x},\tau} = \int d^3x d\tau$, Δ is the Laplace operator and we suppressed the arguments $\mathbf{x}$, τ of the fields. Furthermore the Grassmann-valued, fermionic fields $\psi_\uparrow$ and $\psi_\downarrow$ represent $\uparrow$- and $\downarrow$-spin fermions of equal mass m. Note that we work in units $\hbar = 2m = 1$ and $\sigma = (\uparrow, \downarrow)$. The associated chemical potentials μ_σ are adjusted such that the $\uparrow$-fermions have a finite density $n_\uparrow = k_F^3/(6\pi^2)$, with k_F the Fermi momentum, while there is only a single $\downarrow$-atom. In this limit the action Eq. (49) describes the problem of a single impurity immersed in a Fermi sea. The detuning ν_ϕ, together with the coupling h determines the interaction strength between the $\downarrow$- and $\uparrow$-atoms which is mediated by the exchange of the field ϕ.

The impurity is dressed by fluctuations in the fermionic background. It becomes a quasi-particle, the Fermi polaron, which is characterised by particle-like properties such as an energy E_p, and a quasi-particle weight Z_p. The quasiparticle properties depend on the interaction between the impurity and the Fermi gas. Due to the presence of bound states, this interaction cannot be described within perturbation theory and requires non-perturbative approximations. Hence, it presents an ideal testing ground for methods such as the FRG.

In the following the prediction of E_p will serve as our observable to study the regulator dependences occurring in the RG evaluation of the model Eq. (49). The Fermi polaron problem is particularly interesting for our study since accurate numerical predictions for various quantities exist based on a bold diagrammatic Monte Carlo scheme [66]. For instance at unitarity, where the infrared scattering amplitude at zero scattering momentum – given by the scattering length a_s – diverges, $k_F a_s \to \infty$, the ground state energy is predicted to approach the value of $E_p = -0.615\,\epsilon_F$ [66], with ϵ_F the Fermi energy.

We note that also other non-relativistic systems of coupled bosons and fermions are described by Eq. (49). For instance for a chemical potential $\mu_\downarrow > -E_p$ the system exhibits the BEC–BCS crossover at low temperature as interactions are varied [70–72]. This crossover has been studied extensively by FRG methods [60,61,73–75].

5.1. Truncation and flow equations

In the following we will solve the FRG flow equation (2) for the truncation of the effective action

$$\begin{aligned}\Gamma_k = \int_{\mathbf{p},\omega} \Big\{ &\psi_\uparrow^*[-i\omega + \mathbf{p}^2 - \mu_\uparrow]\psi_\uparrow + \psi_\downarrow^* G_{\downarrow,k}^{-1}(\omega,\mathbf{p})\psi_\downarrow \\ &+ \phi^* G_{\phi,k}^{-1}(\omega,\mathbf{p})\phi \Big\} + \int_{\vec{x},\tau} h(\psi_\uparrow^*\psi_\downarrow^*\phi + h.c.),\end{aligned} \tag{50}$$

where $\int_{p,\omega} = \int \frac{d^3p}{(2\pi)^3} \int \frac{d\omega}{2\pi}$. In this truncation the only RG scale k dependent quantities are $G_{\downarrow,k}$ and $G_{\phi,k}$. While in previous work the flow of fully momentum dependent propagators $G_{\downarrow,k}$ and $G_{\phi,k}$ has been considered [63], we study here the regulator dependences arising in a field expansion of the derivative expansion where

$$\begin{aligned}G_{\downarrow,k}^{-1}(\omega,\mathbf{p}) &= S_\downarrow[-i\omega + \mathbf{p}^2] + m_\downarrow^2, \\ G_{\phi,k}^{-1}(\omega,\mathbf{p}) &= S_\phi[-i\omega + \mathbf{p}^2/2] + m_\phi^2,\end{aligned} \tag{51}$$

with scale-dependent wave function renormalisations $S_\downarrow, S_\phi$. We neglect higher orders of bosonic interactions. In [63] it has been found that this approximation allows to obtain accurate results for the low-energy properties of the impurity–bath system, such as the ground state energy or the impurity quasi-particle weight $Z_\downarrow = 1/S_\downarrow$. The RG scale dependent coupling constants $m_\downarrow^2$ and m_ϕ^2 are related to the flowing static self-energies $\Sigma_{\downarrow,\phi}(0,\mathbf{0})$, e.g. $m_\downarrow^2 = -\mu_\downarrow - \Sigma_\downarrow(0,\mathbf{0})$. In the impurity problem the majority fermions are not renormalised, $S_\uparrow = 1$, and the density of the Fermi sea is determined by the chemical potential $\mu_\uparrow = \varepsilon_F = k_F^2$. In summary, from this truncation, we obtain the four flow

equations, cf. Appendix H,

$$\begin{aligned}
\partial_t m_\phi^2 &= \frac{h^2}{2\pi^2}\int_{k_F}^{\infty} dp \frac{p^2(\partial_t R_\downarrow + S_\downarrow \partial_t R_\uparrow)}{[m_\downarrow^2 + R_\downarrow + S_\downarrow(2p^2 - \mu_\uparrow + R_\uparrow)]^2}\\
\partial_t S_\phi &= -\frac{h^2}{\pi^2}\int_{k_F}^{\infty} dp \frac{S_\downarrow p^2(\partial_t R_\downarrow + S_\downarrow \partial_t R_\uparrow)}{[m_\downarrow^2 + R_\downarrow + S_\downarrow(2p^2 - \mu_\uparrow + R_\uparrow)]^3}\\
\partial_t m_\downarrow^2 &= \frac{h^2}{2\pi^2}\int_{0}^{k_F} dp \frac{p^2(\partial_t R_\phi - S_\phi \partial_t R_\uparrow)}{[m_\phi^2 + R_\phi - S_\phi(p^2/2 - \mu_\uparrow + R_\uparrow)]^2}\\
\partial_t S_\downarrow^2 &= -\frac{h^2}{\pi^2}\int_{0}^{k_F} dp \frac{S_\phi p^2(\partial_t R_\phi - S_\phi \partial_t R_\uparrow)}{[m_\phi^2 + R_\phi - S_\phi(p^2/2 - \mu_\uparrow + R_\uparrow)]^3}.
\end{aligned} \tag{52}$$

We study the dependence of the predictions from the FRG using a continuous set of regulators $R_{\downarrow,\uparrow}$ and R_ϕ which are dependent on various parameters. We choose

$$\begin{aligned}
R_\phi^k(p) &= c_\phi \frac{S_\phi k^2}{2}(a_\phi - b_\phi y)\frac{y^{n_\phi}}{e^{y^{n_\phi}} - 1}\\
R_\downarrow^k(p) &= c_\downarrow S_\downarrow k^2(a_\downarrow - b_\downarrow y)\frac{y^{n_\downarrow}}{e^{y^{n_\downarrow}} - 1}\\
R_\uparrow^k(p) &= c_\uparrow S_\uparrow k^2(a_\uparrow - b_\uparrow y)\sigma(p^2 - \mu_\uparrow)\frac{y^{n_\uparrow}}{e^{y^{n_\uparrow}} - 1}
\end{aligned} \tag{53}$$

where $y \equiv p^2/(c_i k^2)$ and $\sigma(x) = 1, (-1)$ for $x > 0$ ($x \leq 1$) for the impurity $\psi_\downarrow$ and boson field ϕ. These regulators are similar to the regulators studied in the relativistic models in the previous sections, cf. Table 1. Note however that for the bath fermions the pole structure due to the Fermi surface has to be accounted for so that here $y \equiv (|\mathbf{p}^2 - \mu_\uparrow|)/(c_\downarrow k^2)$. Similar to the definitions used in Section 4, the parameters c_i ($i = \phi, \uparrow, \downarrow$) allow for the study of changing the relative scales at which the various field are integrated out, while the other parameters allow for deformations of the regulator shape, cf. Fig. 8.

5.2. Initial conditions

As discussed in Section 2.1, first the initial values at the UV scale $k = \Lambda$ have to be set. The initial value of m_ϕ^2 is determined by the interaction strength between the impurity and fermions in the Fermi sea. This interaction strength is given by the low-energy scattering length a_s. The latter is determined by the evaluation of the tree-level exchange of the molecule field ϕ in the two-body problem where $\mu_{\downarrow,\uparrow} = 0$. This results in the initial value

$$m_{\phi\Lambda}^2 = -\frac{h^2}{8\pi a_s} - \frac{h^2}{2\pi^2}\int_0^\infty dp\, p^2 \left[\frac{1}{2p^2 + R_\uparrow^\Lambda + R_\downarrow^\Lambda} - \frac{1}{2p^2}\right].$$

For large cutoff scales Λ this implies the scaling $m_{\phi\Lambda}^2 \sim \mu(r)h^2\Lambda$ with $\mu(r)$ being a regulator dependent number. It is the non-relativistic equivalent to the UV scaling discussed earlier, cf. Eq. (9). Furthermore, while the initial value of $S_\downarrow$ is determined by its classical value $S_{\downarrow\Lambda} = 1$, we choose $S_\phi = 0$ so that the bosonic field becomes a pure auxiliary field.

The Fermi momentum allows to define the dimensionless interaction parameter $1/(k_F a_s)$. In the following we work in units where $k_F = 1$. Finally the initial value of $m_{\downarrow\Lambda}^2$ has to be chosen such that the self-energy acquired by the impurity leads to the fulfilment of the infrared condition $m_{\downarrow k=0}^2 = 0$. This condition ensures that the system is just on the verge of occupying a finite number of impurity atoms, which is the defining property of the impurity problem. This condition implies $E_p = -m_{\downarrow\Lambda}^2 = \mu_\downarrow$ [63].

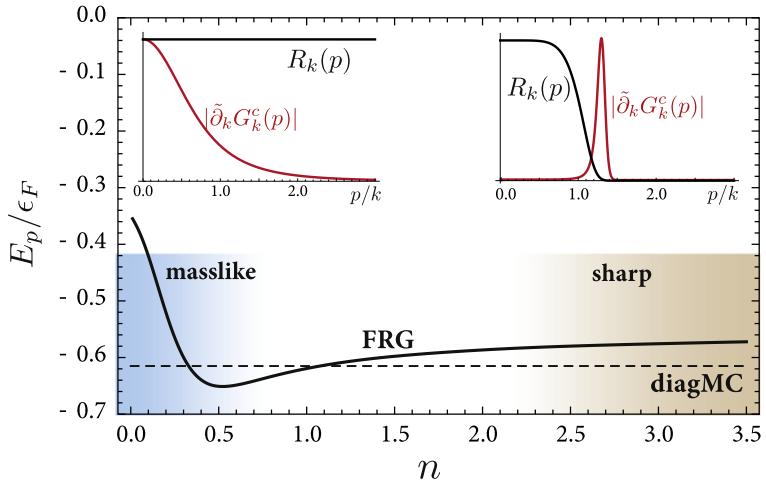

Fig. 13. Dependence of the polaron energy E_p/ϵ_F on the regulator shape. The FRG results (blue) are shown for a crossover from a k^2 to a sharp regulator by changing the exponent $n \equiv n_i$ in all regulators at constant prefactors $a_i = 100$ and $b = 0$. The exact result from diagrammatic Monte Carlo [66] is shown as dashed black line. The insets illustrate the structure of the regulator $R_\downarrow(p)$ and single scale propagator $\tilde{\partial}_k G^c_{\downarrow,k} = -(G^c_{\downarrow,k})^2 \partial_k R_k$ in dependence of momentum p. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

5.3. Regulator dependencies

Shape dependence. First we study the dependence of the results on the shape of the regulators when integrating out the bosonic and fermionic fields synchronously for the choice $c_\phi = c_\downarrow = c_\uparrow = 1$. Specifically, we monitor the energy of the polaron, E_p. In Fig. 13 we show the result for the polaron energy at unitary interactions, $k_F a = \infty$, as a function of the shape parameters $n \equiv n_\uparrow = n_\downarrow = n_\phi$ (blue line). We have studied such a variation previously in Section 3.4 in the context of a relativistic φ^4 theory. Here, we choose $b_i = 0$ and $a_i = 100$ so that the regulator interpolates between a masslike k^2 and a sharp regulator. The black dashed line corresponds to the result obtained from diagrammatic Monte Carlo, $E_p/\epsilon_F \approx -0.615$ [66]. We note that a non-selfconsistent T-matrix approximation yields the result $E_p/\epsilon_F \approx -0.607$ [76]. This approximation (leading order $1/N$ expansion) corresponds to the sequence where first the dimer selfenergy is evaluated and then inserted into the self-energy of the impurity.

The result from the FRG calculation is shown as blue line. We observe a strong regulator shape dependence for small values of n (masslike regulator) as, here, the regulator leads to a non-local integration of field modes in momentum space.

In contrast, for $n \to \infty$ (sharp regulator), the regulator becomes very local in momentum, and the results show only a small shape dependence. In the flows, the single scale propagator $\tilde{\partial}_k G^c_k = -(G^c_k)^2 \partial_k R_k$ determines the locality of the regulator in momentum space which is illustrated in the insets in Fig. 13. Here, we show the form of the single scale propagator as well as $R_\downarrow(p = |\mathbf{p}|)$ evaluated at zero frequency $\omega = 0$. We emphasise that momentum locality of the loop integration is but one of the important conditions for the optimisation. Regularity of the flow is a further important one, and the sharp cutoff fails in this respect. Indeed, it is the latter property which is crucial for critical exponents.

For $n \to 0$ the non-local structure of field integration leads to a great sensitivity of the RG flow of Γ_k in theory space and hence a large sensitivity to the truncation chosen. This also finds an interpretation in terms of physics: due to the non-local structure of $R_k(p)$ the RG flow does not separate between the few-body (vacuum) physics at large momenta on the one hand and on the other hand the emergence of corrections to the vacuum flow due to finite density at small momenta. Such an unphysical mixture of physically vastly separated energy scales leads to an artificially strong dependence of the FRG results. We indicate the regime of artificial non-local, non-physical regulator choices by the grey shaded area in Fig. 13. The results illustrate the significance of the statement that for local truncations non-local cutoffs are a particularly bad choice of regularisation of RG flows. Instead general RG

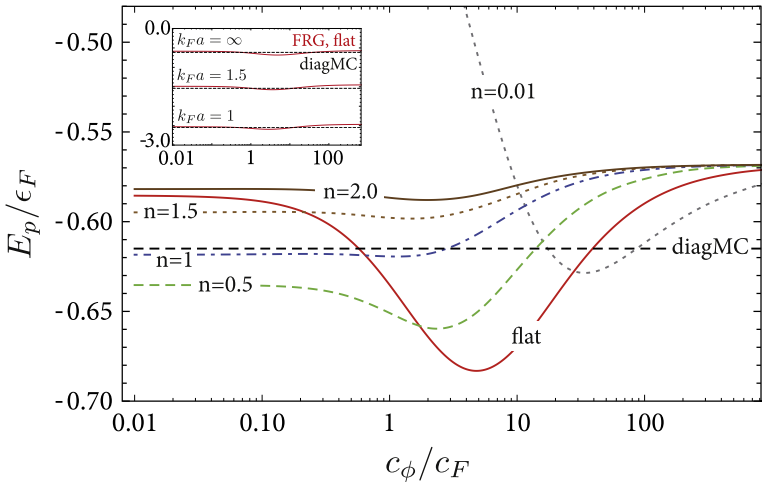

Fig. 14. Dependence of the polaron energy E_p/ϵ_F on the relative bosonic–fermionic cutoff scale for various choices of the regulator shape. Results are shown for parameters as employed in Fig. 13 and also for the flat (Litim) regulator, cf. Table 1. The inset shows the relative cutoff scale dependence for the flat regulator for various interaction strengths $1/k_F a \neq 0$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

flows should be kept sufficiently local. However, also an extremely local regulator such as the sharp regulator is not desired as it prevents an interference of closely related momentum/length scales; it lacks regularity. We have indicated this regime as shaded area at large values of n where the single scale propagator becomes strictly peaked and interference of close-by momentum scales is heavily suppressed. The non-shaded regime corresponds to regulator choices which satisfy the criterion of sufficient interference of momentum scales while still avoiding an unphysical, non-local flow.

In summary, both extreme choices lack crucial properties of optimised FRG flows. This is also reflected in the fact that both limits do not do well in the optimisation criterion in its representations (32) and (37). Indeed, the combination of a sharp cutoff and a mass cutoff gives the worst result within the optimisation as it combines both failures, momentum nonlocality and lack of regularity.

Dependence on relative cutoff scale. Next, we investigate the dependence of the results on the relative scale at which the fermionic and bosonic degrees of freedom are integrated out. As in Section 4, this is achieved by changing the parameter c_ϕ relative to the choice of $c_F \equiv c_\downarrow = c_\uparrow$ in Eq. (53). The result is shown in Fig. 14 where we choose the same regulator shape parameters as in Fig. 13. Also we show the result for a flat regulator choice (red curves).

For $c_\phi/c_F \to \infty$ the flow is equivalent to a purely fermionic flow since here the auxiliary bosons are integrated out only in the last step of the RG flow. In this last step the fermions are not subject to an RG gap R_σ anymore. In consequence, since the flow of the boson propagator is solely dependent on fermions, its self-energy Σ_ϕ reached its final RG value already before this last RG step is taken. This leads to results which are independent of the regulator shape, $E_p/\epsilon_F \to -0.57$, and which is the result obtained from an leading order $1/N$ expansion [77] within our truncation for the momentum dependence of the bosonic propagator P_ϕ.

Contrary, for $c_\phi/c_F \to 0$ the bosonic field is integrated out first. The flow of the boson propagator, being only a functional of the fermionic Green's functions, is then completely suppressed in the first stage of the RG flow. This stage correspondingly amounts to a mere reversion of the introduction of the bosonic field ϕ as an auxiliary degree of freedom mediating the atom–atom interaction. Since P_ϕ cannot acquire any momentum dependence in this step, the resulting – now purely fermionic – theory has a truncation with a completely momentum independent coupling constant $\lambda \sim -h^2/m^2_{\phi,\Lambda}$.

Such a truncation of the effective flowing action Γ_k is of course a very poor one so that strong regulator dependencies are expected as also observed in Fig. 14. This result also represents an example supporting the discussion given in Section 2.2, cf. also Fig. 2: by choosing a poor truncation the flow is particularly sensitive to its path in theory space and hence can lead to strong regulator dependences of infrared quantities. Furthermore this result reflects the observation that the integrability

condition (12) is more severely violated when the effective action is truncated to a larger degree (here by loosing the momentum dependence of interactions altogether).

Having shown that the regulator dependences in the two extreme limits for c_ϕ/c_F can be understood in simple terms we now turn to the intermediate regime where the bosonic and fermionic fields are integrated out synchronously. In this regime we observe a variation of the result for E_p/ϵ_F on the order of $\pm 10\%$, with the exact result $E_p/\epsilon_F = -0.615$ being in the vicinity of the predicted result by the FRG.

We also show the result when applying the flat regulator (red line) which shows similar variations with the relative cutoff scale. Our results indicate that within the regime of an 'informed regulator choice', indicated by the non-shaded region in Fig. 13, regulator dependences in FRG flows might allow for determining an error estimate on its own predictions.

6. Conclusion

In this work, we have presented a systematic investigation of the impact of different regulator choices on renormalisation group flows in given approximation schemes. To this end, we studied the functional RG which is based on the scale-dependent effective action. As an important aspect, this exact flow equation clearly exhibits the role played by the regulator within the RG, see (2), as it is directly proportional to its scale derivative. This already indicates the need for a thorough understanding of regulator dependencies. Such an understanding is not only important to the functional RG, in particular, but, more comprehensively, extends to the analysis of approximations schemes in the renormalisation group framework in general.

Here, we focused on three key aspects of how the regulator choice affects RG results: First, we discussed how the choice of a specific regulator influences FRG flows by integrating over flow trajectories along closed loops in the space of action functionals varying both, the regulator scale and its shape function. For these flows we have discussed an integrability condition, [7], which is violated in the presence of truncations. Consequently, an educated regulator choice is mandatory to extract the best possible results from the RG in a given truncation. To this end we have extended the work on functional renormalisation in [7]. For the construction of such an optimised regulator, we have introduced the definition of the length of an RG trajectory which is minimal for an optimised regulator. This provides a pragmatic optimisation procedure which at the example of a single scalar field yields the flat regulator as a unique and analytical solution. A comparison of the lengths of these trajectories can also be set up straightforwardly in more complex models in order to identify optimised regulators. We leave explicit applications of this procedure for future work.

As a second aspect, we have investigated systems with two field degrees of freedom which both have to be regularised. Here a choice of relative cutoff scales is required. In given momentum-independent approximations this choice has a severe impact on the RG results and, hence, for the described physics. At the example of relativistic bosonic two-field models, we have discussed the consequences of a variation of the relative cutoff scales as well as its truncation dependence. We have shown that a crossover between different universality classes can be induced, triggered by the regulator-dependence of physical parameters in truncated flows. This entails that the relative cutoff scale has to be chosen carefully for a reliable description of a physical system in a given approximation. A controlled approach towards devising an optimised choice of relative cutoff scales can be provided by our optimisation procedure.

Third, we also have exhibited corresponding dependencies on relative cutoff scales and regulator shapes in non-relativistic continuum models of coupled fermionic and bosonic fields. At the example of the Fermi polaron problem in three spatial dimensions, we have illustrated such dependences and showed how to interpret them in physical terms. We suggested that, in the regime of an informed regulator choice, regulator dependences in FRG flows can provide error estimates. This has been discussed here at the example of a coupled non-relativistic many-body model. It will be interesting to investigate these capabilities further in more elaborate many-body models. Finally, it is of great interest to extend the functional optimisation framework laid out here and in [7] to an approach for general systematic error estimates in the functional RG.

Acknowledgements

We thank N. Prokofiev and B. Svistunov for providing their diagMC data. Further, we thank A. Rodigast and I. Boettcher for discussions. R.S. was supported by the NSF through a grant for the Institute for Theoretical Atomic, Molecular, and Optical Physics at Harvard University and the Smithsonian Astrophysical Observatory. S.W. acknowledges support by the Heidelberg Graduate School of Fundamental Physics. This work is supported by the Helmholtz Alliance HA216/EMMI, and the grant ERC-AdG-290623.

Appendix A. Effective action and RG transformations

Note also, that the above renormalisation scheme dependence carries over to the full quantum effective action $\Gamma[\phi] = \Gamma_{k=0}[\phi]$. It satisfies the standard homogeneous RG equation

$$s\frac{d\Gamma_{k=0}[\phi]}{ds} = 0, \tag{A.1}$$

(A.1) is non-trivially achieved as all correlation functions $\Gamma^{(n)}$ transform according to the anomalous dimension of the fields,

$$\left(\partial_s + n\gamma_\phi\right)\Gamma^{(n)} = 0\,, \quad \text{with} \quad \gamma_\phi\phi = s\frac{d\phi}{ds}. \tag{A.2}$$

For the purpose of the present work the RG-transformations of the full effective action are not relevant. Hence, from now we shall identify observables that are identical up to RG transformation of the underlying theory. Note however, that this identification does not remove the relevant UV scaling carried by (9).

Appendix B. General one-parameter flows

We have introduced one-parameter flows, referring to general changes of the cutoff scale k with $k(s)$, changes of the shape of the regulator, r_s as well as reparametrisations of the theory. The corresponding flow equation has the same form as that for the k-flow in (2). It reads

$$s\frac{d\Gamma[\phi,R]}{ds} = \frac{1}{2}\mathrm{Tr}\,G[\phi,R]\left(\partial_s + 2\gamma_\phi\right)R, \tag{B.1}$$

where the total derivative w.r.t. s also includes reparamterisations of the fields with $d\phi/ds = \gamma_\phi\phi$ reflected in the term proportional to the anomalous dimension γ_ϕ on the right hand side of (B.1). The representation of the total s derivative similar to (14) is simply given by

$$\frac{d}{ds} = \left(-\frac{1}{2}\mathrm{Tr}\,G[\phi,R]\,\left(\partial_s + 2\gamma_\phi\right)R\,G[\phi,R]\frac{\delta^2}{\delta\phi^2}\right). \tag{B.2}$$

Note that (B.2) has to vanish as an operator if it represents a reparameterisation of the theory at hand, that is a standard renormalisation group transformation in the presence of a regulator. We infer, [7],

$$\left(\partial_s + 2\gamma_\phi\right)R \stackrel{!}{=} 0. \tag{B.3}$$

Eq. (B.3) entails that the regulator has to be transformed as a two-point correlation function under RG-transformations in order to fully reparameterise the theory.

Appendix C. Integrability condition and self-consistency of approximations

In case the integrability condition (15) holds, the flow necessarily has a (local) representation as a total derivative w.r.t. s, and hence it can be written as a total derivative of a diagrammatic representation. This entails that the integrated flow has a diagrammatic representation in terms of full vertices and propagators in the given approximation to the effective action Γ_k.

A simple example for such an approximation is perturbation theory: at perturbative n-loop order the integrated flow simply reproduces renormalised perturbation theory within a generalised BPHZ-scheme. Note however that the ordering scheme is an expansion in the fundamental coupling of the theory for both, the effective action and the flow equation, rather than one in expansion coefficients of the effective action such as the vertex expansion in terms of $\Gamma_k^{(n)}$. A more interesting example are 2PI-resummation schemes such as 2PI perturbation theory or $1/N$-expansions: it has been worked out how to implement renormalised versions of these schemes in the FRG, see [78,79]. Hence in this case integrated flows provide renormalised perturbative or $1/N$ 2PI-resummations, and the integrability condition (15) is satisfied to any order of such an expansion. Again we note that the ordering scheme is an expansion in the fundamental coupling of the theory or the number of fields for both, the effective action and the flow equation. Similarly it is possible to find approximation schemes that lead to renormalised solutions of Dyson–Schwinger equations.

We emphasise that in both cases discussed above the flow operators (14), (B.2) evaluated on the solution of the effective action in the given approximation, does not satisfy the integrability condition. In the case of n-loop perturbation theory the flow operators (14), (B.2) then generates n-loop FRG-resummed perturbation theory which fails to satisfy (15). In the case of the n-loop 2PI approximation, the flow operators then generate n-loop FRG-resummed 2PI perturbation theory. To summarise, the violation of the integrability condition is a measure for the incompleteness, in terms of the full quantum theory, of fully non-perturbative resummation schemes.

Appendix D. Threshold functions

Scalar model

The scalar model from Section 2.4 requires the threshold function $I(\omega)$ defined in (19). Here, we explicitly give the analytical expressions for this integral for the cases of the flat regulator $r_{\rm L}$ and the sharp regulator $r_{\rm sharp}$. For the flat regulator, we obtain $I(\omega) \to I^{(\rm L)}(\omega)$

$$I^{(\rm L)}(\omega) = v_d \frac{4}{d} \frac{1}{1+\omega}. \tag{D.1}$$

Choosing the sharp regulator yields $I(\omega) \to I^{(\rm sharp)}(\omega)$

$$I^{(\rm sharp)}(\omega) = -2 v_d \log(1+\omega)\,. \tag{D.2}$$

Two-field-model

For the two-field models from Section 4, we have introduced similar threshold functions reading

$$\begin{aligned} I_{R,\phi}(\omega_\phi, \omega_\chi, \omega_{\phi\chi}) &= v_d \int_0^\infty y^{\frac{d}{2}+1} dy \left(-2 r_\phi'(y)\right) \\ &\quad \times \frac{y(1+r_\chi(y)) + \omega_\chi}{(y(1+r_\phi(y)) + \omega_\phi)(y(1+r_\chi(y)) + \omega_\chi) - \omega_{\phi\chi}}, \end{aligned} \tag{D.3}$$

$$\begin{aligned} I_{R,\chi}(\omega_\phi, \omega_\chi, \omega_{\phi\chi}) &= v_d \int_0^\infty y^{\frac{d}{2}+1} dy \left(-2 r_\chi'(y)\right) \\ &\quad \times \frac{y(1+r_\phi(y)) + \omega_\phi}{(y(1+r_\phi(y)) + \omega_\phi)(y(1+r_\chi(y)) + \omega_\chi) - \omega_{\phi\chi}}, \end{aligned} \tag{D.4}$$

and

$$I_{G,i}(x) = v_d \int_0^\infty y^{\frac{d}{2}+1} dy \, \frac{-2 r_i'(y)}{y(1+r_i(y)) + x}, \tag{D.5}$$

with $i \in \{\phi, \chi\}$.

Two-field model & separate cutoff scales

For the discussion of the two-field model in Section 4, we use the flat regulator functions with a relative cutoff scale, as given in Eq. (47). With these shape functions, we obtain the threshold functions for the Goldstone modes

$$I_{G,\phi}(x) = \frac{4v_d}{d\,(x+1)}\,, \quad I_{G,\chi}(x) = \frac{4v_d\,c^{d/2+1}}{d\,(x+c)}. \tag{D.6}$$

The threshold function including radial modes are given by the expressions

$$\begin{aligned} I_{R,\phi}(\omega_\phi,\omega_\chi,\omega_{\phi\chi}) &= 2v_d\theta(1-c)F_1(\omega_\phi,\omega_\chi,\omega_{\phi\chi}) \\ &\quad + \frac{4v_d(c+\omega_\chi)\left(c^{d/2}\theta(1-c)+\theta(c-1)\right)}{d\left((\omega_\phi+1)(c+\omega_\chi)-\omega_{\phi\chi}\right)}, \end{aligned} \tag{D.7}$$

$$\begin{aligned} I_{R,\chi}(\omega_\phi,\omega_\chi,\omega_{\phi\chi}) &= -2c\,v_d\theta(c-1)F_2(\omega_\phi,\omega_\chi,\omega_{\phi\chi}) \\ &\quad + \frac{4v_d(\omega_\phi+1)\left(c^{\frac{d}{2}+1}\theta(1-c)+c\,\theta(c-1)\right)}{d\left((\omega_\phi+1)(c+\omega_\chi)-\omega_{\phi\chi}\right)}, \end{aligned} \tag{D.8}$$

where we have introduced the two integral functions

$$F_1(\omega_\phi,\omega_\chi,\omega_{\phi\chi}) = \int_c^1 \mathrm{d}y \frac{y^{\frac{d}{2}-1}(\omega_\chi+y)}{(\omega_\phi+1)(\omega_\chi+y)-\omega_{\phi\chi}}, \tag{D.9}$$

$$F_2(\omega_\phi,\omega_\chi,\omega_{\phi\chi}) = \int_c^1 \mathrm{d}y \frac{y^{\frac{d}{2}-1}(\omega_\phi+y)}{(c+\omega_\chi)(\omega_\phi+y)-\omega_{\phi\chi}}. \tag{D.10}$$

This completes our list of required threshold functions for the two-field model with shape functions defined on separate cutoff scales.

Appendix E. Flow equations for the couplings

Scalar model

In the symmetric regime, where $\kappa = 0$, we obtain the flow of the coupling constants $\partial_t\lambda_i = (\partial_t u)^{(i)}|_{\rho=0}$ from the ith derivative of $\partial_t u(\rho)$ with respect to ρ.

Analogously, in the symmetry broken regime, where $\lambda_1 = 0$ and $\kappa > 0$, we get

$$\partial_t\kappa = -\frac{(\partial_t u)'}{\lambda_2}\Big|_{\rho=\kappa}, \quad \partial_t\lambda_{i\geq 2} = (\partial_t u)^{(i)} + u^{(i+1)}\partial_t\kappa\,\Big|_{\rho=\kappa}.$$

Two-field model

Projecting the flow equation on the definition of u gives us the system of beta functions for the couplings

$$\begin{aligned} \partial_t\kappa_\phi &= -\frac{\lambda_{02}(\partial_t u)^{(1,0)} - \lambda_{11}(\partial_t u)^{(0,1)}}{\lambda_{20}\lambda_{02}-\lambda_{11}^2}\Big|_{\kappa_\phi,\kappa_\chi}, \\ \partial_t\kappa_\chi &= -\frac{\lambda_{20}(\partial_t u)^{(0,1)} - \lambda_{11}(\partial_t u)^{(1,0)}}{\lambda_{20}\lambda_{02}-\lambda_{11}^2}\Big|_{\kappa_\phi,\kappa_\chi}, \end{aligned} \tag{E.1}$$

and

$$\partial_t\lambda_{ij} = (\partial_t u)^{(i,j)} + u^{(i+1,j)}\partial_t\kappa_\phi + u^{(i,j+1)}\partial_t\kappa_\chi\,\Big|_{\kappa_\phi,\kappa_\chi},$$

where the field invariants ρ_j are understood to be evaluated at their scale dependent expansion points κ_j.

Appendix F. Calculation of explicit loop flows

Explicit loop flows I

Starting with Eqs. (21) and (23), we can calculate a flow which translates continuously from one regulator to another as long as both regulators are finite. These equations can be easily extended to an $O(N)$ model

$$\frac{\mathrm{d}}{\mathrm{d}s}u = J(u' + 2\rho\, u'') + (N-1)J(u') + \frac{\partial_s k(s)}{k(s)}\big(-d\,u + (d-2)\rho\, u'\big)\,. \tag{F.1}$$

A commonly used regulator is the sharp regulator $r_{\rm sharp}(y) = c/y\;\theta(1-y)|_{c\to\infty}$, which is infinite in [0,1]. In order to interpolate between $r_{\rm sharp}$ and other regulators in a continuous manner we need to extend our calculations. Here, we interpolate between $r_{\rm sharp}$ and $r_{\rm L}$ using an interpolation which shifts the cutoff scale in $r_{\rm sharp}$ by a factor $a(s) \in [0,\,1]$.

$$r_s(y) = r_{\rm L}(y) + r_{\rm sharp}\big(y/a(s)^2\big)\,. \tag{F.2}$$

Hence, $a(s) \to 0$ causes $r_{\rm sharp}$ to vanish. On the other hand, if $a(s) = 1$, then $r_{\rm sharp}$ causes the regulator to diverge on $[0,\,1]$ such that there is no residual influence of $r_{\rm L}$. The threshold function can be decomposed into two parts

$$\begin{aligned} J(\omega) &= v_d \int_0^\infty y^{\frac{d}{2}}\,\mathrm{d}y \frac{\frac{d}{ds}r_s(y)}{y(1+r_s(y))+\omega} \\ &= J^{\rm A}(\omega) + J^{\rm B}(\omega)\,, \end{aligned} \tag{F.3}$$

where $J^{\rm A}$ contains the regulator derivative from r_{flat} such that it can be inferred from (D.1)

$$\begin{aligned} J^{\rm A}(\omega) &= \frac{k'(s)}{k(s)}(1-a(s)^d)I^{(\rm L)}(\omega) \\ &= v_d \frac{k'(s)}{k(s)}\frac{4}{d}\frac{1}{1+w}(1-a(s)^d). \end{aligned} \tag{F.4}$$

Similarly, $J^{\rm B}$ corresponds to the regulator derivative of r_{sharp} and can be calculated by inserting (D.2)

$$\begin{aligned} J^{\rm B}(\omega) &= \Big(\frac{a'(s)}{a(s)} + \frac{k'(s)}{k(s)}\Big)a(s)^{d-2} I^{(\rm sharp)}(\omega) \\ &= -2v_d\Big(\frac{a'(s)}{a(s)} + \frac{k'(s)}{k(s)}\Big)a(s)^{d-2}\,\log(1+\omega)\,. \end{aligned} \tag{F.5}$$

Explicit loop flows II

In case we insist on a linear superposition between $r_{\rm sharp}$ and other finite regulators, the solution of the flow equation will show a discontinuity. As soon as we allow for a small contribution from $r_{\rm sharp}$, it will dominate over all finite regulators because it is infinite in the region $[0,\,1]$. This discontinuity can be seen in Figs. 6 and F.15 for the couplings λ_1 and λ_2. In order to calculate the magnitude of this discontinuity, we start at the flow equation at fixed k for an $O(1)$-model, Eq. (21)

$$\frac{\mathrm{d}}{\mathrm{d}s}u|_{k=\rm const} = J(\omega)\,,\;\; \omega = u' + 2\rho\, u''\,, \tag{F.6}$$

which we want to solve from $s = 0$ to $s = 1$. Since the only change occurs at $s = 0$, we can simply denote the magnitude of the discontinuity $\Delta u = u(s=1) - u(s=0)$. The threshold function

$$J(\omega) = v_d \int_0^\infty y^{\frac{d}{2}}\,\mathrm{d}y \frac{\partial_s r_s(y)}{y(1+r_s(y))+\omega}, \tag{F.7}$$

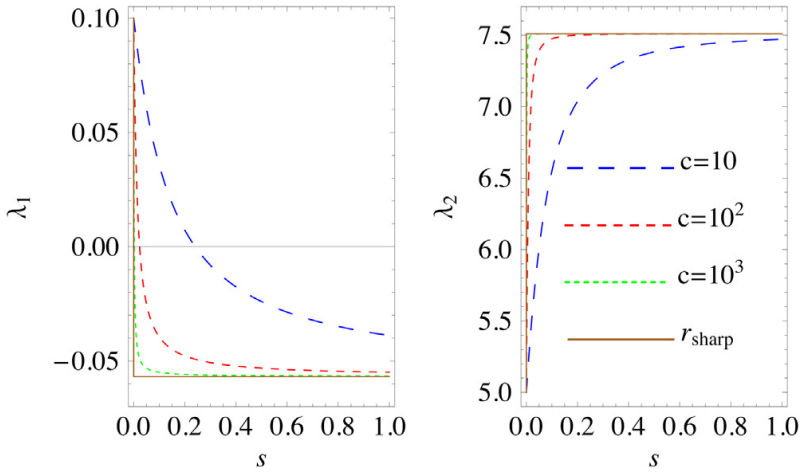

Fig. 15. Change of the initial conditions for different choices of interpolations between the flat regulator and the step like regulator, for various c and the sharp regulator.

only depends on the change of shape $\partial_s r_s(y)$, but not on the scale change, since we are keeping k fixed. Inserting a linear superposition between $r_{\rm sharp}$ and $r_{\rm L}$

$$r_s(y) = (1-s) r_{\rm L}(y) + s\, r_{\rm sharp}(y), \tag{F.8}$$

evaluates to

$$J(w) = v_d \int_0^1 y^{\frac{d}{2}-1} {\rm d}y \, \frac{c(1+1/c-y/c)}{1+sc(1-1/c+y/c)+w}\Big|_{c\to\infty}.$$

We now shift the flow variable from $s \in [0,\,1]$ to $\bar{s} = s\,c \in [0,\,c]$ and take the limit $c \to \infty$. The new modified flow equation, reading

$$\frac{\rm d}{{\rm d}\bar{s}} u|_{k={\rm const}} = \frac{2v_d}{d} \frac{1}{\bar{s}+1+w}, \tag{F.9}$$

must be solved from $\bar{s} = 0$ to $\bar{s} = \infty$. This equation leads to a logarithmic divergence if $\bar{s} \to \infty$. However, the divergent part is just a constant shift of the effective potential which can be removed by subtracting it

$$\frac{\rm d}{{\rm d}\bar{s}} u|_{k={\rm const}} = -\frac{2v_d}{d} \frac{1}{\bar{s}+1} \frac{\omega}{\bar{s}+1+\omega}. \tag{F.10}$$

Our construction ensures that the discontinuity is expressed as $\Delta u = u(\bar{s}=\infty) - u(\bar{s}=0)$ which can be evaluated in a continuous flow equation.

Appendix G. Two-field-model & critical exponents

In Fig. 16, we show a variant of Fig. 12, exhibiting the deviation of the critical exponent y_1 in the $O(2) \oplus O(2)$ model from its value at $c = 1$ without taking the absolute value and the weighting factor from (48).

Appendix H. Flows for the polaron problem

The flow equations, graphically represented in Fig. 17, are given by

$$\begin{aligned}
\partial_k P_{\downarrow,k}(Q) &= h^2 \tilde{\partial}_k \int_P G^c_{\phi,k}(P) G^c_{\uparrow,k}(P+Q) \\
\partial_k P_{\phi,k}(Q) &= -h^2 \tilde{\partial}_k \int_P G^c_{\downarrow,k}(P) G^c_{\uparrow,k}(Q-P),
\end{aligned} \tag{H.1}$$

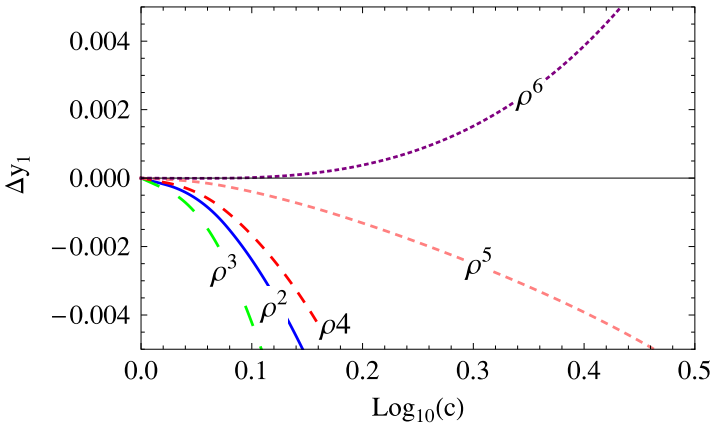

Fig. 16. Deviation from the critical exponent y_1 using separated cutoff scales in the $O(2) \oplus O(2)$ model in LPA from order ρ^2 to order ρ^6.

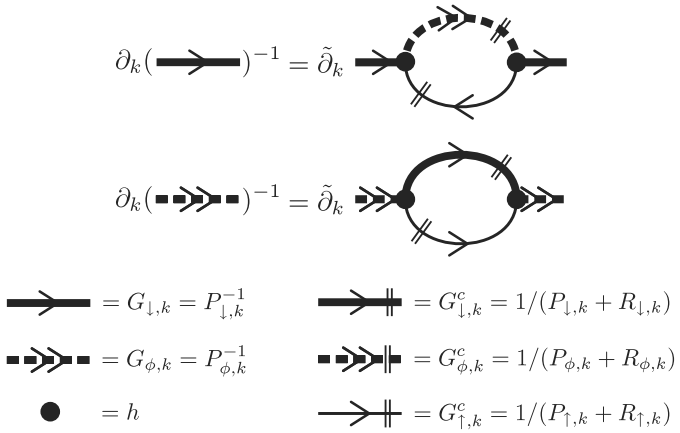

Fig. 17. Feynman diagrams for the flow of the dimer and impurity inverse propagators.

where $P \equiv (\omega, \mathbf{p})$. The flowing inverse propagators $P_k \equiv G_k^{-1}$ on the left-hand side are defined without the regulators, while the regulated propagators G_k^c are given by

$$G_k \equiv 1/P_k G_k^c \equiv 1/(P_k + R_k), \tag{H.2}$$

and $\tilde{\partial}_k$ implies that the derivative acts only on the regulator term R_k inside the cutoff propagators G_k^c.

References

[1] C. Wetterich, Phys. Lett. B 301 (1993) 90.
[2] D.F. Litim, J.M. Pawlowski, The Exact Renormalization Group. Proceedings, Workshop, Faro, Portugal, September 10–12, 1998, 1998, pp. 168–185. hep-th/9901063. URL http://alice.cern.ch/format/showfull?sysnb=0302190.
[3] K. Aoki, Internat. J. Modern Phys. B 14 (2000) 1249.
[4] C. Bagnuls, C. Bervillier, Phys. Rep. 348 (2001) 91. hep-th/0002034.
[5] J. Berges, N. Tetradis, C. Wetterich, Phys. Rep. 363 (2002) 223. hep-ph/0005122.
[6] J. Polonyi, Central Eur. J. Phys. 1 (2003) 1. hep-th/0110026.
[7] J.M. Pawlowski, Ann. Physics 322 (2007) 2831. hep-th/0512261.
[8] H. Gies, Lecture Notes in Phys. 852 (2012) 287. hep-ph/0611146.
[9] B.-J. Schaefer, J. Wambach, Phys. Part. Nucl. 39 (2008) 1025. hep-ph/0611191.
[10] B. Delamotte, Lecture Notes in Phys. 852 (2012) 49. cond-mat/0702365.
[11] P. Kopietz, L. Bartosch, F. Schutz, Introduction to the Functional Renormalization Group, in: Lect. Notes Phys., 2010.
[12] M.M. Scherer, S. Floerchinger, H. Gies, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 368 (2011) 2779. arXiv:1010.2890.
[13] O.J. Rosten, Phys. Rep. 511 (2012) 177. arXiv:1003.1366.

[14] J. Braun, J. Phys. G 39 (2012) 033001. arXiv:1108.4449.
[15] D.F. Litim, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 369 (2011) 2759. arXiv:1102.4624.
[16] R. Percacci, Time and Matter, 2011, pp. 123–142. arXiv:1110.6389. URL http://inspirehep.net/record/943400/files/arXiv:1110.6389.pdf.
[17] L. von Smekal, Nuclear Phys. Proc. Suppl. 228 (2012) 179. arXiv:1205.4205.
[18] M. Reuter, F. Saueressig, New J. Phys. 14 (2012) 055022. arXiv:1202.2274.
[19] I. Boettcher, J.M. Pawlowski, S. Diehl, Nuclear Phys. Proc. Suppl. 228 (2012) 63. arXiv:1204.4394.
[20] P.M. Stevenson, Phys. Rev. D 23 (1981) 2916. URL http://link.aps.org/doi/10.1103/PhysRevD.23.2916.
[21] L. Canet, B. Delamotte, D. Mouhanna, J. Vidal, Phys. Rev. B 68 (2003) 064421. hep-th/0302227.
[22] L. Canet, Phys. Rev. B 71 (2005) 012418. URL http://link.aps.org/doi/10.1103/PhysRevB.71.012418.
[23] F. Benitez, J.-P. Blaizot, H. Chate, B. Delamotte, R. Mendez-Galain, et al., Phys. Rev. E 85 (2012) 026707. arXiv:1110.2665.
[24] F. Benitez, J.-P. Blaizot, H. Chaté, B. Delamotte, R. Mé, N. Wschebor, Phys. Rev. E 80 (2009) 030103. URL http://link.aps.org/doi/10.1103/PhysRevE.80.030103.
[25] R.D. Ball, P.E. Haagensen, I. Latorre, Jose, E. Moreno, Phys. Lett. B 347 (1995) 80. hep-th/9411122.
[26] S.-B. Liao, J. Polonyi, M. Strickland, Nuclear Phys. B 567 (2000) 493. hep-th/9905206.
[27] D.F. Litim, Phys. Lett. B 486 (2000) 92. hep-th/0005245.
[28] D.F. Litim, Phys. Rev. D 64 (2001) 105007. hep-th/0103195.
[29] D.F. Litim, Internat. J. Modern Phys. A 16 (2001) 2081. hep-th/0104221.
[30] D.F. Litim, Nuclear Phys. B 631 (1–2) (2002) 128–158. hep-th/0203006.
[31] D.F. Litim, Renormalization Group. Proceedings, 5th International Conference, RG 2002, Tatranska Strba, Slovakia, March 10–16, 2002, Acta Phys. Slovaca 52 (2002) 635–642. hep-th/0208117. URL http://weblib.cern.ch/abstract?CERN-TH-2002-197.
[32] L. Canet, B. Delamotte, D. Mouhanna, J. Vidal, Phys. Rev. D 67 (2003) 065004. hep-th/0211055.
[33] D.F. Litim, J. High Energy Phys. 0507 (2005) 005. hep-th/0503096.
[34] C.S. Fischer, A. Maas, J.M. Pawlowski, Ann. Physics 324 (2009) 2408. arXiv:0810.1987.
[35] L. Fister, J.M. Pawlowski, 2011, arXiv:1112.5440.
[36] D. Schnoerr, I. Boettcher, J.M. Pawlowski, C. Wetterich, 2013, arXiv:1301.4169.
[37] S. Diehl, H. Gies, J. Pawlowski, C. Wetterich, Phys. Rev. A 76 (2007) 053627. cond-mat/0703366.
[38] D.F. Litim, J.M. Pawlowski, Phys. Rev. D 66 (2002) 025030. hep-th/0202188.
[39] C. Pagani, 2016, arXiv:1603.07250.
[40] I. Nandori, J. High Energy Phys. 1304 (2013) 150. arXiv:1208.5021.
[41] S. Nagy, B. Fazekas, L. Juhasz, K. Sailer, Phys. Rev. D 88 (2013) 116010. arXiv:1307.0765.
[42] A.J. Helmboldt, J.M. Pawlowski, N. Strodthoff, Phys. Rev. D 91 (2015) 054010. arXiv:1409.8414.
[43] A. Eichhorn, D. Mesterházy, M.M. Scherer, Phys. Rev. E 88 (2013) 042141. arXiv:1306.2952.
[44] A. Eichhorn, D. Mesterházy, M.M. Scherer, Phys. Rev. E 90 (2014) 052129. arXiv:1407.7442.
[45] I. Boettcher, Phys. Rev. E 91 (2015) 062112. arXiv:1503.07817.
[46] A. Eichhorn, T. Helfer, D. Mesterházy, M.M. Scherer, 2015, arXiv:1510.04807.
[47] L. Classen, I.F. Herbut, L. Janssen, M.M. Scherer, 2015, arXiv:1510.09003.
[48] J.M. Kosterlitz, D.R. Nelson, M.E. Fisher, Phys. Rev. B 13 (1976) 412. URL http://link.aps.org/doi/10.1103/PhysRevB.13.412.
[49] A. Aharony, Phys. Rev. B 8 (1973) 4270. URL http://link.aps.org/doi/10.1103/PhysRevB.8.4270.
[50] D.R. Nelson, J.M. Kosterlitz, M.E. Fisher, Phys. Rev. Lett. 33 (1974) 813. URL http://link.aps.org/doi/10.1103/PhysRevLett.33.813.
[51] P. Calabrese, A. Pelissetto, E. Vicari, Phys. Rev. B 67 (2003) 054505. URL http://link.aps.org/doi/10.1103/PhysRevB.67.054505.
[52] S. Diehl, H.C. Krahl, M. Scherer, Phys. Rev. C 78 (2008) 034001. URL http://link.aps.org/doi/10.1103/PhysRevC.78.034001.
[53] S. Moroz, S. Floerchinger, R. Schmidt, C. Wetterich, Phys. Rev. A 79 (2009) 042705. URL http://link.aps.org/doi/10.1103/PhysRevA.79.042705.
[54] R. Schmidt, S. Moroz, Phys. Rev. A 81 (2010) 052709. URL http://link.aps.org/doi/10.1103/PhysRevA.81.052709.
[55] M.C. Birse, B. Krippa, N.R. Walet, Phys. Rev. C 87 (2013) 054001. URL http://link.aps.org/doi/10.1103/PhysRevC.87.054001.
[56] Y. Tanizaki, Progr. Theoret. Exp. Phys. 2013 (2013). URL http://ptep.oxfordjournals.org/content/2013/11/113A01.abstract.
[57] B. Jaramillo Ávila, M.C. Birse, Phys. Rev. A 92 (2015) 023601. URL http://link.aps.org/doi/10.1103/PhysRevA.92.023601.
[58] S. Floerchinger, S. Moroz, R. Schmidt, 51, (2011) 153. URL http://dx.doi.org/10.1007/s00601-011-0231-z.
[59] N. Dupuis, K. Sengupta, Europhys. Lett. 80 (2007) 50007. URL http://stacks.iop.org/0295-5075/80/i=5/a=50007.
[60] S. Diehl, H. Gies, J.M. Pawlowski, C. Wetterich, Phys. Rev. A 76 (2007) 021602. URL http://link.aps.org/doi/10.1103/PhysRevA.76.021602.
[61] S. Floerchinger, M. Scherer, S. Diehl, C. Wetterich, Phys. Rev. B 78 (2008) 174528. URL http://link.aps.org/doi/10.1103/PhysRevB.78.174528.
[62] S. Floerchinger, C. Wetterich, Phys. Rev. A 77 (2008) 053603. URL http://link.aps.org/doi/10.1103/PhysRevA.77.053603.
[63] R. Schmidt, T. Enss, Phys. Rev. A 83 (2011) 063620.
[64] A. Rançon, N. Dupuis, Phys. Rev. B 83 (2011) 172501. URL http://link.aps.org/doi/10.1103/PhysRevB.83.172501.
[65] I. Boettcher, J.M. Pawlowski, C. Wetterich, Phys. Rev. A 89 (2014) 053630. arXiv:1312.0505.
[66] N. Prokof'ev, B. Svistunov, Phys. Rev. B 77 (2008) 020408.
[67] A. Schirotzek, C.-H. Wu, A. Sommer, M.W. Zwierlein, Phys. Rev. Lett. 102 (2009) 230402.
[68] R. Schmidt, T. Enss, V. Pietilä, E. Demler, Phys. Rev. A 85 (2012) 021602.

[69] C. Kohstall, M. Zaccanti, M. Jag, A. Trenkwalder, P. Massignan, G.M. Bruun, F. Schreck, R. Grimm, Nature 485 (2012) 615. URL http://dx.doi.org/10.1038/nature11065.
[70] R. Haussmann, W. Rantner, S. Cerrito, W. Zwerger, Phys. Rev. A 75 (2007) 023610.
[71] I. Bloch, J. Dalibard, W. Zwerger, Rev. Modern Phys. 80 (2008) 885.
[72] W. Zwerger, The BCS-BEC Crossover and the Unitary Fermi Gas, Vol. 836, Springer Science & Business Media, 2011.
[73] S. Diehl, C. Wetterich, Nuclear Phys. B 770 (2007) 206. cond-mat/0510407.
[74] S. Diehl, C. Wetterich, Phys. Rev. A 73 (2006) 033615. cond-mat/0502534.
[75] L. Bartosch, P. Kopietz, A. Ferraz, Phys. Rev. B 80 (2009) 104514. URL http://link.aps.org/doi/10.1103/PhysRevB.80.104514.
[76] R. Combescot, A. Recati, C. Lobo, F. Chevy, Phys. Rev. Lett. 98 (2007) 180402.
[77] T. Enss, Phys. Rev. A 86 (2012) 013616. URL http://link.aps.org/doi/10.1103/PhysRevA.86.013616.
[78] J.-P. Blaizot, J.M. Pawlowski, U. Reinosa, Phys. Lett. B 696 (2011) 523. arXiv:1009.6048.
[79] M.E. Carrington, W.-J. Fu, D. Pickering, J.W. Pulver, Phys. Rev. D 91 (2015) 025003. arXiv:1404.0710.

4 Unsupervised Learning of Phase Transitions: From Principal Component Analysis to Variational Autoencoders

Wetzel, Sebastian J.

Phys. Rev. E 96, 022140, Published 18 August 2017 [2]

Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders

Sebastian J. Wetzel
Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany
(Received 13 March 2017; revised manuscript received 9 June 2017; published 18 August 2017)

We examine unsupervised machine learning techniques to learn features that best describe configurations of the two-dimensional Ising model and the three-dimensional XY model. The methods range from principal component analysis over manifold and clustering methods to artificial neural-network-based variational autoencoders. They are applied to Monte Carlo–sampled configurations and have, *a priori*, no knowledge about the Hamiltonian or the order parameter. We find that the most promising algorithms are principal component analysis and variational autoencoders. Their predicted latent parameters correspond to the known order parameters. The latent representations of the models in question are clustered, which makes it possible to identify phases without prior knowledge of their existence. Furthermore, we find that the reconstruction loss function can be used as a universal identifier for phase transitions.

DOI: 10.1103/PhysRevE.96.022140

I. INTRODUCTION

Inferring macroscopic properties of physical systems from their microscopic description is an ongoing work in many disciplines of physics such as condensed matter, ultracold atoms, or quantum chromodynamics. The most drastic changes in the macroscopic properties of a physical system occur at phase transitions, which often involve a symmetry-breaking process. The theory of such phase transitions was formulated by Landau as a phenomenological model [1] and later devised from microscopic principles using the renormalization group [2,3]. Different phases can be identified by an order parameter that is zero in the disordered phase and nonzero in the ordered phase. Whereas in many known models the order parameter can be determined by symmetry considerations of the underlying Hamiltonian, there are states of matter where such a parameter can only be defined in a complicated nonlocal way [4]. These systems include topological insulators, quantum spin Hall states [5], or quantum spin liquids [6]. Therefore, it is important to develop new methods to identify parameters capable of describing phase transitions in these systems.

Such methods might be borrowed from machine learning. With the development of more powerful computers and artificial neural networks, machine learning has become one of the most influential disciplines of this century. It has been shown that such neural networks can approximate any continuous function under mild assumptions [7,8]. They quickly found applications in image classification [9], speech recognition [10], and natural language understanding [11].

In recent years physicists have started to employ machine learning techniques. Most of the tasks were tackled by supervised learning algorithms or with the help of reinforcement learning [12–24]. In supervised learning the algorithm is trained on labeled data to assign labels to data points. After successful training it can predict the labels of previously unseen data with high accuracy.

In addition to supervised learning, there exist unsupervised learning algorithms that can find structure in unlabeled data. It is already possible to employ unsupervised learning techniques to reproduce Monte Carlo–sampled states of the Ising model [25]. Phase transitions were found in an unsupervised manner using principal component analysis [26,27]. In this article we examine several other unsupervised learning algorithms such as manifold methods, clustering, and autoencoders. As a result of this examination we conclude that principal component analysis and variational autoencoders are the most promising among them to reveal phase transitions. This motivates us to transition to the employment of variational autoencoders and test how the latter handles different physical models. This algorithm finds a low-dimensional latent representation of the physical system that coincides with the correct order parameter. Furthermore, we find that autoencoders can reconstruct samples more accurately in the ordered phase, which suggests the use of the reconstruction error as a universal identifier for phase transitions.

Whereas for physicists this work is a promising way to find order parameters of systems where they are hard to identify, computer scientists and machine learning researchers might find an interpretation of the latent parameters.

II. MODELS

A. Ising model in two dimensions

The Ising model is one of the most studied and well understood models in physics. Whereas the one-dimensional classical Ising model does not possess a phase transition, the two-dimensional model does. The Hamiltonian of the Ising model on the square lattice with vanishing external magnetic h field reads

$$H(\mathbf{S}) = -J \sum_{\langle i,j\rangle_{\mathrm{NN}}} s_i s_j, \tag{1}$$

with uniform interaction strength J and spins $s_i \in \{+1 = \uparrow, -1 = \downarrow\}$ on each site $i = 1, \ldots, N$. The notation $\langle i,j\rangle_{\mathrm{NN}}$ indicates a summation over nearest neighbors. A spin configuration $\mathbf{S} = (s_1, \ldots, s_N)$ is a fixed assignment of a spin to each lattice site and Λ denotes the set of all possible configurations $\mathbf{S}$. We set the Boltzmann constant $k_B = 1$ and the interaction strength $J = 1$ for the ferromagnetic case and $J = -1$ for the antiferromagnetic case. A spin configuration $\mathbf{S}$ can be

expressed in matrix form as

$$\underline{S} \hat{=} \begin{pmatrix} \uparrow & \downarrow & \uparrow & \cdots & \uparrow \\ \downarrow & \uparrow & \uparrow & \cdots & \uparrow \\ \vdots & \vdots & \vdots & & \vdots \\ \downarrow & \downarrow & \uparrow & \cdots & \downarrow \end{pmatrix}_{L\times L}. \tag{2}$$

Onsager solved the two-dimensional Ising model [28] and thereby calculated the critical temperature $T_c = 2/\ln(1+\sqrt{2}) = 2.269$.

For the purpose of this work, we assume a square lattice with length $L = 28$ such that $L \times L = N = 784$ and periodic boundary conditions. We sample the Ising model using a Monte Carlo algorithm [29] at temperatures $T \in [0,5]$ to generate 50 000 samples in the ferromagnetic case and 10 000 samples in the antiferromagnetic case. The Ising model obeys a discrete $\mathbb{Z}_2$ symmetry, which is spontaneously broken below T_c. The magnetization of a spin sample is defined as

$$M(\mathbf{S}) = \frac{1}{N}\sum_i s_i. \tag{3}$$

The partition function

$$Z = \sum_{\mathbf{S}\in\Lambda} \exp\left[-H(\mathbf{S})/T\right] \tag{4}$$

allows us to define the corresponding order parameter. It is the expectation value of the absolute value of the magnetization at fixed temperature

$$\langle \|M(T)\| \rangle = \frac{1}{Z}\sum_{\mathbf{S}\in\Lambda} \|M(\mathbf{S})\| \exp[-H(\mathbf{S})/T]. \tag{5}$$

Similarly, with the help of the matrix $A_{ij} = (-1)^{i+j}$, we define the order parameter of the antiferromagnetic Ising model, the expectation value of the staggered magnetization. The latter is calculated from an elementwise product with a matrix form of the spin configurations

$$M_{\text{st}} = M(\underline{S} \odot A). \tag{6}$$

B. The XY model in three dimensions

The Mermin-Wagner-Hohenberg theorem [30,31] prohibits continuous phase transitions in $d \leqslant 2$ dimensions at finite temperature when all interactions are sufficiently short ranged. Hence, we choose the XY model in three dimensions as a model to probe the ability of a variational autoencoder to classify phases of models with continuous symmetries. The Hamiltonian of the XY model reads

$$H(\mathbf{S}) = -J\sum_{\langle i,j\rangle_{\text{NN}}} \mathbf{s}_i \cdot \mathbf{s}_j, \tag{7}$$

with spins on the one-sphere $\mathbf{s}_i \in \mathbb{R}^2$, $\|\mathbf{s}_i\| = 1$. Employing $J = 1$, the transition temperature of this model is $T_c = 2.2017$ [32]. Using a cubic lattice with $L = 14$ such that $N = L^3 = 2744$, we perform Monte Carlo simulations to create 10 000 independent sample spin configurations in the temperature range of $T \in [0,5]$. The order parameter is defined analogously to the Ising model magnetization (5), but with the L^2 norm of a magnetization consisting of two components.

III. METHODS

Principal component analysis [33] is an orthogonal linear transformation of the data to an ordered set of variables, sorted by their variance. The first variable, which has the largest variance, is called the first principal component, the variable with the second largest variance is the second principal component, and so on. The linear function $\langle\cdot,\mathbf{w}\rangle$, which maps a collection of spin samples $(\mathbf{S}_{(1)},\ldots,\mathbf{S}_{(n)})$ to its first principal component, is defined as

$$\arg\max_{\|\mathbf{w}\|=1}\left[\sum_j [(\mathbf{S}_{(j)} - \mu)\cdot\mathbf{w}]^2\right], \tag{8}$$

where μ is the vector of mean values of each spin averaged over the whole data set. Further principal components are obtained by subtracting the already calculated principal components and repeating Eq. (8) on the remaining subspace.

Kernel principal component analysis [34] projects the data into a kernel space in which the principal component analysis is then performed. In this work the nonlinearity is induced by a radial basis functions kernel.

Traditional neural-network-based autoencoders [35,36] consist of two artificial neural networks stacked on top of each other. They are created from an encoding artificial neural network, which outputs a latent representation of the input data, and a decoding neural network that tries to accurately reconstruct the input data from its latent representation (see Fig. 1). Very shallow versions of autoencoders can reproduce the results of principal component analysis [37]. The parameters of this algorithm are trained by performing gradient descent updates in order to minimize the reconstruction loss (reconstruction error) between input data and output data.

Variational autoencoders [38] are a modern version of autoencoders that impose additional constraints on the encoded representations, i.e., the latent variables in Fig. 1. These constraints transform the autoencoder to an algorithm that learns a latent variable model for its input data. Whereas the neural networks of traditional autoencoders learn an arbitrary function to encode and decode the input data, variational autoencoders learn the parameters of a probability distribution modeling the data. After learning the probability distribution, one can sample parameters from it and then let the encoder network generate samples closely resembling the training data. To achieve this, variational autoencoders employ the assumption that one can sample the input data from a unit Gaussian distribution of latent parameters. The weights of the model are trained by simultaneously optimizing two loss functions, a reconstruction loss and the Kullback-Leibler

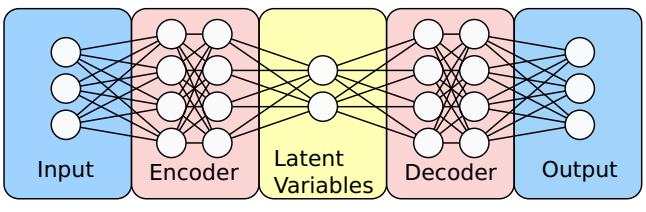

FIG. 1. Autoencoder neural network architecture. The encoder network translates the input to its latent representation, from which the decoder reconstructs an approximation of the input as output.

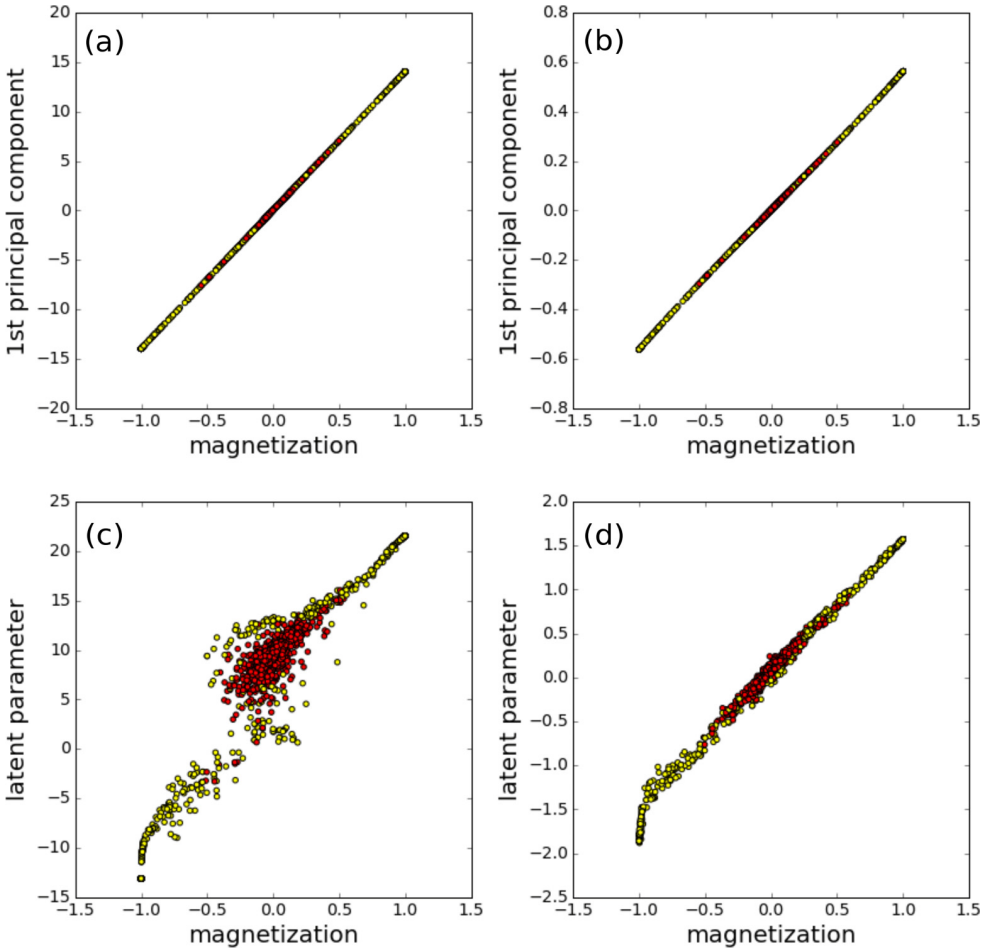

FIG. 2. Ferromagnetic Ising model showing the principal components and latent representations versus magnetization for different algorithms: (a) principal component analysis, (b) kernel principal component analysis, (c) autoencoder, and (d) variational autoencoder. Red points correspond to configurations of the unordered phase, while yellow points belong to the ordered phase.

divergence between the learned latent distribution and a prior unit Gaussian.

In this work we use autoencoders and variational autoencoders with one fully connected hidden layer in the encoder as well as one fully connected hidden layer in the decoder, each consisting of 256 neurons. The number of latent variables is chosen to match the model from which we sample the input data. The activation functions of the intermediate layers are rectified linear units. The activation function of the final layer is a sigmoid in order to predict probabilities of spin $\uparrow$ or $\downarrow$ in the Ising model or tanh for predicting continuous values of spin components in the XY model. We do not employ any L^1, L^2, or dropout regularization. However, we tune the relative weight of the two loss functions of the variational autoencoder to fit the problem at hand. The Kullback-Leibler divergence of the variational autoencoder can be regarded as regularization of the traditional autoencoder. In our autoencoder the reconstruction loss is the cross-entropy loss between the input and output probability of discrete spins, as in the Ising model. The reconstruction loss is the mean-square error between the input and the output data of continuous spin variables in the XY model.

To understand why a variational autoencoder can be a suitable choice for the task of classifying phases, we recall what happens during training. The weights of the autoencoder learn two things. On the one hand, they learn to encode the similarities of all samples to allow for an efficient reconstruction. On the other hand, they learn a latent distribution of the parameters that encode the most information possible to distinguish between different input samples. Let us translate these considerations to the physics of phase transitions. If all the training samples are in the unordered phase, the autoencoder learns the common structure of all samples. The autoencoder fails to learn any random entropy fluctuations, which are averaged out over all data points. However, in the ordered phase there exists a common order in samples belonging to the same phase. This common order translates to a nonzero latent parameter, which encodes correlations on each input sample. It turns out that in our cases this parameter is the order parameter corresponding to the broken symmetry. It is not necessary to find a perfect linear transformation between the order parameter and the latent parameter as is the case in Fig. 2. A one-to-one correspondence is sufficient, such that one is able to define a function that maps these parameters onto each other and captures all discontinuities of the derivatives of the order parameter.

Principal component analysis and autoencoders seem very different, but they share common characteristics.

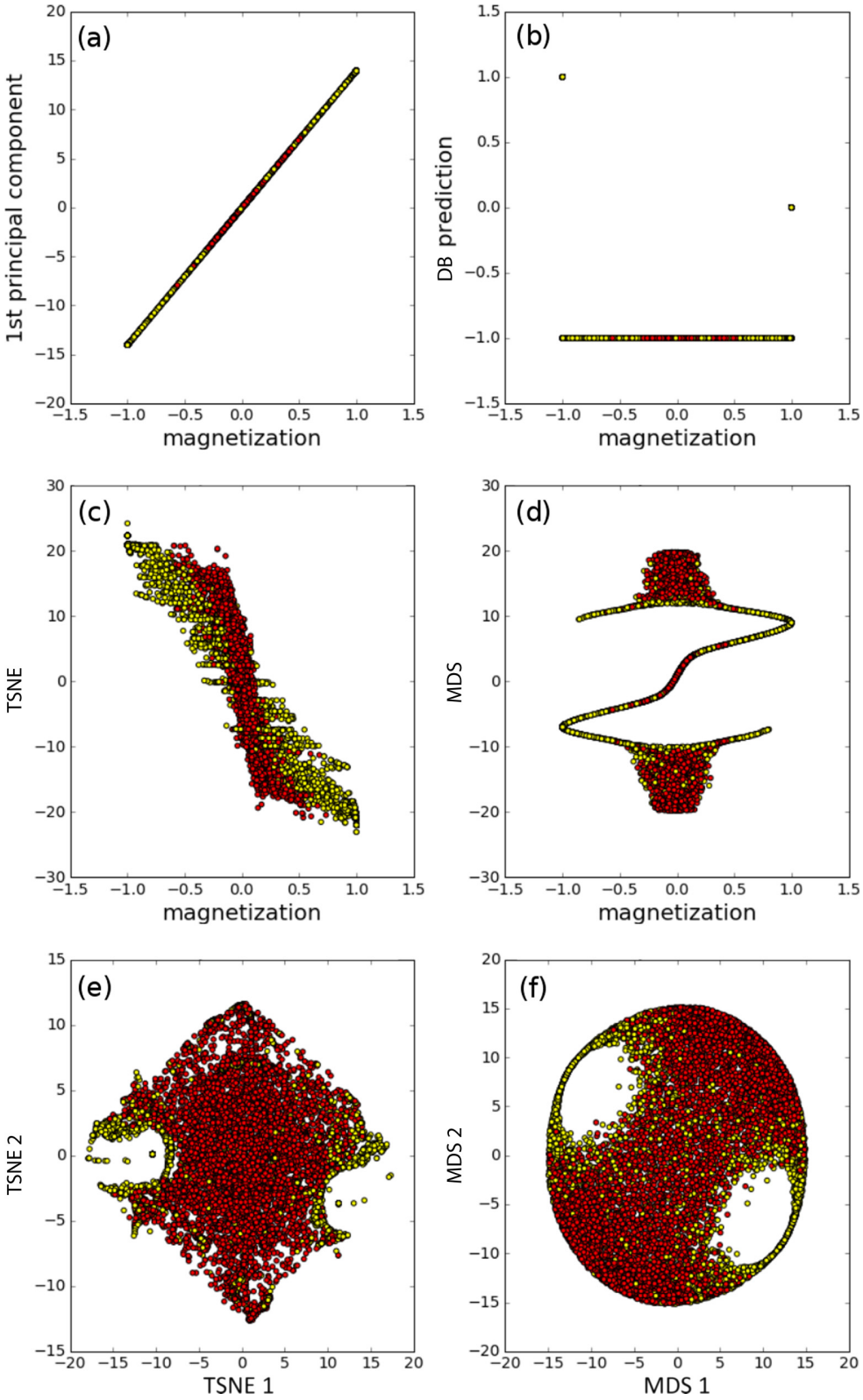

FIG. 3. Ferromagnetic Ising model showing a comparison of manifold and clustering methods: (a) PCA, (b) DBSCAN, (c) one-component TSNE, (d) one-component MDS, (e) two-component TSNE, and (f) two-component MDS.

Reconstructing the input data from its principal components minimizes the mean-square reconstruction error. Hence, a single-layer autoencoder with linear activation functions closely resembles principal component analysis [37]. Principal component analysis is much easier to apply and in general is characterized by fewer parameters than autoencoders. Autoencoders based on convolutional layers can have a reduced number of parameters. In extreme cases this number can be even less than the parameters of principal component analysis. Furthermore, such autoencoders can promote locality of features in the data.

We further examine manifold methods, where high-dimensional data are embedded in a low-dimensional manifold in which similar data points are represented close to each other. For this purpose we employ t-distributed stochastic neighbor embedding (TSNE) [39], a technique that is particularly sensitive to local structure. We also consider multidimensional scaling (MDS) [40], which seeks a low-dimensional representation of the data in which the distances respect the distances in the original high-dimensional space. In the following sections we embed the spin configurations into a one- (one-component TSNE or MDS) or two-dimensional manifold (two-component TSNE or MDS).

As an example of clustering methods we try to employ density-based spatial clustering of applications with noise (DBSCAN) [41]. It associates clusters with areas of high density separated by areas of low density. The DBSCAN can also find convex-shaped clusters.

For this work we employed the PYTHON libraries sklearn [42] and keras [43]. A detailed introduction to autoencoders can be found in [44].

IV. RESULTS

A. Ferromagnetic Ising model

First we compare the four most successful algorithms applied to the Ising model: principal component analysis (PCA), kernel principal component analysis, autoencoders, and variational autoencoders. They all share the characteristic that the first principal components or the latent parameters (Fig. 2) show a clear correlation to the magnetization. However, the traditional autoencoder fails to capture this correlation in the vicinity of magnetization $M(S) = 0$; this fact lets us favor variational autoencoders over traditional autoencoders. The principal component methods show the most accurate results, slightly better than the variational autoencoder. This is to be expected, since the former are modeled by fewer parameters.

In Fig. 3 we compare manifold and clustering methods to principal component analysis. All algorithms are employed on raw data in the temperature range $T \in [0,5]$. The PCA is the only method that successfully manages to approximate the magnetization. The DBSCAN completely fails to find clusters on raw data. Moreover, the one-component TSNE can separate clusters of positive and negative magnetization. One-component MDS finds a structure that separates the ordered phase (yellow) from the unordered phase (red). However, this structure cannot be interpreted in a physical sense. Two-component TSNE and two-component MDS both successfully place samples belonging to the same phases close to each other. They also distinguish between two sorts of magnetization. We expect that in the case of supervised learning, where one can cluster the training data, one would obtain much clearer results.

In the following results we concentrate on the variational autoencoder as the most promising and powerful algorithm for unsupervised learning.

As a starting point, we choose the number of latent parameters in the variational autoencoder to be one. After training the network for 50 epochs and observing a saturation of the training loss, we visualize the results in Fig. 4. In Fig. 4(a) we see a close linear correlation between the latent parameter and the magnetization. In Fig. 4(b) there is a histogram of spin configurations encoded into their latent parameter. The model learned to classify the configurations into three clusters. The identification of the latent parameter as a close approximation of the magnetization $M(S)$ allows

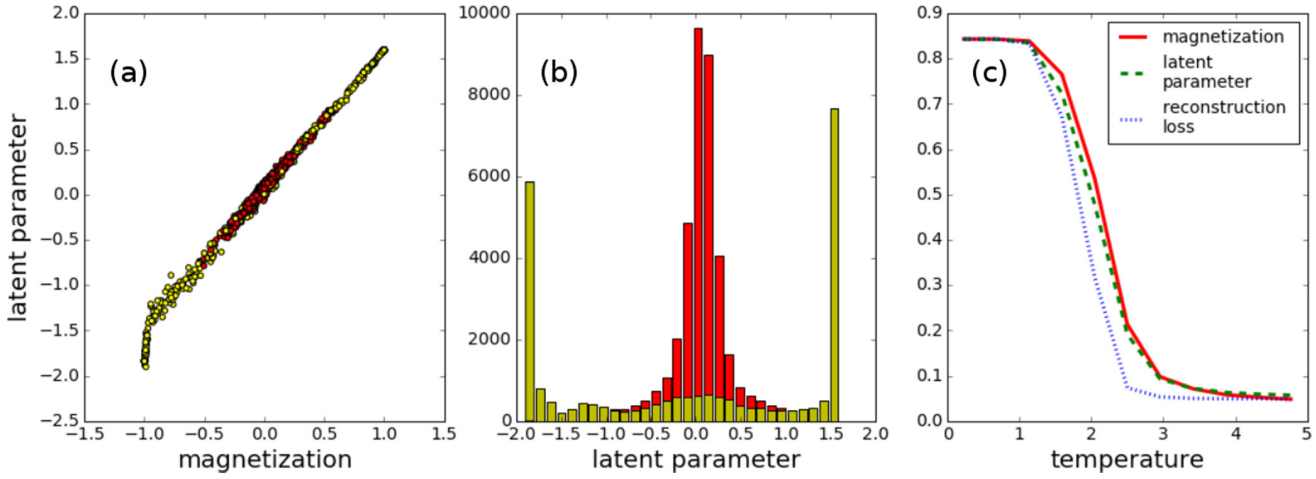

FIG. 4. Ferromagnetic Ising model. (a) The correlation between the latent parameter and magnetization is shown for each spin sample. Red dots indicate points in the unordered phase, while yellow dots correspond to the ordered phase. (b) The histogram counts occurrences of latent parameters. (c) One can see the average values at fixed temperature of the absolute value of magnetization, the absolute value of the latent parameter, and the cross-entropy reconstruction loss. The reconstruction loss is mapped on the $T = 0$ and 5 values of the magnetization and the latent parameter is rescaled to the magnetization at $T = 0$.

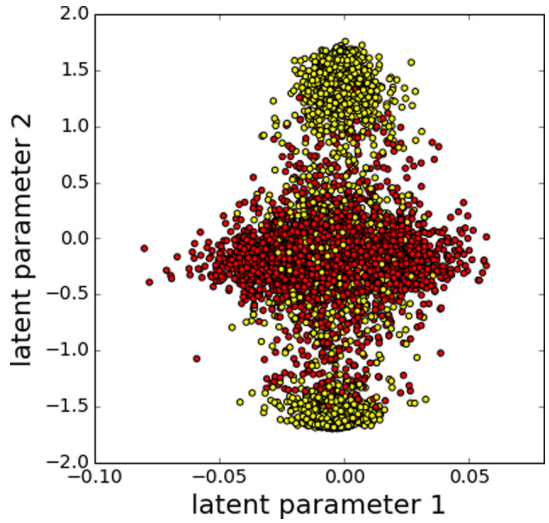

FIG. 5. Ferromagnetic Ising model with visualization of data in a two-dimensional latent space. Red dots indicate points in the unordered phase, while yellow dots correspond to the ordered phase. The axis for parameter 1 has a smaller range than the axis for parameter 2.

us to interpret the properties of the clusters. The right and left clusters in Fig. 4(b) correspond to an average magnetization of $M(S) \approx \pm 1$, while the middle cluster corresponds to the magnetization $M(S) \approx 0$. Employing a different viewpoint, from Fig. 4 we conclude that the parameter that holds the most information on how to distinguish Ising spin samples is the order parameter. In Fig. 4(c) the average of the magnetization, the latent parameter, and the reconstruction loss are shown as a function of the temperature. A sudden change in the magnetization at $T_c \approx 2.269$ defines the phase transition between paramagnetism and ferromagnetism. Even without previous knowledge of this order parameter, we can use the results of the autoencoder to infer the position of the phase transition. As an approximate order parameter, the average absolute value of latent parameter also shows a steep change at T_c. The averaged reconstruction loss also changes drastically at T_c. While the latent parameter is different for each physical model, the reconstruction loss can be used as a universal parameter to identify phase transitions. In conclusion, without any knowledge of the Ising model and its order parameter but sample configurations, we can find a good estimation for its order parameter and the occurrence of a phase transition.

It is *a priori* not clear how to determine the number of latent neurons in the creation of the neural network of the autoencoder. Due to the lack of theoretical groundwork, we find the optimal number by experimentation. If we expand the number of latent dimensions by one (see Fig. 5), the results of our analysis change only slightly. The second parameter contains much less information compared to the first, since it stays very close to zero. Hence, for the Ising model, one parameter is sufficient to store most of the information of the latent representation.

B. Antiferromagnetic Ising model

After having identified variational autoencoders as the most promising unsupervised learning algorithms to determined phase transitions, we concentrate on this algorithm in the following and present the results of the other algorithms in the Appendix.

While the ferromagnetic Ising model serves as an ideal starting ground, in the next step we are interested in models where different sites in the samples contribute in a different manner to the order parameter. We do this in order to show that our model is even sensitive to structure on the smallest scales. For the magnetization in the ferromagnetic Ising model, all spins contribute with the same weight. In contrast, in the antiferromagnetic Ising model, neighboring spins contribute with opposite weight to the order parameter (6).

Again the variational autoencoder manages to capture the traditional order parameter. The staggered magnetization is strongly correlated with the latent parameter (see Fig. 6). The three clusters in the latent representation make it possible to

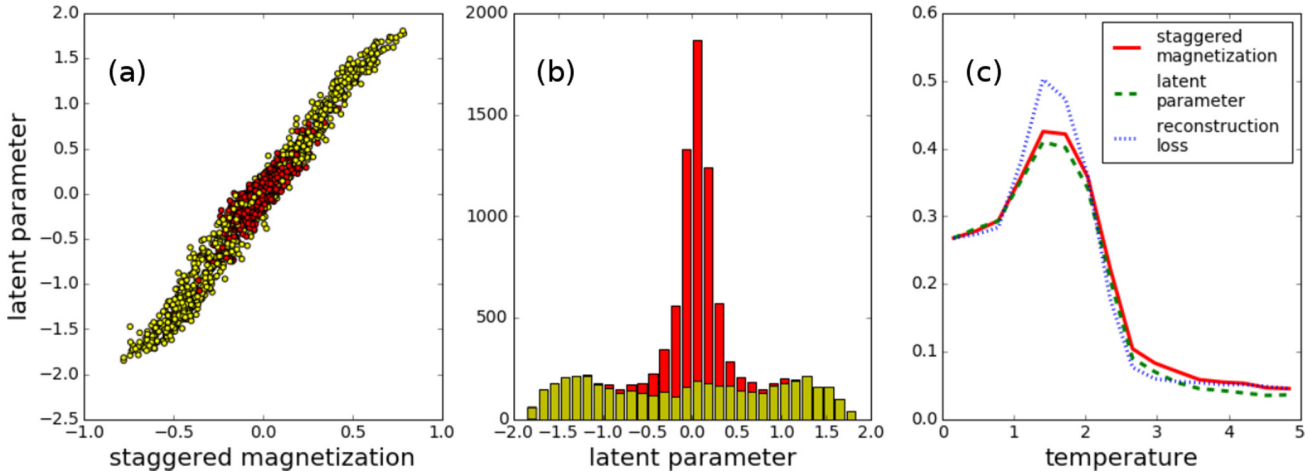

FIG. 6. Antiferromagnetic Ising model. (a) Correlation between the latent parameter and staggered magnetization for each spin sample. Red dots indicate points in the unordered phase, while yellow dots indicate points in the ordered phase. (b) The histogram counts occurrences of latent parameters. (c) Average at fixed temperature of the absolute value of staggered magnetization, the absolute value of the latent parameter, and the cross-entropy reconstruction loss. The reconstruction loss is mapped on the $T = 0$ and 5 values of the staggered magnetization and the latent parameter is rescaled to the magnetization at $T = 0$.

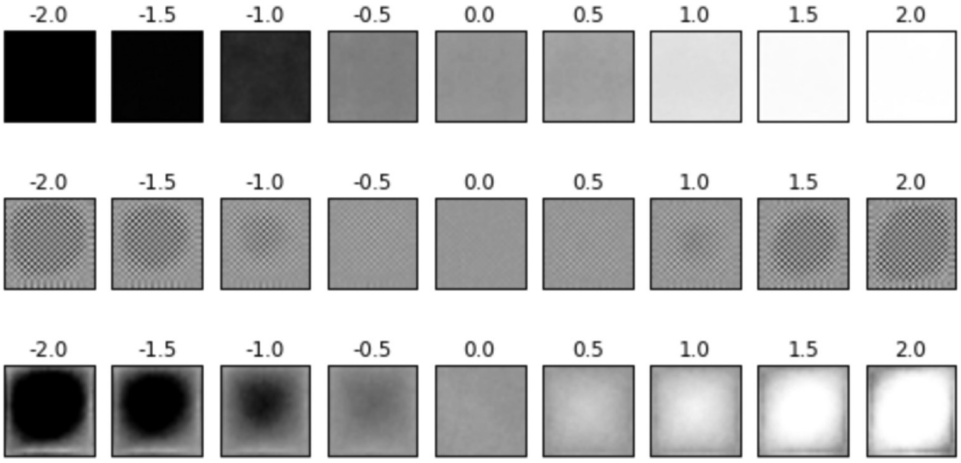

FIG. 7. Reconstruction of images, each consisting of 28×28 pixels, from the latent parameter. The brightness indicates the probability of the spin to be up [white indicates $p(\uparrow) = 1$ and black $p(\downarrow) = 1$]. The first row is a reconstruction of sample configurations from the ferromagnetic Ising model. The second row corresponds to the antiferromagnetic Ising model. The third row is the prediction from the AF latent parameter, where each second spin is multiplied by -1, to show that the second row indeed predicts an antiferromagnetic state.

interpret different phases. Furthermore, we note that all three averaged quantities, the magnetization, the latent parameter, and the reconstruction loss, can serve as indicators of a phase transition.

Figure 7 demonstrates the reconstruction from the latent parameter. In the first row we see the reconstruction from samples of the ferromagnetic Ising model; the latent parameter encodes the whole spin order in the ordered phase. Reconstructions from the antiferromagnetic Ising model are shown in the second and third rows. Since the reconstructions clearly show an antiferromagnetic phase, we infer that the autoencoder encodes the spin samples even to the most microscopic level.

C. The XY model

In the XY model we examine the capabilities of a variational autoencoder to encode models with continuous symmetries. The application of other algorithms is compared in the Appendix. In models like the Ising model, where discrete symmetries are present, the autoencoder only needs to learn a discrete set, which is often finite, of possible representations of the symmetry-broken phase. If a continuous symmetry is broken, there are infinitely many possibilities of how the ordered phase can be realized. Hence, in this section we test the ability of the autoencoder to embed all these different realizations into latent variables.

The variational autoencoder handles this model equally well as the Ising model. We find that two latent parameters model the phase transition best. The latent representation in Fig. 8(b) shows the distribution of various states around a central cluster. The radial symmetry in this distribution leads to the assumption that a sensible order parameter is constructed from the L^2 norm of the latent parameter vector. In Fig. 8 one sees the correlation between the magnetization and the

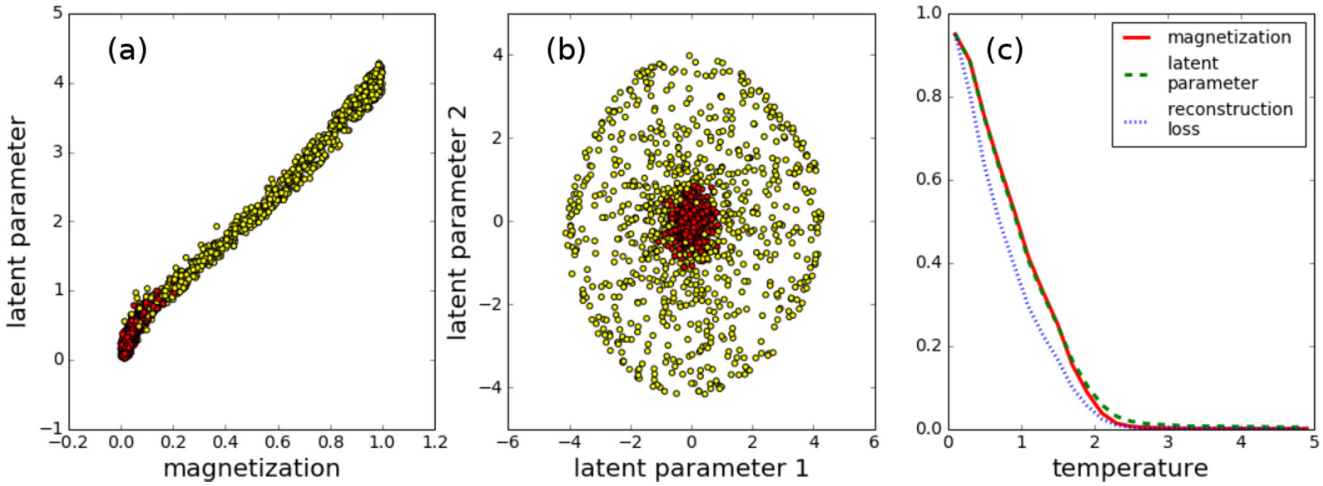

FIG. 8. The XY model. (a) Correlation between the L^2 norm of the latent parameter vector and the L^2 norm of the magnetization for each spin sample. Red dots indicate points in the unordered phase, while yellow dots indicate points in the ordered phase. (b) Representation of the spin configurations in two-dimensional latent space. (c) For each L^2 norm of the magnetization, the L^2 norm of latent parameter, and the average of the square root of the mean-square error reconstruction loss, we plot the average at fixed temperature. The reconstruction loss is mapped on the $T = 0$ and 5 values of the magnetization and the latent parameter is rescaled to the magnetization at $T = 0$.

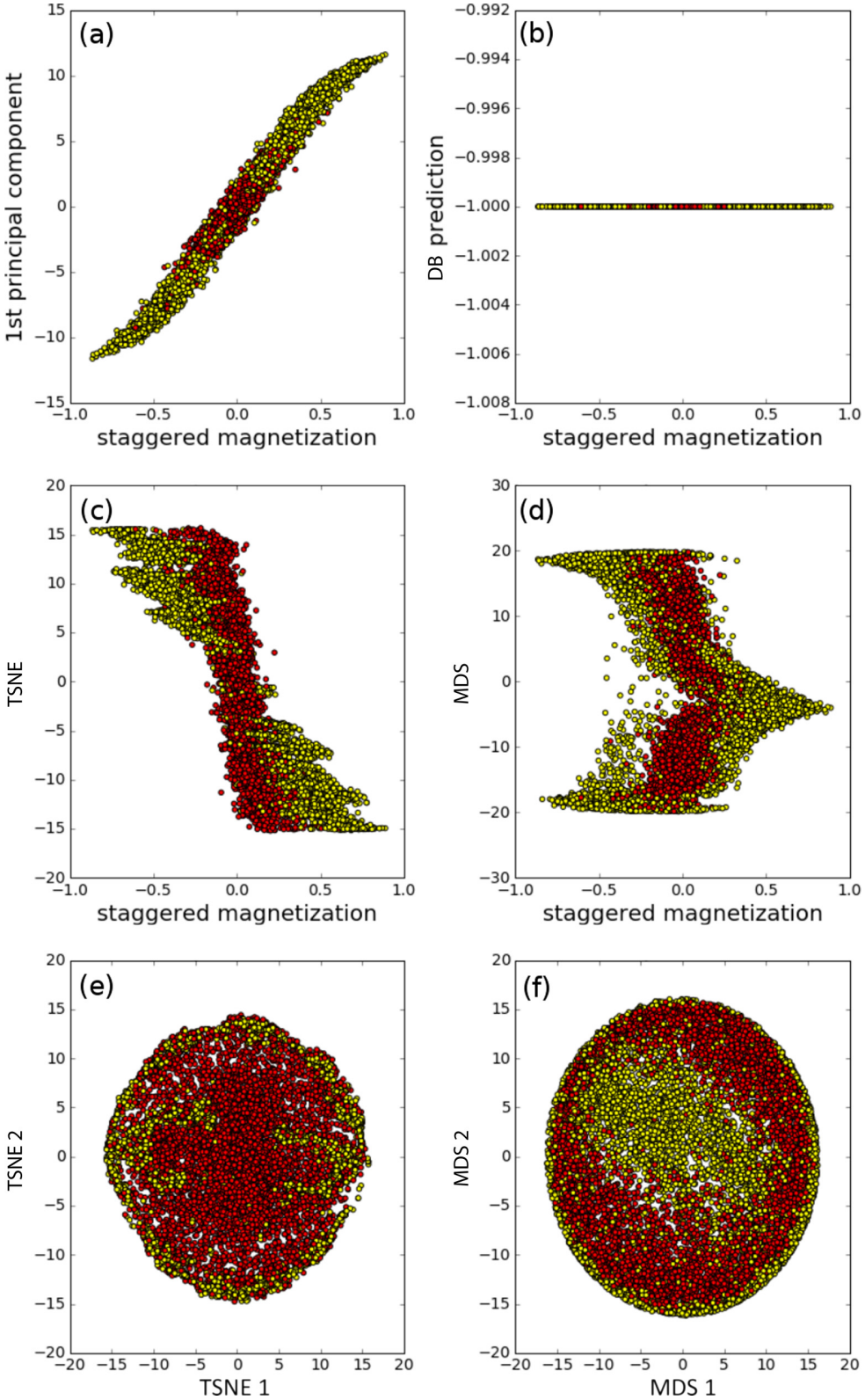

FIG. 9. Antiferromagnetic Ising model showing a comparison of the manifold and clustering methods: (a) PCA, (b) DBSCAN, (c) one-component TSNE, (d) one-component MDS, (e) two-component TSNE, and (f) two-component MDS.

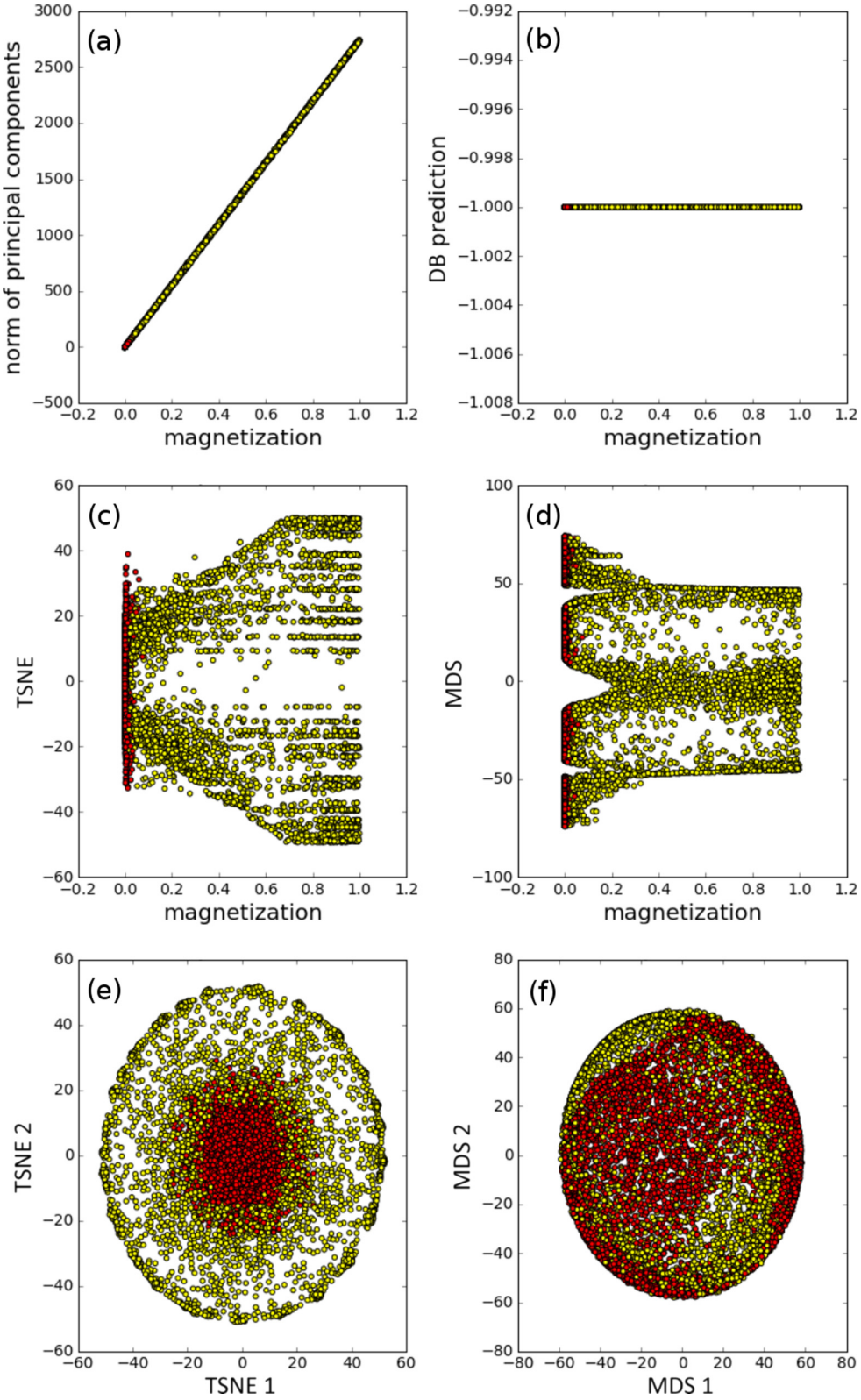

FIG. 10. The XY model showing a comparison of the manifold and clustering methods: (a) PCA, (b) DBSCAN, (c) one-component TSNE, (d) one-component MDS, (e) two-component TSNE, and (f) two-component MDS.

absolute value of the latent parameter vector. Averaging the samples for the same temperature hints at the fact that the latent parameter and the reconstruction loss can serve as an indicator for the phase transition.

V. CONCLUSION

We have shown that it is possible to observe phase transitions using unsupervised learning. We compared different unsupervised learning algorithms and found that principal component analysis and variational autoencoders are the best algorithms for examining phase transitions. We were motivated by the need for an upgrade of the traditional autoencoder to a variational autoencoder. The weights and latent parameters of the variational autoencoder are able to store information about microscopic and macroscopic properties of the underlying systems. The most distinguished latent parameters coincide with the known order parameters. Furthermore, we have established the reconstruction loss as a universal indicator for phase transitions. We expanded the toolbox of unsupervised learning algorithms in physics by powerful methods, most notably the variational autoencoder, which can handle nonlinear features in the data and scale very well to huge data sets. In the future one may employ autoencoders to capture phase transitions with nonlinear order parameters, which PCA cannot reproduce. These theories include lattice gauge theories, where the order parameters are defined by a loop along several lattice sites. We expect the prediction of unseen phases or exposure of unknown order parameters, e.g., in quantum spin liquids. We look forward to the development of deep convolutional autoencoders that have a reduced number of parameters compared to fully connected autoencoders to probe locality in feature selection. Furthermore, since there exists a connection between deep neural networks and renormalization group [45], it may be helpful to employ deep convolutional autoencoders to further expose this connection.

ACKNOWLEDGMENTS

The author would like to thank Timo Milbich, Björn Ommer, Michael Scherer, Manuel Scherzer, and Christof Wetterich for useful discussions as well as Mathias Neidig and Shirin Nkongolo for helpful suggestions. S.W. acknowledges support from the Heidelberg Graduate School of Fundamental Physics.

APPENDIX: PCA, MANIFOLD, AND CLUSTERING RESULTS

1. Antiferromagnetic Ising model

In Fig. 9 we compare different unsupervised learning algorithms applied to the antiferromagnetic Ising model. Principal component analysis approximates the magnetization by its first principal component equally well as variational autoencoders (Fig. 6). The DBSCAN fails to find any meaningful structure. One-component TSNE finds two clusters of staggered magnetizations. One-component MDS is able to discover structure in the data, however we cannot relate it to any physical quantity. Two-component TSNE projects the ordered phase (yellow) to the outside of the manifold, while the unordered phase (red) is mapped to the inside. It also finds substructure in the unordered phase. Two-component MDS is able to separate points belonging to the unordered phase from points belonging to the ordered phase.

2. The XY model

Figure 10 compares the results of different unsupervised learning algorithms applied to raw Monte Carlo simulations of the XY model. The Euclidean norm of the first two principal components analysis is perfectly correlated with the magnetization. Principal component analysis thus yields similar results as variational autoencoders (Fig. 8). One-component TSNE (one-component MDS) finds a structure where two (three) points in the embedded manifold correspond to a finite magnetization. The most promising result is achieved by two-component TSNE: Spin configurations belonging to the unordered phase are centered at the origin, while configurations belonging to the ordered phase are mapped to the exterior. This result is similar to the clustering from the variational autoencoder (see Fig. 8). Two-component MDS finds two clusters of a higher density of data points belonging to the ordered phase; the result is less pronounced compared to TSNE.

[1] L. Landau, Phys. Z. Sowjetunion **11**, 26 (1937).
[2] L. P. Kadanoff, Physics **2**, 263 (1966).
[3] K. G. Wilson, Rev. Mod. Phys. **47**, 773 (1975).
[4] X.-G. Wen, *Quantum Field Theory of Many-Body Systems* (Oxford University Press, Oxford, 2004).
[5] C.-C. Kaun and T. Seideman, Phys. Rev. Lett. **94**, 226801 (2005).
[6] P. Anderson, Mater. Res. Bull. **8**, 153 (1973).
[7] G. Cybenko, Math. Control Signals Syst. **2**, 303 (1989).
[8] K. Hornik, Neural Networks **4**, 251 (1991).
[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, in *Advances in Neural Information Processing Systems 25*, edited by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Curran, Red Hook, 2012), pp. 1097–1105.
[10] G. Hinton *et al.*, IEEE Signal Process. Mag. **29**, 82 (2012).
[11] Y. Goldberg, arXiv:1510.00726.
[12] S. Curtarolo, D. Morgan, K. Persson, J. Rodgers, and G. Ceder, Phys. Rev. Lett. **91**, 135503 (2003).
[13] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, Phys. Rev. Lett. **108**, 058301 (2012).
[14] Z. Li, J. R. Kermode, and A. D. Vita, Phys. Rev. Lett. **114**, 096405 (2015).
[15] E. LeDell, Prabhat, D. Y. Zubarev, B. Austin, and W. A. Lester, Jr., J. Math. Chem. **50**, 2043 (2012).
[16] G. Pilania, J. E. Gubernatis, and T. Lookman, Phys. Rev. B **91**, 214302 (2015).
[17] Y. Saad, D. Gao, T. Ngo, S. Bobbitt, J. R. Chelikowsky, and W. Andreoni, Phys. Rev. B **85**, 104104 (2012).

[18] O. S. Ovchinnikov, S. Jesse, P. Bintacchit, S. Trolier-McKinstry, and S. V. Kalinin, Phys. Rev. Lett. **103**, 157203 (2009).
[19] L.-F. Arsenault, A. Lopez-Bezanilla, O. A. von Lilienfeld, and A. J. Millis, Phys. Rev. B **90**, 155136 (2014).
[20] J. C. Snyder, M. Rupp, K. Hansen, K.-R. Müller, and K. Burke, Phys. Rev. Lett. **108**, 253002 (2012).
[21] G. Hautier, C. C. Fischer, A. Jain, T. Mueller, and G. Ceder, Chem. Mater. **22**, 3762 (2010).
[22] J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).
[23] G. Kasieczka, T. Plehn, M. Russell, and T. Schell, JHEP 05 (2017) 006.
[24] G. Carleo and M. Troyer, Science **355**, 602 (2017).
[25] G. Torlai and R. G. Melko, Phys. Rev. B **94**, 165134 (2016).
[26] L. Wang, Phys. Rev. B **94**, 195105 (2016).
[27] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nat. Phys. **13**, 435 (2017).
[28] L. Onsager, Phys. Rev. **65**, 117 (1944).
[29] N. Metropolis and S. Ulam, J. Am. Stat. Assoc. **44**, 335 (1949).
[30] N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).
[31] P. C. Hohenberg, Phys. Rev. **158**, 383 (1967).
[32] A. P. Gottlob and M. Hasenbusch, Physica A **201**, 593 (1993).
[33] K. Pearson, Philos. Mag. Series 6 **2**, 559 (1901).
[34] B. Scholkopf, A. Smola, and K.-R. Müller, *Advances in Kernel Methods* (MIT Press, Cambridge, 1999), pp. 327–352.
[35] H. Bourlard and Y. Kamp, Biol. Cybern. **59**, 291 (1988).
[36] G. E. Hinton and R. S. Zemel, in Advances in *Neural Information Processing 6*, edited by J. D. Cowan, G. Tesauro, and J. Alspector (Kaufmann, San Mateo, 1994).
[37] P. Baldi and K. Hornik, Neural Networks **2**, 53 (1989).
[38] D. P. Kingma and M. Welling, arXiv:1312.6114.
[39] L. van der Maaten and G. Hinton, J. Mach. Learn. Res. **9**, 2579 (2008).
[40] J. Kruskal, Psychometrika **29**, 115 (1964).
[41] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, edited by E. Simoudis, J. Han, and U. Fayyad (AAAI, Palo Alto, 1996), pp. 226–231.
[42] F. Pedregosa *et al.*, J. Mach. Learn. Res. **12**, 2825 (2011).
[43] F. Chollet, keras, 2015, https://github.com/fchollet/keras.
[44] F. Chollet, Building autoencoders in keras, 2014, https://blog.keras.io/building-autoencoders-in-keras.html.
[45] P. Mehta and D. J. Schwab, arXiv:1410.3831.

5 Machine Learning of Explicit Order Parameters: From the Ising Model to SU(2) Lattice Gauge Theory

Wetzel, Sebastian J.; Scherzer Manuel

Phys. Rev. B 96, 184410, Published 8 November 2017 [3]

PHYSICAL REVIEW B **96**, 184410 (2017)

Machine learning of explicit order parameters: From the Ising model to SU(2) lattice gauge theory

Sebastian J. Wetzel and Manuel Scherzer
Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany
(Received 27 May 2017; revised manuscript received 27 October 2017; published 8 November 2017)

We present a solution to the problem of interpreting neural networks classifying phases of matter. We devise a procedure for reconstructing the decision function of an artificial neural network as a simple function of the input, provided the decision function is sufficiently symmetric. In this case one can easily deduce the quantity by which the neural network classifies the input. The method is applied to the Ising model and SU(2) lattice gauge theory. In both systems we deduce the explicit expressions of the order parameters from the decision functions of the neural networks. We assume no prior knowledge about the Hamiltonian or the order parameters except Monte Carlo–sampled configurations.

DOI: 10.1103/PhysRevB.96.184410

I. INTRODUCTION

Machine learning enables computers to learn from experience and generalize their gained knowledge to previously unseen problems. The development of better hardware and algorithms, most notably artificial neural networks, propelled machine learning to one of the most transformative disciplines of this century. Nowadays such algorithms are used to classify images [1], to recognize language [2] or to beat humans in complex games [3]. Recently, machine learning has even been successfully employed to tackle highly complex problems in physics [4–15] and in turn physics has also inspired developments in machine learning [16–24]. It is now possible to classify phases of matter in the context of supervised learning [25–31] only from Monte Carlo samples. Phases can also be found without any information about their existence by unsupervised learning [32–35].

These algorithms suffer from a huge drawback: there is no comprehensive theoretical understanding of what they actually learn [36–40]. Without knowing if the neural networks base their decision on physical quantities one has no reason to trust the results if applied to an unknown system. Previous works suggest that machine learning discriminates phases of the Ising model by the order parameter [25,41]; others use the weights of the neural network to formulate a new order parameter [30].

In this paper (a) we propose a method to fully interpret neural networks, provided their decision function is sufficiently symmetric, (b) we explain this method at the Ising model and demonstrate its power at SU(2) gauge at theory, (c) we thereby justify the use of neural networks to classify phases, (d) this method yields such a clear interpretation that it can be used to determine the nature of the ordered phase.

To this end we introduce the correlation probing neural network. It can reduce the complexity of sufficiently symmetric decision functions. Since physical quantities are typically highly symmetric, this network is ideal for probing whether a physical quantity is responsible for the learned decision function. After reducing the complexity, we show that it is possible to fully reconstruct the explicit mathematical expression of the decision function in a simple form. From this expression one can extract the quantities by which the neural network distinguishes between phases.

This procedure is introduced at the Ising model, where we show that neural networks at distinguish between phases by the expected energy per spin (dominant) and the magnetization (subleading). We apply our method to SU(2) lattice gauge theory, where we find that the decision function is based on a nonlocal order parameter, the Polyakov loop.

II. ARTIFICIAL NEURAL NETWORKS

In this work we employ feed-forward artificial neural networks as a tool to distinguish between two classes in the context of supervised learning. After being successfully trained, the algorithm is able to predict the class of unseen test samples with high accuracy. We consider a neural network as an approximation of the decision function D. The decision function assigns to each sample S a probability $P \in [0,1]$ to be in class 1. The decision boundary is a hyperplane in the space of the parameters of sample configurations defined by $D(S) = 0.5$, where the neural network is most unsure about the correct label. If there exists an explicit quantity $Q(S)$ which is learned by the neural network, and which is responsible for the distinction between phases, we expect that a change in the quantity Q is always related to a change in the prediction probability, hence $\nabla Q || \nabla D$ in the vicinity of the decision boundary. In our neural networks the output can be written as $D(S) = \text{sigmoid}(\xi(S))$, where $\text{sigmoid}(x) = 1/[1 + \exp(-x)]$ maps the latent prediction $\xi(S)$ to a probability. It follows that $\nabla Q || \nabla \xi$ and thus Q can be expressed as a linear function of ξ in a linearized regime close to the decision boundary $\xi(S) = w\, Q(S) + b$. The decision function of neural networks is encoded in a highly elusive and highly nonlinear way. In order to decode the decision function, we present a type of neural network that is tailored to probe if specific correlations between different variables contribute to the decision function of the neural network. We call it the *correlation probing neural network*; see Fig. 1. The neural network architecture can be found in Appendix B.

The idea is to construct a tunable neural network which is able to interpolate between a traditional feed-forward neural network in one limit and an optimal minimal neural network, that still yields a similar classification performance, in the other limit. A neural network is an algorithm that excels in identifying hierarchical structure on data. These hierarchical functions can in principle be decomposed into simpler subfunctions. To this end the correlation probing neural network is decomposed into subnetworks of which each can only learn a specific function. The subfunctions are unique

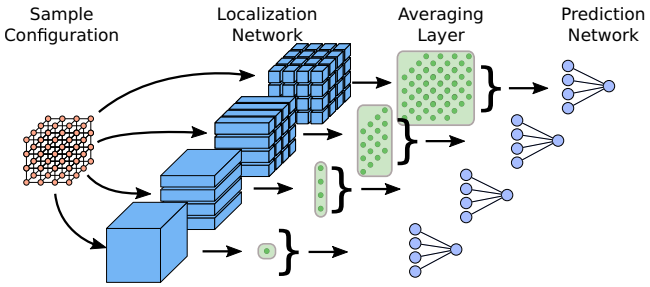

FIG. 1. The correlation probing neural network consists of three types of neural networks stacked on top of each other. The localization network is a fully convolutional neural network which prohibits connections outside of the receptive field of each output neuron and thus only recognizes correlations in the receptive field. The averaging layer averages over the input from the localization network, similarly to how the magnetization averages over all spins. The prediction network is a fully connected neural network, which transforms the output of the averaging layer to a prediction probability.

up to a linear transformation. The procedure for finding the optimal minimal neural network is to reduce the capacity of each of these subnetworks in an ordered manner until the neural network experiences a significant drop in the classification performance. The decision function of the optimal minimal neural network can then be written in a simplified form. If the quantity by which the neural network classifies the input is highly symmetric it is often possible to read off the quantity from the decision function. This is the case in the classification of phases in many physical systems.

III. ISING MODEL

The Hamiltonian of the ferromagnetic nearest-neighbor Ising model on the square lattice with vanishing external magnetic field is

$$H(S) = -J \sum_{\langle i,j\rangle_{nn}} s_i s_j, \tag{1}$$

with $J = 1$, $S = (s_1, \ldots, s_N)$ denotes a spin configuration, where $s_i \in \{1, -1\}$. It is a simple, well studied, and exactly solvable model from statistical physics that undergoes a second-order phase transition at $T_c = 2/[k_B \ln(1+\sqrt{2})]$ [42]. At T_c the specific heat $C_V = \partial\langle E\rangle/\partial\beta$ diverges, as does the temperature derivative of the expectation value of the absolute value of the magnetization $\langle M\rangle$, where $M(S) = |1/N \sum_i s_i|$.

The existence of different phases in the Ising model in the low- and high-temperature limit is known from unsupervised learning [32,33,35]. Using this knowledge we train the correlation probing neural network to predict phases of Monte Carlo–sampled configurations of size 28×28 below $T = 1.6$ in the ordered phase and above $T = 2.9$ in the unordered phase. More information about Monte Carlo simulation can be found in Appendix A. Using the full receptive field of 28×28, we allow the neural network to learn all possible spin correlations to approximate its decision function. In this case, the correlation probing network is equivalent to a standard convolutional neural network. Training and validation losses close to zero indicate that the neural network has found all necessary information it needs to reliably classify the phases. By successively lowering the receptive field size, we do not observe a drop in performance, except from 1×2 to 1×1 and from 1×1 to the baseline classifier; see Table I. In each of these steps the neural network loses important information about the samples. In Fig. 2(c) we can see the average classification probability, as a function of the temperature, of both networks. The phase-transition temperature can be found where $P = 0.5$. This is at $T = 2.5 \pm 0.5$ for the 1×1 network and $T = 2.25 \pm 0.25$ for the 1×2 network. An accurate estimation can be found in [25]. We however focus on examining what information got lost while lowering the receptive field size.

TABLE I. Ising model: Losses of neural networks with different receptive fields of the neurons in the localization network. Smaller numbers mean better performance. The baseline classifier is a random classifier which predicts each phase with a probability of $p = 0.5$.

Receptive field size	Train loss	Validation loss
28×28	6.1588×10^{-4}	0.0232
$\mathbf{1 \times 2}$	$\mathbf{1.2559 \times 10^{-4}}$	$\mathbf{1.2105 \times 10^{-7}}$
$\mathbf{1 \times 1}$	**0.2015**	**0.1886**
Baseline	0.6931	0.6931

By construction, the decision function D of the 1×1 neural network can be expressed as

$$D(S) = F\left(\frac{1}{N}\sum_i f(s_i)\right) = \text{sigmoid}\left[\xi\left(\frac{1}{N}\sum_i f(s_i)\right)\right], \tag{2}$$

where F is the function approximated by the prediction network and f is the function approximated by the localization network. The function f can be Taylor expanded:

$$f(s_i) = f_0 + f_1\, s_i + f_2 \underbrace{s_i^2}_{1} + f_3 \underbrace{s_i^3}_{s_i} + \cdots. \tag{3}$$

Since $s_i^2 = 1$, all higher-order terms can be neglected. The constants f_0 and f_1 can be absorbed by the bias and the weights of the prediction network approximating F. Thus, the decision function reduces to

$$D(S) = F\left(\frac{1}{N}\sum_i s_i\right). \tag{4}$$

In order to determine the function F, we compare the latent prediction ξ of the neural network, with the argument of F: $1/N \sum_i s_i$, in the vicinity of the decision boundary; see Fig. 2(a). This knowledge allows us to construct the decision function

$$D(S) \approx \text{sigmoid}\left(w\left|\frac{1}{N}\sum_i s_i\right| + b\right), \tag{5}$$

with weight w and bias b of the prediction neuron. The perfect correlation between the latent prediction $\xi(S)$ and $|1/N \sum_i s_i|$ further reinforces that our above deduction was correct. Until this point we have not used any information about the Ising model except Monte Carlo configurations. We have found that the decision function determines the phase by the quantity $Q(S) = |1/N \sum_i s_i|$. This function is the magnetization.

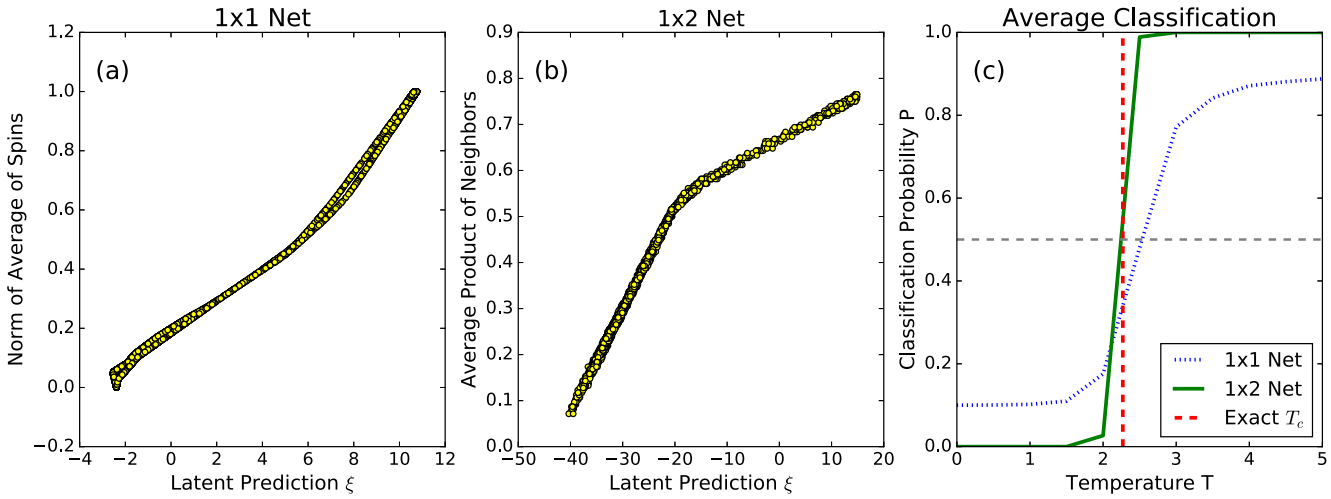

FIG. 2. Results of the correlation probing neural network applied to the Ising model. The latent prediction ξ is the argument of the sigmoid function in the last layer of the prediction network. (a) The latent prediction is perfectly correlated with the absolute value of the average of spins, i.e., the magnetization in the 1×1 network, for all sampled configurations. (b) The latent prediction of the 1×2 network is perfectly correlated with the average product of neighbors, i.e., the expected energy per site.

By examining the 1×2 network, we require by construction that the decision function is of the form

$$D(S) = F\left(\frac{1}{N}\sum_{\langle i,j\rangle_T} f(s_i,s_j)\right). \tag{6}$$

Here the sum only goes over transversal nearest neighbors, collecting each spin only once. The Taylor expansion,

$$\begin{aligned} f(s_i,s_j) =& f_{0,0} + f_{1,0}\, s_i + f_{0,1}\, s_j \\ &+ f_{2,0}\, s_i^2 + f_{1,1}\, s_i\, s_j + f_{0,2}\, s_j^2 + \cdots, \end{aligned} \tag{7}$$

contains only three terms of note; all other terms can be reduced to simpler ones by using $s_i^2 = 1$. The terms $f_{1,0}\, s_i$ and $f_{0,1}\, s_j$ represent the magnetization. From Table I and the analysis of the 1×1 network, we know that these terms contain less information than the quantity we are looking for. So the leading term must be $f_{1,1} s_i s_j$. Thus, the decision function can be written as

$$D(S) \approx F\left(\frac{1}{N}\sum_{\langle i,j\rangle_T} s_i s_j\right). \tag{8}$$

In Fig. 2(b) we see the perfect correlation between the latent prediction $\xi(S)$ and $1/N\sum_{\langle i,j\rangle_T} s_i s_j$. This also means that the correction from the subleading terms $f_{1,0}\, s_i$ and $f_{0,1}\, s_j$ is indeed negligible. Hence, we end up with the decision function

$$D(S) \approx \text{sigmoid}\left[w\left(\frac{1}{N}\sum_{\langle i,j\rangle_T} s_i s_j\right) + b\right]. \tag{9}$$

By translational and rotational symmetry, the sum can be generalized to all neighbors $Q(S) = \frac{1}{N}\sum_{\langle i,j\rangle_{nn}} s_i s_j$. This quantity is, up to a minus sign, the average energy per spin site. It is worth noting that the energy per site can be used to distinguish between phases more reliably than the magnetization; see Table I.

IV. SU(2) LATTICE GAUGE THEORY

We examine SU(2) lattice gauge theory, which shows confinement, one of the most distinct features of QCD. It builds on the idea of discretizing the Euclidean path integral of SU(2) Yang-Mills theory. Lattice configurations are defined by a set of link variables $U_\mu^x \in \text{SU}(2)$. Each matrix connects two sites on a four-dimensional $x \in N_\tau \times N_s^3$ space-time lattice with $N_\tau = 2$ (temporal direction) and $N_s = 8$ (spatial volume). The direction is indicated by $\mu \in \{\tau, x, y, z\}$. A sample lattice configuration collects all link variables on the lattice $S = (\{U_\mu^x\})$. Each U_μ^x is parametrized by four real parameters,

$$U_\mu^x = a_\mu^x \mathbb{1} + i\left(b_\mu^x \sigma_1 + c_\mu^x \sigma_2 + d_\mu^x \sigma_3\right), \tag{10}$$

where σ_i are the Pauli matrices; the coefficients obey $(a_\mu^x)^2 + (b_\mu^x)^2 + (c_\mu^x)^2 + (d_\mu^x)^2 = 1$. The trace of U_μ^x is given by $2\, a_\mu^x$, since the Pauli matrices are traceless. We employ the lattice version of the Yang-Mills action, the Wilson action [43],

$$S_{\text{Wilson}}[U] = \beta_{\text{latt}} \sum_x \sum_{\mu<\nu} \text{Re tr}\left(\mathbb{1} - U_{\mu\nu}^x\right), \tag{11}$$

where β_{latt} is the lattice coupling. Here $U_{\mu\nu}^x = U_\mu^x U_\nu^{x+\hat{\mu}} U_{-\mu}^{x+\hat{\mu}+\hat{\nu}} U_{-\nu}^{x+\hat{\nu}}$ is the smallest possible closed rectangular loop. The order parameter for the deconfinement

TABLE II. SU(2): Losses of neural networks with different receptive fields of the neurons in the localization network (* no hidden layers in the prediction net).

Receptive field size	Train loss	Validation loss
$2 \times 8 \times 8 \times 8$	1.0004×10^{-4}	2.6266×10^{-4}
$\mathbf{2 \times 1 \times 1 \times 1}$	$\mathbf{8.8104} \times 10^{-8}$	$\mathbf{6.8276} \times 10^{-8}$
$\mathbf{2 \times 1 \times 1 \times 1}^*$	$\mathbf{2.2292} \times 10^{-7}$	$\mathbf{4.2958} \times 10^{-7}$
$1 \times 1 \times 1 \times 1$	0.6620	0.9482
Baseline	0.6931	0.6931

phase transition is the expectation value of the Polyakov loop

$$L(\vec{x}) = \mathrm{tr}\left(\prod_{x_0=0}^{N_\tau-1} U_\tau^x\right) \overset{N_\tau=2}{=} \mathrm{tr}\big(U_\tau^{0,\vec{x}} U_\tau^{1,\vec{x}}\big)$$
$$= 2\big(a_\tau^{0,\vec{x}} a_\tau^{1,\vec{x}} - b_\tau^{0,\vec{x}} b_\tau^{1,\vec{x}} - c_\tau^{0,\vec{x}} c_\tau^{1,\vec{x}} - d_\tau^{0,\vec{x}} d_\tau^{1,\vec{x}}\big). \quad (12)$$

It is the trace of a closed loop that winds around time direction using periodic boundary conditions. The expectation value of the Polyakov loop is zero in the confined phase and finite in the deconfined phase. More details on the simulations can be found in Appendix A.

The existence of different phases in SU(2) lattice gauge theory can be found by unsupervised learning; see Appendix D. This knowledge allows us to train the correlation probing neural networks with different receptive fields, to classify phases on Monte Carlo–sampled configurations at lattice coupling $\beta \in [1,1.2]$ in one phase and $\beta \in [3.3,3.5]$ in the other phase. We test the neural network in $\beta \in [1.3,3.2]$ to predict a phase transition at $\beta = 1.99 \pm 0.10$ ($2\times1\times1\times1$ network) and $\beta = 1.97 \pm 0.10$ ($2\times8\times8\times8$ network); see Fig. 3(c). Our direct lattice calculation reveals $\beta = 1.880 \pm 0.025$. By successively lowering the receptive field size we lose important information for classifying phases below $2\times1\times1\times1$; see Table II. This means that crucial information about the phase transition is contained in this specific structure.

The decision function of the $2\times1\times1\times1$ network is constrained to

$$D(S) = F\left(\frac{2}{N}\sum_{\vec{x}} f\big(\{U_\mu^{x_0,\vec{x}}\}\big)\right), \quad (13)$$

where the arguments of f are eight matrices at spatial location $\vec{x}$. A general approach to find F and f is presented in Appendix F. A simpler approach is based on the observation that we do not need any hidden layers in the prediction network, i.e., we only keep the output neuron; see Table II. Then the decision function simplifies to $D(S) = \mathrm{sigmoid}[w\,Q(S) + b]$, where

$$Q(S) = \frac{2}{N}\sum_{\vec{x}} f\big(\{U_\mu^{x_0,\vec{x}}\}\big) \quad (14)$$

reduces to a sum of functions acting only on a single patch of size $2\times1\times1\times1$ each. This allows us to split all samples to a minimum size of $2\times1\times1\times1$. We train a new local neural network to classify the phases of each local sample. By performing a regression on the latent prediction of the local neural network, we find that a second-order polynomial performs best (a comparison of different algorithms is found in Appendix E). The regression approximates the latent prediction by a sum of 561 terms,

$$f\big(\{U_\mu^{x_0}\}\big) \approx +7.3816\, a_\tau^0 a_\tau^1 + 0.2529\, a_\tau^1 b_\tau^1 + \cdots$$
$$- 0.2869\, d_\tau^0 c_\tau^1 - 7.2279\, b_\tau^0 b_\tau^1$$
$$- 7.3005\, c_\tau^0 c_\tau^1 - 7.4642\, d_\tau^0 d_\tau^1. \quad (15)$$

We only keep the leading contributions and assume that the differences between the leading contributions originate from approximation errors. Since overall factors and intercepts can be absorbed in the weights and biases of the neural network, we can simply rescale the above parameter to

$$f\big(\{U_\mu^{x_0}\}\big) \approx a_\tau^0 a_\tau^1 - b_\tau^0 b_\tau^1 - c_\tau^0 c_\tau^1 - d_\tau^0 d_\tau^1 = \mathrm{tr}\big(U_\tau^0 U_\tau^1\big). \quad (16)$$

This is the Polyakov loop on a single spatial lattice site (12). We promote $f(\{U_\mu^{x_0}\}) \to f(\{U_\mu^{x_0,\vec{x}}\})$ to act on the full lattice, such that we can formulate the decision function of the neural network with the full receptive field as

$$D(S) \approx \mathrm{sigmoid}\left[w\left(\frac{2}{N}\sum_{\vec{x}} f\big(\{U_\mu^{x_0,\vec{x}}\}\big)\right) + b\right]. \quad (17)$$

Here $Q(S) = \frac{2}{N}\sum_{\vec{x}} f(\{U_\mu^{x_0,\vec{x}}\})$ is the Polyakov loop on the full lattice. A confirmation of this deduction can be seen in the perfect correlation between the latent prediction and the Polyakov loop in Figs. 3(a) and 3(b).

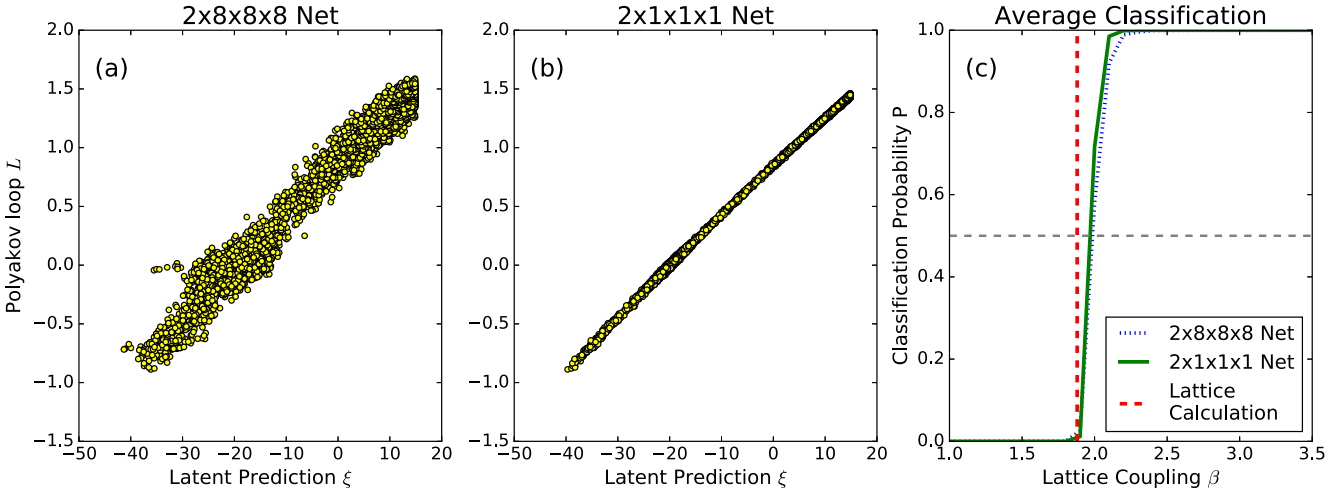

FIG. 3. Results of the correlation probing network applied to SU(2) lattice gauge theory. (a),(b) The latent prediction shows a strong correlation with the Polyakov loop in both the $2\times8\times8\times8$ network and the $2\times1\times1\times1$ network. (c) The average prediction probability of the two networks.

TABLE III. Ising model neural network. A, B, C determine the receptive field size of each neuron in the averaging layer.

Layer	Output shape	Kernel size
InputLayer	(784, 1)	
Convolution1D	$(784/(A),n_A)$	A
Convolution1D	$(784/(A \times B),n_B)$	B
Convolution1D	$(784/(A \times B \times C),n_C)$	C
Average pooling	$(1, n_C)$	
Flatten	(n_C)	
Dense	(n_D)	
Dense	(1)	

TABLE V. Scores of different regression algorithms. Higher is better.

Order of regression	Train score	Validation score
Polynomial regression		
1	0.00128	-0.00042
2	**0.72025**	**0.72395**
3	0.75675	0.69129
Support vector regression		
1	-0.08943	-0.08988
2	0.64048	0.65367
3	-0.08434	-0.08963

V. CONCLUSION

We proposed and demonstrated a method to fully interpret neural networks, which is based on the correlation probing neural network. The method was introduced at the Ising model on the square lattice, where the neural network predicts phases via the magnetization (5) or the expected energy per site (9). We then demonstrated the power of this method at SU(2) lattice gauge theory, where the reconstructed decision function reveals the explicit mathematical expression of the Polyakov loop (17), a nonlinear, nonlocal order parameter. This method provides the means to judge whether neural networks have learned physical properties and thus whether their results can be trusted. Furthermore, our procedure can be used to deduce the explicit formulas of physical order parameters. Since our approach is vastly different than conventional methods, it could determine the nature of phases where conventional methods have not yet succeeded.

A first application could be identifying if machine learning methods classify sign problematic models by physical quantities [26]. Then we could reliably determine the phase diagram of QCD at finite density [44–46] or examine the pseudogap [47–50] or the competition between d-wave and antiferromagnetic order [51–54] in the two-dimensional Hubbard model.

ACKNOWLEDGMENTS

We would like to thank J. M. Pawlowski, M. Salmhofer, I.-O. Stamatescu, and C. Wetterich for useful discussions. We thank M. Neidig and S. Nkongolo for reviewing the manuscript. S.J.W. acknowledges support by the Heidelberg Graduate School of Fundamental Physics. M.S. was supported by the DFG via Project No. STA283/16-2.

TABLE IV. SU(2) Neural network. A, B, C determine the receptive field size of each neuron in the averaging layer.

Layer	Output shape	Kernel size
InputLayer	(1024, 16)	
Convolution1D	$(1024/(A),n_A)$	A
Convolution1D	$(1024/(A \times B),n_B)$	B
Convolution1D	$(1024/(A \times B \times C),n_C)$	C
Average pooling	$(1,n_C)$	
Flatten	(n_C)	
Dense	(n_D)	
Dense	(1)	

APPENDIX A: MONTE CARLO SIMULATIONS

In statistical physics and lattice gauge theory, Markov Chain Monte Carlo algorithms are used to sample lattice configurations from the Boltzmann factor. This is done by constructing a stochastic sequence that starts at some random initial configuration. This stochastic sequence is constructed such that the configurations obey Boltzmann statistics in the equilibrium. For more details on algorithm requirements and algorithms see, e.g., [55].

Observables are then computed by taking the average over many spin or lattice configurations from the equilibrium distribution

$$\langle \mathcal{O} \rangle = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \mathcal{O}_i. \tag{A1}$$

Taking the limit in the last equality is practically not possible. Hence, the expectation value of the observable is approximated by large N and gives rise to a statistical error. It is important to take enough configurations such that ergodicity is achieved. In the case of two distinct regions of phase space, this can take a very long simulation time.

For the Ising model, we produced a total of 55 000 spin configurations, of size 28×28, equally distributed over 11 equidistant temperature values $T \in [0,5]$ by employing the Metropolis-Hastings algorithm [56] with simulated annealing.

For SU(2), we used the Heatbath algorithm [57] to produce a total of 15 600 decorrelated configurations equally distributed over 26 values in the range of $\beta_{\text{latt}} = 4/g^2 \in [1,3.5]$. In the context of this paper it is important to have decorrelated data, since neural networks are good at finding structures, and thus correlations between configurations, if existent. Due to center symmetry breaking, in the deconfined phase the average Polyakov loop can take either positive or negative values of equal magnitude. In our simulations, we initiated all links with the unit matrix, hence we introduced a bias for large values of β_{latt}, i.e., our simulations are not fully ergodic. At large values of β_{latt}, this will prevent a full exploration of phase space. If we were to employ neural networks to extract the position of the phase transition, this nonergodicity leads to a shift in the value of critical β_{latt}. Generally speaking, ergodicity can be retained by doing more simulations and employing algorithms such as

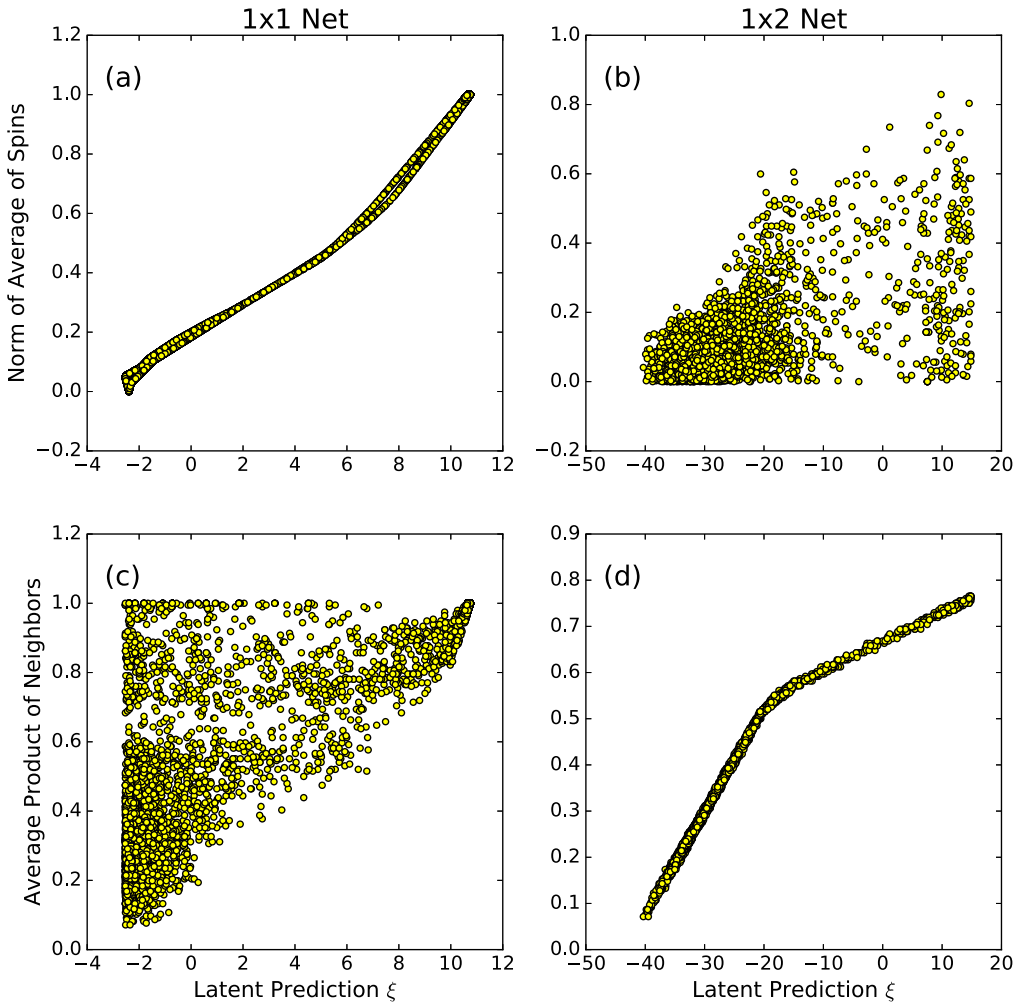

FIG. 4. Results of the correlation probing neural network applied to the Ising model. (a) The latent prediction is perfectly correlated with the absolute value of the average of spins, i.e., the magnetization in the 1×1 network. (b) The latent prediction of the 1×2 network is not correlated with the absolute value of the average of spins. (d) The latent prediction of the 1×2 network is perfectly correlated with the average product of neighbors, i.e., the expected energy per site. (c) The latent prediction of the 1×1 network is not correlated with the average of neighbors.

simulated annealing or overrelaxation, thus in principle it is possible to extract the critical temperature reliably.

APPENDIX B: NEURAL NETWORK ARCHITECTURE

We constructed our machine learning pipeline using Scikit-learn [58] and Keras [59]. The neural network architectures are presented in Tables III and IV. Since there is no Convolutional4D in Keras, we just rearranged our samples to fit a Convolutional1D layer. We used neural networks with number of filters $n_A, n_B, n_D \in \{1,4,8,32,256,1024\}$. The kernel sizes A, B, C are used to set the receptive field size. For our problems, $n_C = 1$ is sufficient to capture the structure of the order parameter. This was probed in the same manner as finding the optimal receptive field size. In other models one might need a higher n_C, e.g., in the Heisenberg model, $n_C = 3$ could be optimal. Hence, this can already be an early indicator for the type of the broken symmetry. The activation functions are rectified linear units $\mathrm{relu}(x) = \max(0,x)$ between all layers and the sigmoid function $\mathrm{sigmoid}(x) = 1/[1+\exp(-x)]$ in the last layer. We do not employ any sort of regularization. The training objective is minimizing the binary cross entropy loss function

$$C(Y,P) = -\frac{1}{N}\sum_i [y_i \ln p_i + (1-y)\ln(1-p_i)], \qquad \text{(B1)}$$

where $Y = y_i$ is a list of labels and $P = p_i$ is the corresponding list of predictions. Our baseline classifier is the classifier which assigns each label with a probability of $p_i = 0.5$. This means that this classifier just assigns a label to each sample randomly. The binary cross entropy then evaluates to 0.6931. The neural networks learn by optimizing the weights and biases via RMSprop gradient descent. The neural networks were trained for 300 epochs or less, if the loss already saturated in earlier epochs. The validation set is 20% of the training data.

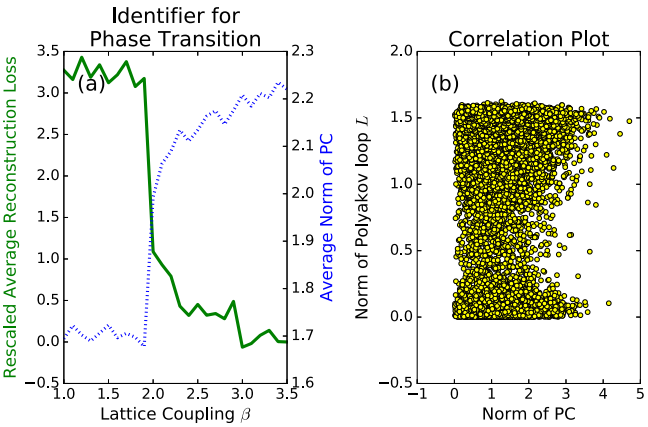

FIG. 5. (a) Finding a possible phase transition in SU(2) lattice gauge theory with PCA. Green solid: The average mean squared error reconstruction loss as a function of temperature is a universal identifier for a phase transition. It was calculated in 100 independent incremental PCA runs with two principal components (PC), measured in units of $\times 10^{-5}$ and shifted by the value at $\beta = 3.5$. Blue dotted: The average norm of the PC also indicates a phase transition. (b) There is no correlation between the principal components and the Polyakov loop.

APPENDIX C: CROSS COMPARISON OF ISING MODEL NEURAL NETWORK RESULTS

In Fig. 4 we show that latent parameters of the neural networks applied to the Ising model cannot be simultaneously correlated with the magnetization and the average energy per spin.

APPENDIX D: UNSUPERVISED LEARNING OF PHASE TRANSITIONS IN SU(2) LATTICE GAUGE THEORY

We assume no prior knowledge of the phase transition, even its existence. Hence, we employ unsupervised learning to find any possible indications for a phase transition. For the sake of simplicity we employ principal component analysis (PCA) [32,60] with two principal components. PCA is an orthogonal linear transformation of the input samples to a set of variables, sorted by their variance. Here, unsupervised learning algorithms that are based on the reconstruction loss like autoencoders [33] are doomed to fail, since the states are gauge invariant. The autoencoder would need to predict a matrix which is not unique.

Even though the Polyakov loop is a nonlinear order parameter, PCA captures indications of a phase transition at $\beta \in [1.8, 2.2]$, which is demonstrated in Fig. 5(a). Here we employed the average reconstruction loss [33] and the Euclidean norm of the principal components as identifiers for a phase transition. In Fig. 5(b) we show that there is no correlation between the Polyakov loop and the principal components.

It is worth noting that this example shows that PCA can capture phase indicators even when the principal components cannot approximate any order parameter.

APPENDIX E: REGRESSION OF THE POLYAKOV LOOP IN THE LOCAL NEURAL NETWORK

We perform a regression on the latent prediction of the local neural network on only 1% of the local samples of size $2 \times 1 \times 1 \times 1$ and use another 1% as validation set. By comparing different algorithms, we find that a second-order polynomial regression gives the best results; see Table V.

APPENDIX F: FULL REGRESSION OF THE POLYAKOV LOOP IN THE GLOBAL NEURAL NETWORK

Here we present the general procedure for reconstructing the decision function of a neural network applied to SU(2) gauge theory. Since it requires separating the correlation probing network into subnetworks, and transferring weights between different networks, it requires more advanced knowledge of artificial neural network architecture.

The decision function of the $2 \times 1 \times 1 \times 1$ neural network which predicts the lattice SU(2) phase transition is by construction

$$D(S) = F\Bigg(\frac{2}{N}\sum_{\vec{x}} f\big(U_\tau^{0,\vec{x}}, U_x^{0,\vec{x}}, U_y^{0,\vec{x}}, U_z^{0,\vec{x}}, U_\tau^{1,\vec{x}}, U_x^{1,\vec{x}}, U_y^{1,\vec{x}}, U_z^{1,\vec{x}}\big)\Bigg). \tag{F1}$$

In general, we cannot assume that the prediction network consists only of the output neuron. Therefore, we suggest a different procedure for constructing the decision function. We split the full correlation probing net into subnetworks: we extract the localization network plus averaging layer and the prediction network as separate networks. In order to determine $F(S) = \mathrm{sigmoid}(\xi(S))$, we use polynomial regression to fit the latent prediction of the prediction network to the output of the averaging layer. We find a polynomial of degree 1 is enough to fit the data, and ξ is approximated by

$$\xi(x) \approx -0.7101\, x + 9.851\,434\,19. \tag{F2}$$

The slope and intercept can be absorbed by the weight w and bias b of the output neuron, such that we can infer

$$\xi(x) \approx w\, x + b. \tag{F3}$$

The function f requires us to build a new local neural network which only acts on patches of size $2 \times 1 \times 1 \times 1$. By construction this network has the same number of weights and biases as the full neural network acting on the input of size $2 \times 8 \times 8 \times 8$. Instead of training the local neural network, we transfer the weights and biases from the full correlation probing network to the local neural network. Hence, one can obtain the output of the localization network for each patch separately. Again, we employ polynomial regression to fit the input from the local patches to the output of the localization network. The result of a regression of degree 2 with 561 parameters yields

$$\begin{aligned} f\big(\{U_\mu^{x_0}\}\big) \approx &- 26.8354\, a_\tau^0 a_\tau^1 - 2.4972\, d_\tau^0 c_\tau^1 + \cdots \\ &+ 1.5653\, b_\tau^0 c_\tau^0 + 26.5908\, b_\tau^0 b_\tau^1 \\ &+ 27.7054\, c_\tau^0 c_\tau^1 + 27.8939\, d_\tau^0 d_\tau^1. \end{aligned} \tag{F4}$$

After absorbing overall factors and the intercept by the weights and biases of the prediction network and neglecting the subleading terms, we rewrite f as

$$f\left(\{U_\mu^{x_0}\}\right) \approx a_\tau^0 a_\tau^1 - b_\tau^0 b_\tau^1 - c_\tau^0 c_\tau^1 - d_\tau^0 d_\tau^1. \quad \text{(F5)}$$

This is the Polyakov loop on a single lattice site. By employing (F5) as an argument of (F3), we can promote $f(\{U_\mu^{x_0}\}) \to f(\{U_\mu^{x_0,\vec{x}}\})$ to depend on space again. We obtain the definition of the decision function

$$D(S) \approx \text{sigmoid}\left[w\left(\frac{2}{N} \sum_{\vec{x}} f\left(\{U_\mu^{x_0,\vec{x}}\}\right) \right) + b \right], \quad \text{(F6)}$$

where $Q(S) = [\frac{2}{N} \sum_{\vec{x}} f(\{U_\mu^{x_0,\vec{x}}\})]$ is the Polyakov loop on the full lattice.

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, in *Advances in Neural Information Processing Systems 25*, edited by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Curran Associates, Inc., Lake Tahoe, Nevada, USA, 2012), pp. 1097–1105.
[2] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury, IEEE Signal Process. Mag. **29**, 82 (2012).
[3] D. Silver *et al.*, Nature (London) **529**, 484 (2016).
[4] G. Carleo and M. Troyer, Science **355**, 602 (2017).
[5] G. Torlai and R. G. Melko, Phys. Rev. B **94**, 165134 (2016).
[6] D.-L. Deng, X. Li, and S. Das Sarma, arXiv:1609.09060.
[7] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. X **7**, 021021 (2017).
[8] X. Gao and L.-M. Duan, Nat. Commun. **8**, 662 (2017).
[9] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, arXiv:1703.05334.
[10] K.-I. Aoki and T. Kobayashi, Mod. Phys. Lett. B **30**, 1650401 (2016).
[11] G. Kasieczka, T. Plehn, M. Russell, and T. Schell, J. High Energy Phys. 05 (2017) 006.
[12] L. Huang and L. Wang, Phys. Rev. B **95**, 035105 (2017).
[13] J. Liu, Y. Qi, Z. Y. Meng, and L. Fu, Phys. Rev. B **95**, 041101(R) (2017).
[14] N. Portman and I. Tamblyn, J. Comput. Phys. **350**, 871 (2017).
[15] L. Li, T. E. Baker, S. R. White, and K. Burke, Phys. Rev. B **94**, 245129 (2016).
[16] Y. Levine, D. Yakira, N. Cohen, and A. Shashua, arXiv:1704.01552.
[17] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature **549**, 195 (2017).
[18] S. Lloyd, M. Mohseni, and P. Rebentrost, Nat. Phys. **10**, 631 (2014).
[19] M. Kieferova and N. Wiebe, arXiv:1612.05204.
[20] D. Crawford, A. Levit, N. Ghadermarzy, J. S. Oberoi, and P. Ronagh, arXiv:1612.05695.
[21] H. W. Lin, M. Tegmark, and D. Rolnick, J. Stat. Phys. **168**, 1223 (2017).
[22] E. Miles Stoudenmire and D. J. Schwab, *Advances in Neural Information Processing Systems 29*, edited by D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (Curran Associates, Inc., New York, 2016), pp. 4799–4807.
[23] M. H. Amin, E. Andriyash, J. Rolfe, B. Kulchytskyy, and R. Melko, arXiv:1601.02036.
[24] P. Mehta and D. J. Schwab, arXiv:1410.3831.
[25] J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).
[26] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, Sci. Rep. **7**, 8823 (2017).
[27] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Phys. Rev. X **7**, 031038 (2017).
[28] Y. Zhang and E.-A. Kim, Phys. Rev. Lett. **118**, 216401 (2017).
[29] F. Schindler, N. Regnault, and T. Neupert, Phys. Rev. B **95**, 245134 (2017).
[30] A. Tanaka and A. Tomiya, J. Phys. Soc. Jpn. **86**, 063001 (2017).
[31] T. Ohtsuki and T. Ohtsuki, J. Phys. Soc. Jpn. **85**, 123706 (2016).
[32] L. Wang, Phys. Rev. B **94**, 195105 (2016).
[33] S. J. Wetzel, Phys. Rev. E **96**, 022140 (2017).
[34] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nat. Phys. **13**, 435 (2017).
[35] W. Hu, R. R. P. Singh, and R. T. Scalettar, Phys. Rev. E **95**, 062122 (2017).
[36] G. Towell and J. W. Shavlik, in *Advances in Neural Information Processing Systems 4*, edited by J. E. Moody, S. J. Hanson, and R. P. Lippmann (Morgan-Kaufmann, Denver, CO, USA, 1992), pp. 977–984.
[37] A. Mahendran and A. Vedaldi, arXiv:1412.0035.
[38] Z. Ghahramani, Nature (London) **521**, 452 (2015).
[39] Z. C. Lipton, arXiv:1606.03490.
[40] R. Shwartz-Ziv and N. Tishby, arXiv:1703.00810.
[41] P. Ponte and R. G. Melko, arXiv:1704.05848.
[42] L. Onsager, Phys. Rev. **65**, 117 (1944).
[43] K. G. Wilson, Phys. Rev. D **10**, 2445 (1974).
[44] M. G. Alford, A. Schmitt, K. Rajagopal, and T. Schäfer, Rev. Mod. Phys. **80**, 1455 (2008).
[45] M. A. Stephanov, Int. J. Mod. Phys. A **20**, 4387 (2005).
[46] J. B. Kogut and M. A. Stephanov, Cambridge Monogr. Part. Phys., Nucl. Phys., Cosmol. **21**, 1 (2004).
[47] C. M. Varma, Phys. Rev. B **73**, 155113 (2006).
[48] L. Taillefer, Annu. Rev. Condens. Matter Phys. **1**, 51 (2010).
[49] Y. Wang, N. P. Ong, Z. A. Xu, T. Kakeshita, S. Uchida, D. A. Bonn, R. Liang, and W. N. Hardy, Phys. Rev. Lett. **88**, 257003 (2002).
[50] S. Hüfner, M. A. Hossain, A. Damascelli, and G. A. Sawatzky, Rep. Prog. Phys. **71**, 062501 (2008).
[51] S. Friederich, H. C. Krahl, and C. Wetterich, Phys. Rev. B **83**, 155125 (2011).
[52] T. A. Maier, M. Jarrell, T. C. Schulthess, P. R. C. Kent, and J. B. White, Phys. Rev. Lett. **95**, 237001 (2005).
[53] S. Raghu and S. A. Kivelson, Phys. Rev. B **83**, 094518 (2011).
[54] C. J. Halboth and W. Metzner, Phys. Rev. Lett. **85**, 5162 (2000).
[55] C. Gattringer and C. B. Lang, *Quantum Chromodynamics on the Lattice* (Springer, Berlin, 2010).
[56] N. Metropolis and S. Ulam, J. Am. Stat. Assoc. **44**, 335 (1949).
[57] M. Creutz, Phys. Rev. D **21**, 2308 (1980).
[58] F. Pedregosa *et al.*, J. Machine Learn. Res. **12**, 2825 (2011).
[59] F. Chollet, "keras", https://github.com/fchollet/keras (2015).
[60] K. Pearson, Philos. Mag. J. Ser. **2**, 559 (1901).

6 Exploring the Hubbard Model on the Square Lattice at Zero Temperature with a Bosonized Functional Renormalization Approach

Wetzel, Sebastian J.
eprint arXiv:1712.04297 [4]

Exploring the Hubbard Model on the Square Lattice at Zero Temperature with a Bosonized Functional Renormalization Approach

Sebastian J. Wetzel[1]

[1] *Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany*

We employ the functional renormalization group to investigate the phase diagram of the $t-t'$ Hubbard model on the square lattice with finite chemical potential μ at zero temperature. A unified scheme to derive flow equations in the symmetric and symmetry broken regimes allows a consistent continuation of the renormalization flow in the symmetry broken regimes. At the transition from the symmetric regime to the symmetry broken regimes, our calculation reveals leading instabilities in the d-wave superconducting and antiferromagnetic channels. Furthermore, we find a first order transition between commensurate and incommensurate antiferromagnetism. In the symmetry broken regimes our flow equations are able to renormalize around a changing Fermi surface geometry. We find a coexistence of d-wave superconductivity and antiferromagnetism at intermediate momentum scales k. However, there is a mutual tendency of superconductivity and antiferromagnetism to repel each other at even smaller scales k, which leads to the eradication of the coexistence phase in the limit of macroscopic scales.

I. INTRODUCTION

The $t-t'$- Hubbard model [1–3] is a promising model to describe the phase diagram of electrons in CuO planes as they occur in cuprate high-temperature superconductors. The most distinct phases are antiferromagnetism and d-wave superconductivity in close vicinity to each other[4–20]. Early works could already infer d-wave superconductivity by scaling arguments [21–23]. More advanced purely fermionic renormalization calculations revealed the leading instabilities of the Hubbard model in the antiferromagnetic and in the d-wave channel [24–39]. While these methods have problems to calculate results in the symmetry broken phases, a combination of different schemes like renormalization group and mean field calculations can give insights in the interplay of different orders [40–45]. Our approach builds upon a bosonized renormalization group analysis of the Hubbard model which allows the renormalization flow to enter symmetry broken phases [46–53]. For this purpose we employ a scale dependent Hubbard-Stratonovich transformation [54, 55] which is generated during the renormalization flow [56, 57].

We employ a bosonized functional renormalization group calculation to examine the phase diagram of the $t-t'$ Hubbard model with finite chemical potential μ at zero temperature. We formulate a consistent set of flow equations in the symmetric and symmetry broken phases, which are supplied by a renormalization scheme which works in the symmetric and all symmetry broken phases. This allows us to renormalize around a changing Fermi surface in the magnetic symmetry broken phase.

The initial conditions of our flow equations are given by the fermionic Hubbard action in the Matsubara formalism, eq. (4). By examining the leading instabilities of the flow equations in the symmetric regime, we identify three different channels in which symmetry breaking occurs, see fig. 1. These correspond to commensurate antiferromagnetism, incommensurate antiferromagnetism and d-wave superconductivity. An advantage of our method is the way of identifying incommensurate antiferromagnetism. This is because the the minimum of the inverse magnetic propagator is directly related to the type of antiferromagnetism. This property allows for an accurate examination of the transition between commensurate and incommensurate antiferromagnetism, see fig. 4, which we identify as a transition of first order. By continuing the flow equations into the regimes of symmetry breaking, we are able to examine the interplay of the magnetic and d-wave order parameters. We find that at intermediate scales k there is a regime of coexisting antiferromagnetic and d-wave condensate. Our results suggest that these two phases have a tendency to repel each other, which leads to the vanishing of the coexistence phase at macroscopic scales, see fig. 7.

II. FIELD THEORETIC FORMULATION OF THE HUBBARD MODEL

A. Hubbard Model

The t–t' Hubbard model on the square lattice is defined by the Hamiltonian

$$H = -\sum_{ij,\sigma} t_{ij} c^\dagger_{i,\sigma} c_{j,\sigma} + h.c. + U \sum_i (c^\dagger_{i,\downarrow} c_{i,\downarrow})(c^\dagger_{i,\uparrow} c_{i,\uparrow}) \ , \quad (1)$$

where $t_{ij} = t$ for nearest neighbors, $t_{ij} = t'$ for next-to-nearest neighbors and $t_{ij} = 0$ otherwise. We define the energy scales by setting $t = 1$. The fermionic dispersion relation in momentum space is

$$\xi(Q) = -\mu - 2t(\cos(q_x) + \cos(q_y)) - 4t'(\cos(q_x)\cos(q_y)) \ , \quad (2)$$

where we have already included the chemical potential μ, which denotes the level of doping.

B. Functional Renormalization Group

We address the phase diagram of the Hubbard model by calculating the quantum effective action Γ via the flow equation for the effective average action Γ_k [58]

$$\partial_k \Gamma_k = \frac{1}{2}\mathrm{STr}\left(\Gamma_k^{(2)} + R_k\right)^{-1} \partial_k R_k = \frac{1}{2}\mathrm{STr}\,\tilde{\partial}_k \ln\left(\Gamma_k^{(2)} + R_k\right) . \tag{3}$$

We also introduce a short form notation of the scale derivative $\tilde{\partial}_k = (\partial_k R_k)\partial_{R_k}$ acting only on the regulator. For small $k \to 0$, the effective average action becomes the full effective action Γ. In the limiting case of large scales $k \to \Lambda$, the effective action equals the microscopic action $\Gamma_k \to S$. Thus the initial condition for our renormalization group calculation is defined by the Hubbard model action in the Matsubara formalism

$$\begin{aligned} S = &\sum_Q \psi^\dagger(Q)\left(i\omega_Q + \xi(Q)\right)\psi(Q) \\ &+ \frac{U}{2}\sum_{Q_1,\dots,Q_4} \left(\psi^\dagger(Q_1)\psi(Q_2)\right)\left(\psi^\dagger(Q_3)\psi(Q_4)\right) \\ &\times \delta(Q_1 - Q_2 + Q_3 - Q_4) . \end{aligned} \tag{4}$$

The fields depend on a collection of momenta and the Matsubara frequency $Q = (\omega_Q, \vec{q}) = (\omega_Q, q_x, q_y)$. The sum is a short hand notation for an integration over all momenta and Matsubara frequencies $\sum_Q = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \int_{[-\pi,\pi]^2} \frac{dq^2}{(2\pi)^2}$. The electrons are written as four component Grassmann valued fields $\psi(Q) = (\psi(Q)_\downarrow, \psi(Q)_\uparrow)^T$. While the Matsubara frequencies are discrete at nonzero temperature, they take on continuous values at zero temperature. This has a strong effect on the flow equations: The fermionic propagator is no longer gapped by the lowest Matsubara frequencies $\omega = \pm\pi T$, thus fermionic fluctuations contribute even for small energy scales k. This fact is also responsible for the increased effect of the shape of the Fermi surface on the renormalization flow. Hence, special care needs to be taken when dealing with contributions close to the Fermi surface. Furthermore, in the bosonic sector at finite temperature, only the Matsubara zero mode contributes to the flow equations, which induces dimensional reduction at low scales $k \ll \pi T$. In this case, the theory can be influenced only by spatial fluctuations. These two finite temperature properties cannot be exploited in the derivation of flow equations at zero temperature and thus yield an extra challenge for our calculations.

C. Truncation

While the Hubbard action, eq. (4), is the initial condition for the flow equation, other couplings are generated during the renormalization flow. The average effective action can be decomposed into contributions with respect to their fermionic and bosonic content

$$\Gamma_k = \Gamma_{F,k} + \Gamma_{FB,k} + \Gamma_{B,k} . \tag{5}$$

The fermionic part $\Gamma_{F,k}$ contains the Hubbard action (4) and contributions Γ_F^m, Γ_F^d mimicking the magnetic and d-wave contributions of the Hubbard interaction U. They absorb the respective momentum dependence of U arising during the renormalization flow.

$$\begin{aligned} \Gamma_{F,k} = &\sum_Q \psi^\dagger(Q) P_F(Q) \psi(Q) \\ &+ \frac{U}{2}\sum_{Q_1,\dots,Q_4} \left(\psi^\dagger(Q_1)\psi(Q_2)\right)\left(\psi^\dagger(Q_3)\psi(Q_4)\right) \\ &\times \delta(Q_1 - Q_2 + Q_3 - Q_4) + \Gamma_F^m + \Gamma_F^d \end{aligned} \tag{6}$$

While we keep the Hubbard interaction U fixed, we allow for a fermionic wave function renormalization Z_F in the kinetic term $P_F = Z_F\left(i\omega_Q + \xi(Q)\right)$. The fermionic momentum channels

$$\begin{aligned} \Gamma_F^m = &-\frac{1}{2}\sum_{Q_1,\dots,Q_4} \lambda_F^m(Q_1 - Q_2)\delta(Q_1 - Q_2 + Q_3 - Q_4) \\ &\times \left(\psi^\dagger(Q_1)\vec{\sigma}\psi(Q_2)\right)\left(\psi^\dagger(Q_3)\vec{\sigma}\psi(Q_4)\right) \end{aligned} \tag{7}$$

and

$$\begin{aligned} \Gamma_F^d = &-\frac{1}{2}\sum_{Q_1,\dots,Q_4} \lambda_F^d(Q_1 + Q_3)\delta(Q_1 - Q_2 + Q_3 - Q_4) \\ &\times f_d((Q_1 - Q_3)/2) f_d((Q_2 - Q_4)/2) \\ &\times \left(\psi^\dagger(Q_1)\epsilon\psi^*(Q_3)\right)\left(\psi^T(Q_2)\epsilon\psi(Q_4)\right) \end{aligned} \tag{8}$$

will be kept zero during the renormalization flow. Their flow will be redefined as contributions to the Yukawa couplings h_m, h_d of the magnetic and d-wave bosons by flowing bosonization, see appendix B 5.

The purely bosonic part of the effective average action is defined by

$$\begin{aligned} \Gamma_{B,k} = &\frac{1}{2}\sum_Q \vec{m}^T(-Q) P_m(Q) \vec{m}(Q) \\ &+ \sum_Q d^*(Q) P_d(Q) d(Q) \\ &+ \sum_X U_k(\rho_m, \rho_d) . \end{aligned} \tag{9}$$

Here $\vec{m}$ describes a magnetic boson and d a d-wave superconducting Cooper-pair. The effective potential U depends on the symmetry invariants $\rho_m = \frac{1}{2}\vec{m}^T\vec{m}$ and $\rho_d = d^*d$. The kinetic contributions can be decomposed into a frequency dependent part and a momentum dependent part.

$$\begin{aligned} P_m(Q) &= Z_m\omega_Q^2 + A_m F_m(Q) \\ P_d(Q) &= Z_d\omega_Q^2 + A_d F_d(Q) . \end{aligned} \tag{10}$$

The minimum of the spatial shape factor F_d is found at $(0,0)$, however the minimum of the magnetic kinetic shape factor F_m can take different values for different kinds of magnetism. A minimum at $(0,0)$ denotes ferromagnetism, while a minimum at $\vec{\pi} = (\pi,\pi)$ denotes antiferromagnetism. In our phase diagrams we also find incommensurate antiferromagnetism, where the minimum is fourfold degenerate on the axis at $(\pi, \pi \pm \delta_{ic})$ and $(\pi \pm \delta_{ic}, \pi)$.

We parametrize the bosonic propagators with functions which allow for an accurate examination of the expansion around their minima. The magnetic propagator in the case of commensurate antiferromagnetism is given by

$$F_{m,c}(Q) = \frac{D_m |[\vec{q}+\vec{\pi}]|^2}{D_m + |[\vec{q}+\vec{\pi}]|^2} \ , \tag{11}$$

where $[\vec{q}] = ((q_x + \pi \mod 2\pi) - \pi, (q_y + \pi \mod 2\pi) - \pi)$ denotes the projection of momenta into the 1st Brillouin zone $[-\pi,\pi]^2$. In the case of incommensurate antiferromagnetism this parametrization is enhanced by

$$F_{m,ic}(Q) = \frac{D_m \tilde{F}(\vec{q})}{D_m + \tilde{F}(\vec{q})} \ , \tag{12}$$

where

$$\tilde{F}(\vec{q}) = \frac{1}{4\delta_{ic}} \left((|[\vec{q}+\vec{\pi}]|^2 - \delta_{ic}^2)^2 + 4[q_x+\pi]^2[q_y+\pi]^2 \right) \tag{13}$$

is employed to expand around the incommensurate minimum. The d-wave propagator is similarly parametrized by

$$F_d(Q) = \frac{D_d |[\vec{q}]|^2}{D_d + |[\vec{q}]|^2} \ . \tag{14}$$

The interactions in the bosonic sector are contained in the effective potential

$$U_k(\rho_m, \rho_d) = \sum_{n=1}^{\mathrm{ord}} u_n \ , \tag{15}$$

to various orders in ρ_m and ρ_d. The lowest order of the effective potential

$$u_1 = \lambda_{10}(\rho_m - \rho_{m0}) + \lambda_{01}(\rho_d - \rho_{d0}) \tag{16}$$

contains the mass terms $\lambda_{10} = m_m^2$ and $\lambda_{01} = m_d^2$. They are finite in the symmetric phases and zero in the symmetry broken phases. In the latter case we expand the effective potential around the minimum at (ρ_{m0}, ρ_{d0}). The second order interactions

$$\begin{aligned} u_2 = &\frac{\lambda_{20}}{2}(\rho_m - \rho_{m0})^2 + \lambda_{11}(\rho_m - \rho_{m0})(\rho_d - \rho_{d0}) \\ &+ \frac{\lambda_{02}}{2}(\rho_d - \rho_{d0})^2 \end{aligned} \tag{17}$$

determine the curvature around the expansion point. As long as $\det_{BB} = \lambda_{20}\lambda_{02} - \lambda_{11}^2 > 0$, the expansion point is a true minimum. The third order contributions are

$$\begin{aligned} u_3 = &\frac{\lambda_{30}}{6}(\rho_m - \rho_{m0})^3 + \frac{\lambda_{21}}{2}(\rho_m - \rho_{m0})^2(\rho_d - \rho_{d0}) \\ &+ \frac{\lambda_{12}}{2}(\rho_m - \rho_{m0})(\rho_d - \rho_{d0})^2 + \frac{\lambda_{03}}{6}(\rho_d - \rho_{d0})^3 \ . \end{aligned} \tag{18}$$

The interactions between the bosonic and the fermionic sector are mediated by the Yukawa couplings

$$\begin{aligned} &\Gamma_{FB,k} = \\ &- \sum_{Q_1,Q_2,Q_3} h_m(Q_1)\vec{m}(Q_1)\left(\psi^\dagger(Q_2)\vec{\sigma}\psi(Q_3)\right) \\ &\qquad \times \delta(Q_1 - Q_2 + Q_3) \\ &- \sum_{Q_1,Q_2,Q_3} \frac{1}{\sqrt{2}} h_d(Q_1) f_d((Q_2 - Q_3)/2) \\ &\left(d^*(Q_1)(\psi^T(Q_2)\epsilon\psi(Q_3) - d(Q_1)(\psi^\dagger(Q_2)\epsilon\psi^*(Q_3)\right) \\ &\qquad \times \delta(Q_1 - Q_2 + Q_3) \ . \end{aligned} \tag{19}$$

We parametrize the magnetic Yukawa couplings by the momentum-weighted average of a ferromagnetic interaction and an antiferromagnetic interaction $h_m(Q) = \frac{|[\vec{q}]|}{\sqrt{2}\pi} h_m(\Pi) + \frac{\sqrt{2}\pi - |[\vec{q}]|}{\sqrt{2}\pi} h_m(0)$. While solving the flow equations at $T = 0$, we find $\partial_k h_m(0) = 0$ and thus $h_m(0) = 0$. Here $f_d(Q) = \frac{1}{2}(\cos(q_x) - \cos(q_y))$ is the d-wave form factor. The $1/\sqrt{2}$ prefactor of h_d together with a redefinition of the d-wave boson into real fields $d = \frac{1}{\sqrt{2}}(d_1 + i d_2)$ is useful to treat the magnetic and d-wave bosons on an equal footing. Then they can be summarized in a common language in form of a O(2)×O(3) symmetric bosonic submodel.

The regulator function introduces an artificial mass to the kinetic terms of the bosonic and fermionic fields. In our regularization scheme it only acts on the spatial momentum dependent part

$$\begin{aligned} A_m\, F_{m,k}(Q) &= A_m\, F_m(Q) + R_B(F_m(Q)) \\ A_d\, F_{d,k}(Q) &= A_d\, F_d(Q) + R_B(F_d(Q)) \\ Z_F\, \xi_k(Q) &= Z_F\, \xi(Q) + R_F(\xi(Q)) \end{aligned} \tag{20}$$

for slow momentum modes. In our work we chose the Litim regulator [59]. It acts on the bosonic fields as

$$\begin{aligned} R_B(F_m(Q)) &= A_m\,(k^2 - F_m(Q))\Theta(k^2 - F_m(Q)) \\ R_B(F_d(Q)) &= A_d\,(k^2 - F_d(Q))\Theta(k^2 - F_d(Q)) \ , \end{aligned} \tag{21}$$

and on the fermionic fields as

$$R_F(\xi(Q)) = Z_F \,\mathrm{sign}(\xi(Q))(k - |\xi(Q)|)\Theta(k - |\xi(Q)|) \,. \tag{22}$$

The choice of the infrared cutoff function R_F can be continuoulsy extended to the symmetry broken regimes to address the changing Fermi surface which is discussed in section IV A.

III. LEADING INSTABILITIES IN THE SYMMETRIC REGIME

We solve the set of flow equations, see appendix B, in the symmetric and symmetry broken regimes from $\ln(\Lambda) = \ln(k) = 12$ to $\ln(k) = -8$. The initial condition in the form of the Hubbard action (4) dictates the initial conditions for all other couplings. In the fermionic sector the wave function renormalization starts at $Z_F = 1$. The bosonic sector is initially completely decoupled from the fermions, hence $h_m(Q) = h_d = 0$. The bosons start at their Gaussian fixed points, hence $\lambda_{10} = \lambda_{01} = 1$, $P_m(Q) = P_d(Q) = 0$ and $\lambda_{ij} = 0$ if $i + j > 1$. Exemplary we plot the flow of the most important quantities at $\mu = -0.4, t' = -0.1$ in fig. 3. The parameters correspond to a region in the μ, t'-diagram, fig. 1, where the renormalization group flow enters all four possible regimes: the symmetric regime, antiferromagnetism, d-wave superconductivity and coexistence regime of antiferromagnetism and d-wave superconductivity.

A. Phase Diagram of Leading Instabilities

In this section we examine the solutions of the flow equations in the symmetric regime, in order to obtain a phase diagram of leading instabilities in the Hubbard model on the square lattice. More precisely, a leading antiferromagnetic instability occurs if λ_{10} vanishes at a symmetry breaking scale $k_{SB} > 0$, while λ_{01} is still positive at k_{SB}. Similarly, a leading superconducting instability is characterized by λ_{01} vanishing first. The boundary between antiferromagnetism and superconductivity is found where λ_{10} and λ_{01} vanish simultaneously at a common symmetry breaking scale k_{SB}. Finally, in the unordered phase λ_{10} and λ_{01} remain positive for $k \to 0$. The diagram of leading instabilities corresponds to phase diagrams typically computed by purely fermionic flow equations. Indeed, integrating out the bosonic fields leads to a diverging four fermion interaction in the corresponding channels. In the vicinity of the transition between the antiferromagnetic and the superconducting regions the diagram of leading instabilities does not correspond to the true zero temperature phase diagram, since it does not capture the interplay between the two orders. The full phase diagram can only be obtained by following the flow in the spontaneously broken regime to $k \to 0$. This is discussed in section IV B. The free parameters of our model are the chemical potenital μ, the hopping parameters t, t', and the Hubbard on-site interaction U. We set $t = 1$ so that all other quantities are measured in units of t. The on-site interaction is set to $U = 3$, which is lower than found in many cuprates. It is the choice for comparable renormalization group calculations, in order to avoid problems with too high interaction strengths. Previous experiments and calculations for cuprates agree on $U/t \approx 6-8$ [60–62]. Even with lower interaction strength, the important mechanisms are already present and allow us to investigate the emergence of antiferromagnetism and superconductivity in the phase diagram of the Hubbard model from functional renormalization group calculations.

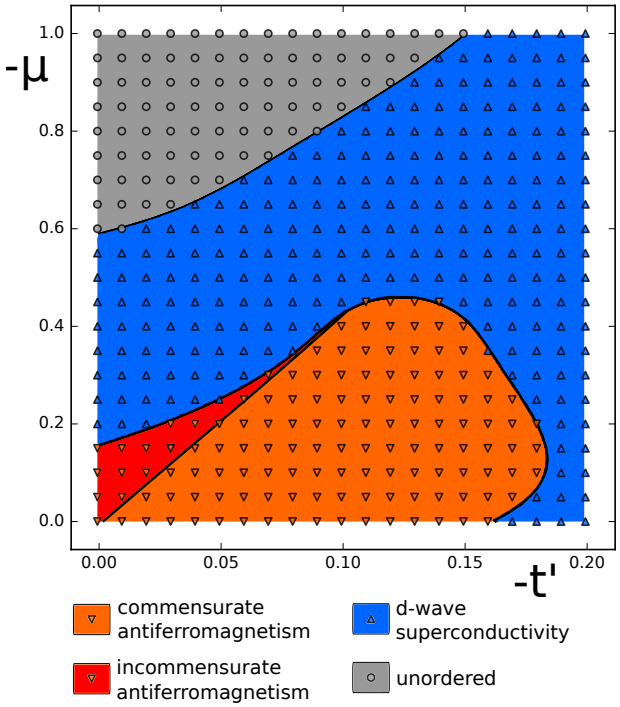

Figure 1: Leading instabilities in the Hubbard model. The diagram was obtained by solving the renormalization flow equations in the symmetric phase for each $\mu - t'$. The truncation includes the effective potential expanded up to ρ^3.

The diagram is calculated on a 21×21 grid containing values of $t' \in [0, -0.2]$ and $\mu \in [0, -1]$. The phase boundaries are statistically optimized using a machine learning algorithm called support vector machine. Doped cuprates, like $La_{2-x}Ba_xCuO_4$ and $La_{2-x}Sr_xCuO_4$, can be found for different levels of doping $x \to \mu$ at $t'/t \approx 0.14 - 0.17$ [63, 64]. Undoped, they exhibit antiferromagnetic order at low temperatures. However doping, or in our picture changing the chemical potential, leads to high-temperature superconductivity.

The diagram of leading instabilities, fig. 1, contains four different phases. At small chemical potential μ and small next-to-nearest neighbor hopping t' there are two

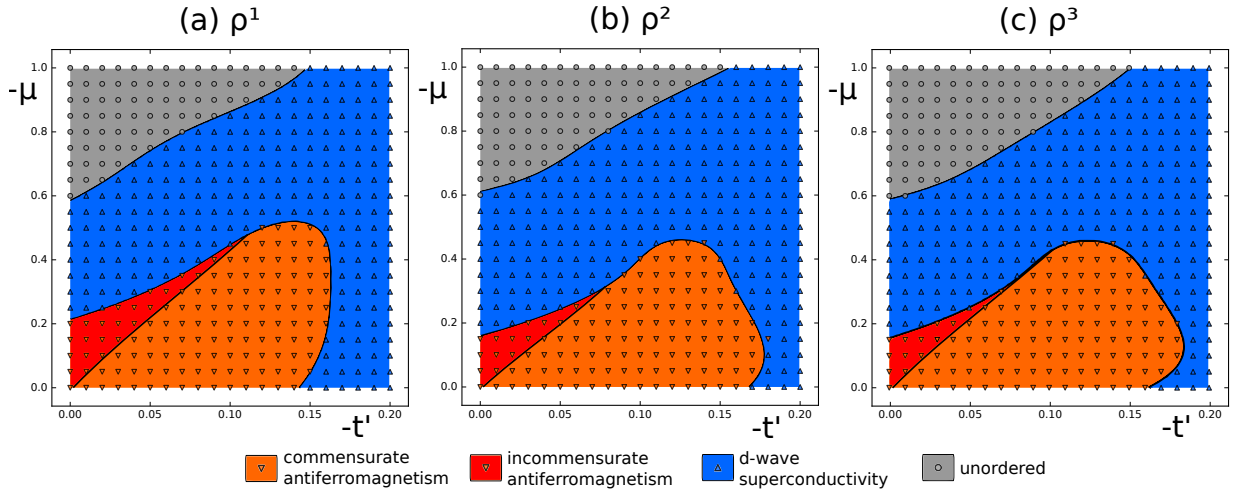

Figure 2: Phase diagrams of leading instabilities calculated with different truncations of the effective potential: expansion up to (a) ρ^1, (b) ρ^2, (c) ρ^3.

different antiferromagnetic regions. The commensurate region is separated from the incommensurate region by the line of Van-Hove singularities $\mu = 4t'$. This transition is further analyzed in section III C. At larger chemical potential and next-to-nearest neighbor hopping the leading instability is the d-wave superconductivity. In section IV B we further demonstrate that it is possible to observe coexistence regions of both antiferromagnetism and d-wave superconductivity. Furthermore, at high chemical potential the ground state is unordered. We are aware that we might have underestimated the size of the antiferromagnetic region, where an accurate quantitative treatment requires the inclusion of charge density and s-wave bosons [53]. Along the Van-Hove line one always observes a broken symmetry in the antiferromagnetic or the d-wave channel. Our approach is not able to resolve a possible ferromagnetic instability [39, 65] for large negative chemical potential and large negative next-to-nearest neighbor hopping on the Van-Hove line. This is a problem of momentum-shell schemes, like ours, and can be circumvented by temperature-flow renormalization techniques for the reasons explained in [30]. Nevertheless, in all other properties the diagram of leading instabilities is in agreement with $\mu - t'$–diagrams from fermionic renormalization group calculations [45].

B. Convergence of Diagrams

It is a priori not clear to which order the effective potential, eq. (15), needs to be expanded to include all necessary effective interactions needed to calculate the phase diagram reliably. We calculated the diagram of leading instabilities for effective potentials to three different powers of the symmetry invariants ρ_m, ρ_d. In fig. 2 one can see on the left (a) a diagram containing only expansion terms to the order ρ^1, these are the bosonic mass terms, eq. (16). In the middle (b) the effective potential contains interactions up to ρ^2 or, in other words, up to quartic interactions for the fields. On the right (c) the diagram corresponds to the solution of the flow equation including terms up to order ρ^3. One can see a convergence of the phase diagrams for higher orders in ρ. We conclude that a truncation containing terms up to ρ^2 is sufficient to reliably calculate the phase diagram of the Hubbard model on the square lattice.

C. Transition between Commensurate and Incommensurate Antiferromagnetism

We examine the phase diagram of leading instabilities with a 16×16 resolution of the inverse magnetic propagator in the positive quadrant of the Brillouin zone $[0, \pi]^2$. We find that it can obtain two distinct minima. A minimum at (π, π) corresponds to commensurate antiferromagnetism. A fourfold degenerate minimum at $(\pi, \pi + \delta_{ic})$,$(\pi, \pi - \delta_{ic})$,$(\pi + \delta_{ic}, \pi)$ and $(\pi - \delta_{ic}, \pi)$ corresponds to incommensurate antiferromagnetism. There is no other possible minimum, that the inverse magnetic propagator can obtain.

Earlier zero temperature calculations suggest that at vanishing next-to-nearest neighbor hopping t' the incommensurability takes on the value $\tilde{\delta}_{ic}(t' = 0) = 2 \arcsin(|\frac{\mu}{2t}|)$[66]. Our results at the $t' = 0$ line of the phase diagram are larger by $\approx 15\%$, $\delta_{ic}(t' = 0) \approx 2.3 \arcsin(|\frac{\mu}{2t}|)$. Considering for example $\mu = -0.3$, the result of the earlier work is $\tilde{\delta}_{ic} = 0.301$, while our result is $\delta_{ic} \approx 0.35$, see fig. 4. Our values of $\delta_{ic}(t' = 0)$ are in agreement with earlier renormalization group studies [51].

The domains of commensurate antiferromagnetism and incommensurate antiferromagnetism are divided by a first order transition around the line $\mu \approx 4t'$. This

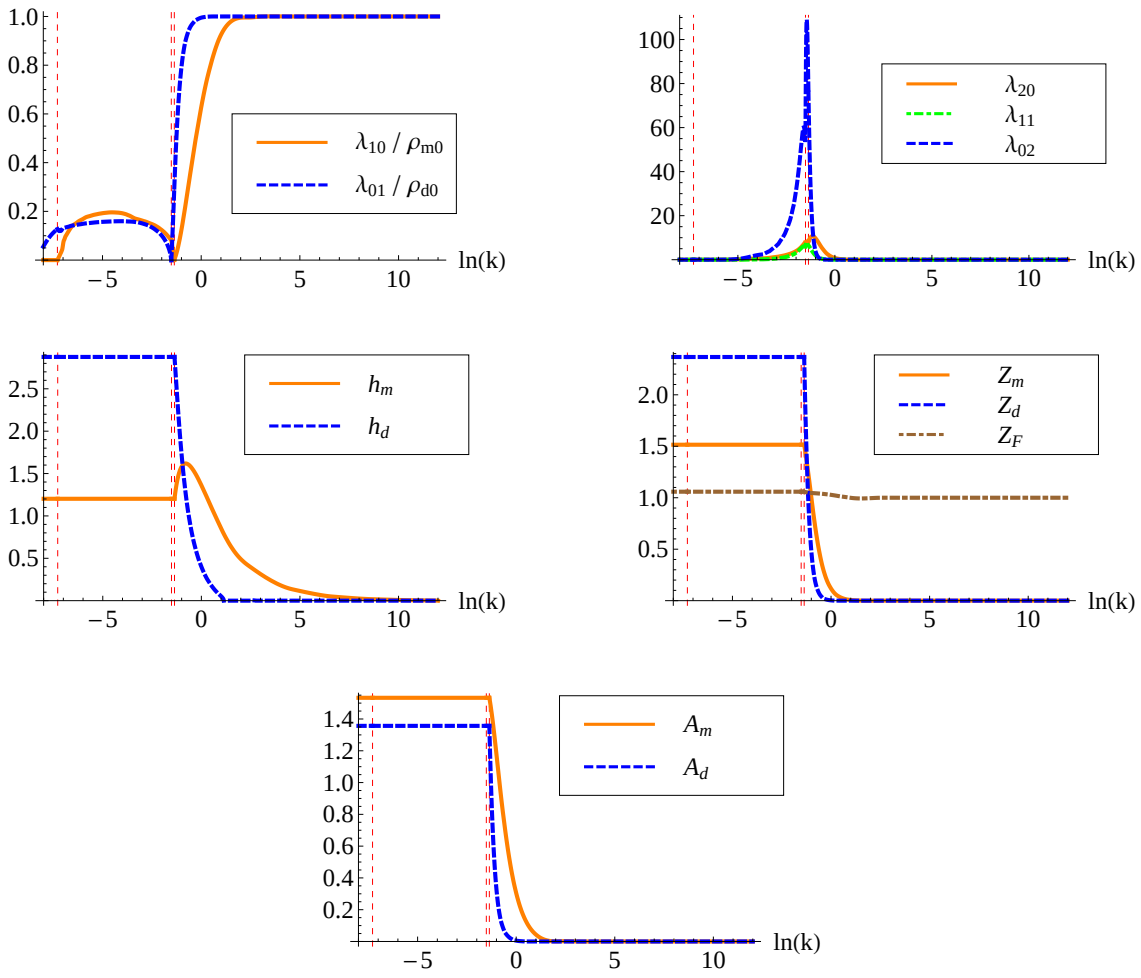

Figure 3: Solutions of the flow equation at $\mu = -0.4, t' = -0.1$ from $\ln(\Lambda) = \ln(k) = 12$ to $\ln(k) = -8$. The horizontal red dashed lines indicate transitions between symmetric and symmetry broken regimes. When the renormalization flow enters any symmetry broken phase, the flow of the masses $\lambda_{10}, \lambda_{01}$ is continued by the flow of the minima of the effective potential ρ_{m0}, ρ_{d0}. This particular flow trajectory visits all four possible regimes: the trajectory starts in the symmetric regime, then enters the antiferromagnetic regime, followed by a coexistence between d-wave and antiferromagnetism, afterwards the flow trajectory enters the d-wave regime. The minima of the effective potential ρ_{m0} and ρ_{d0} are rescaled $\times 20$.

line corresponds to Van-Hove filling. In fig. 4 (a)–(e) we closely examine the transition for fixed $\mu = -0.3$ between the two regions, starting in the commensurate region by lowering t'. Each picture is calculated with 32 sampling points along the axis in the Brillouin zone. In the commensurate domain $\delta_{ic}(t' > 0.075) = 0$ (a) the magnetic propagator begins to flatten as it approaches the transition. At $t' = 0.075$ the system undergoes a first order transition, where a new minimum emerges at a finite distance $\delta_{ic}(t' = 0.075) \approx 2/3 \times \delta_{ic}(t' = 0)$ from the commensurate minimum. This minimum moves away (b),(c),(d). Finally, it converges slowly to $\delta_{ic}(t' = 0) = 0.35$ (e). Also note that there is no significant change of the minimum of the propagator during the renormalization flow. This first order transition was also found in earlier calculations in [44], by examining the jump in the incommensurability δ_{ic}.

IV. RENORMALIZATION FLOW IN THE SYMMETRY BROKEN REGIMES

A. Deformation of the Fermi Surface

The Fermi surface in the antiferromagnetic broken phase changes its geometry according to the formula

$$Z_F^2 \xi(Q)\xi(Q+\Pi) - \Delta_a = 0 \ , \Delta_a = 2h_m(\Pi)^2 \rho_a \ . \qquad (23)$$

In fig. 5 one can see how a finite antiferromagnetic condensate deforms the Fermi surface, such that Fermi pockets emerge. In the symmetric phase $\Delta_a = 0$ the definition of the Fermi surface correctly reduces to $Z_F \xi(Q) = 0$. As the fermionic propagator at zero temperature diverges at the Fermi surface, it naturally occurs in all fermionic contributions to the flow equations in the denominator.

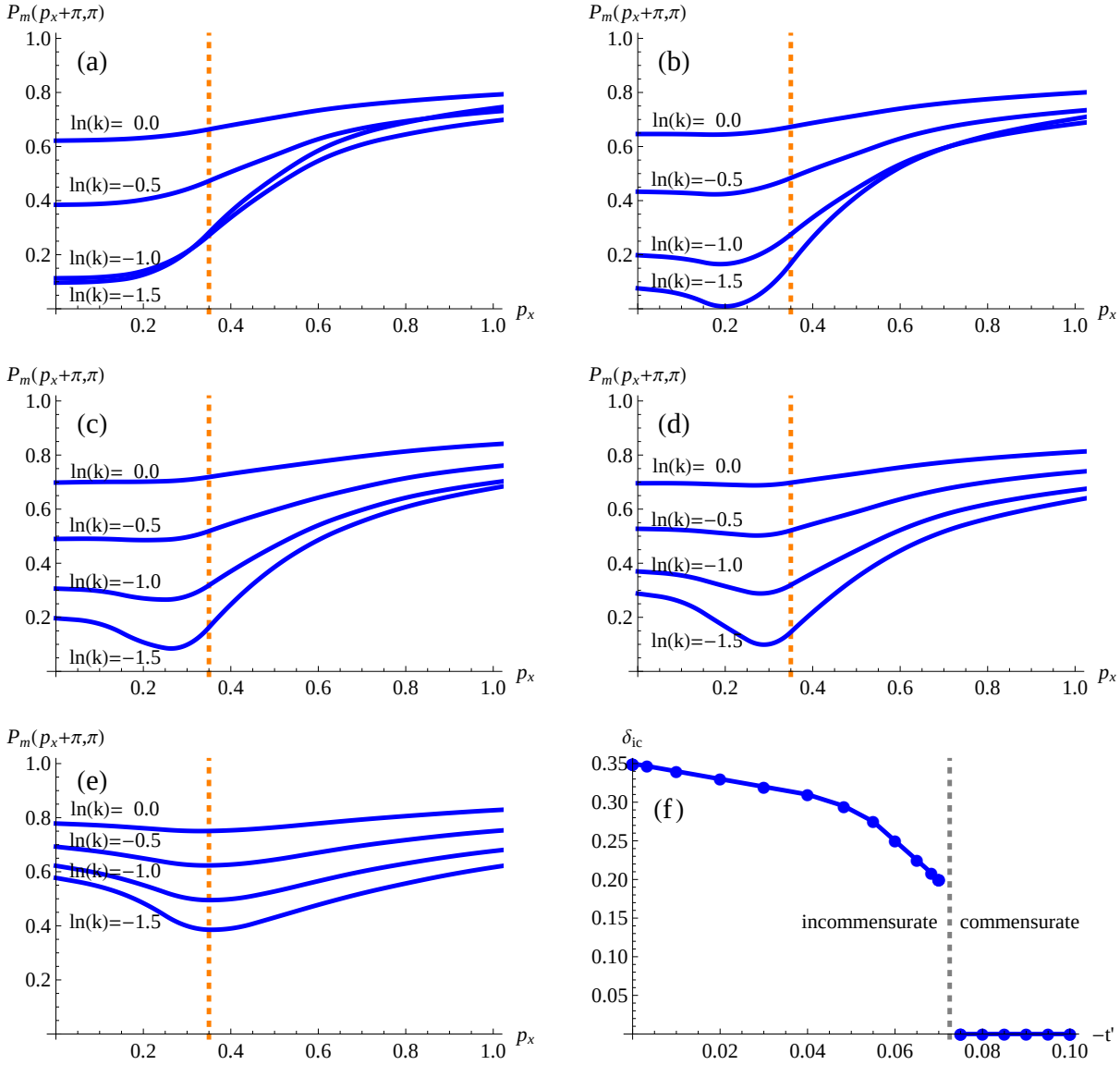

Figure 4: First order transition between incommensurate and commensurate antiferromagnetism at fixed $\mu = 0.3$. (a) $t'/t = -0.08$ (b) $t'/t = -0.07$ (c) $t'/t = -0.06$ (d) $t'/t = -0.05$ (e) $t'/t = -0.00$. The orange dotted line indicates the value of the incommensurate minimum at $t' = 0$. (f) describes the transition of the minimum under a change in t'.

We now consider the effects of the deformation of the Fermi surface on the flow equations. At finite temperature when there is a finite lowest fermionic Matsubara mode, the flow equations are still gapped at sufficiently large scales k in the magnetic symmetry broken phase. Practically this means that the flow equations in the symmetry broken regimes at finite temperature can be solved approximately by regularizing around the Fermi surface belonging to the symmetric regime $Z_F\xi(Q) = 0$. This is scheme is formulated by choosing the regulator, here the Litim regulator [59],

$$R_F(\xi(Q)) = Z_F \,\mathrm{sign}(\xi(Q))(k - |\xi(Q)|)\Theta(k - |\xi(Q)|) \,. \tag{24}$$

A major problem arises at zero temperature. The lowest fermionic Matsubara frequency is zero and can thus no longer introduce a gap in the fermionic propagator. At zero temperature the flow equations would diverge at the Fermi surface at any scale when employing the just introduced Litim regularization scheme. Thus, it is imperative to regularize around the correct Fermi surface (23). In order to capture the deformation of the Fermi surface in the flow equations, we introduce a set of regulators acting differently on different patches in the Brillouin zone,

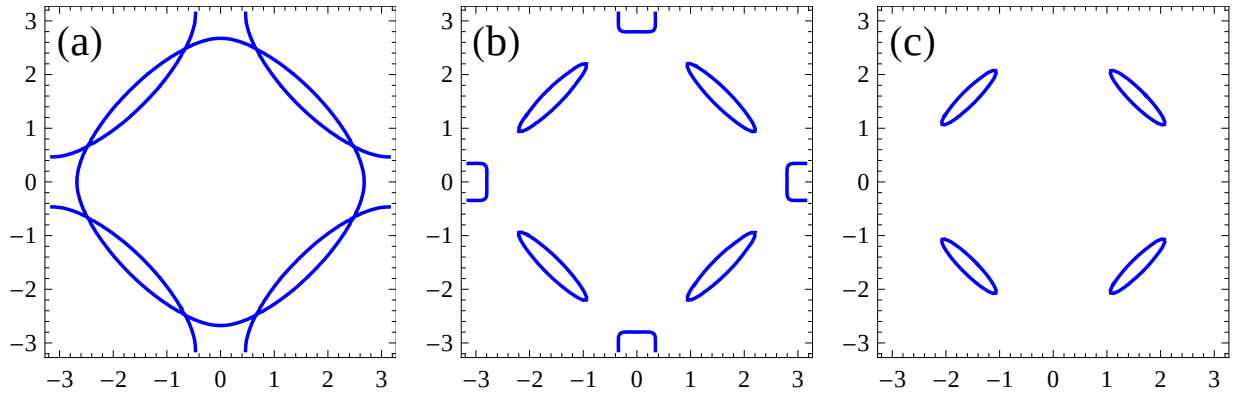

Figure 5: The Fermi surface changes with increasing antiferromagnetic condensate. For next-to-nearest neighbor hopping $t'/t = -0.2$, chemical potential $\mu/t = -0.5$, fermionic wave function renormalization $Z_F = 1$ and different sizes of the commensurate antiferromagnetic condensate (a) $\Delta_a = 0$, (b) $\Delta_a = 0.05$, (c) $\Delta_a = 0.1$.

fig. 6, and thus differently on $\xi(Q)$ and $\xi(Q+\Pi)$,

$$
\begin{aligned}
&R_F(\xi(Q)) = \\
&\theta(\pi - |[q_x]| - |[q_y]|)\, Z_F\, \mathrm{sign}(\xi(Q) - \frac{\Delta_a}{Z_F^2 \xi_k(Q+\Pi)}) \\
&\times (k - |\xi(Q) - \frac{\Delta_a}{Z_F^2 \xi_k(Q+\Pi)}|)\Theta(k - |\xi(Q) - \frac{\Delta_a}{Z_F^2 \xi_k(Q+\Pi)}|) \\
&+ \theta(-\pi + |[q_x]| + |[q_y]|) Z_F\, \mathrm{sign}(\xi(Q)) \\
&\qquad \times (k - |\xi(Q)|)\Theta(k - |\xi(Q)|)\ . \qquad (25)
\end{aligned}
$$

In appendix C we describe how this regulator ensures a gapped propagator in the vicinity of the Fermi surface. This is possible since in all flow equations the contributions from $\xi(Q)$ and $\xi(Q+\Pi)$ occur in pairs. In addition, this regulator ensures that one approaches the Fermi-surface with the correct sign.

While there are other regulators that are capable of capturing the Fermi surface, and in agreement with all requirements of regulators [58] there are more conditions to be fulfilled in order to obtain reliable physical results: (i) In the symmetric case and the microscopic limit the regulator must reduce to the free case, i.e. the regulator (25) reduces to the Litim regulator (24). We experimented with different regulators, not fulfilling this condition, which completely changed the results. (ii) The regulator must be the same in all regimes. It was shown that switching the regulator while solving the flow equations has a strong impact on the results and introduces severe non-physical artifacts [67]. (iii) The Fermi surface needs to be approached with the correct sign. (iv) The regulator may not contain any divergences itself. Our regulator (25) complies with all of these conditions.

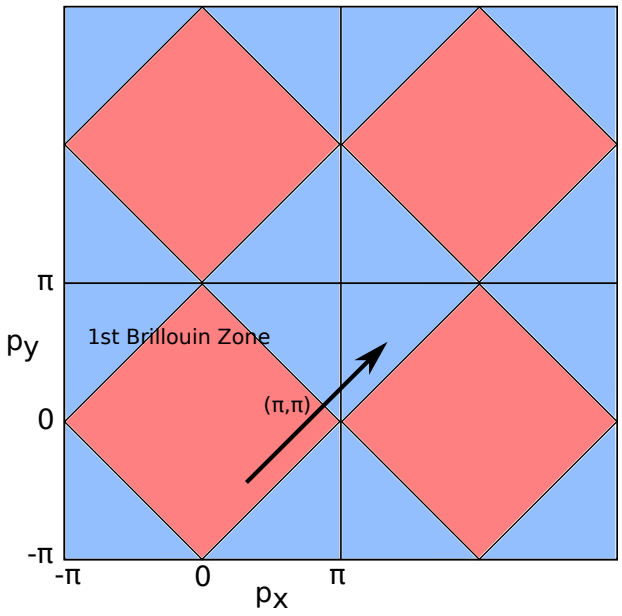

Figure 6: The regulator (25) acts differently on the red and blue patches of the Brillouin zones. A shift by $\Pi = (\pi, \pi)$ induces a switch in the regulator function.

B. Flow Diagram in the Symmetry Broken Regimes

In the symmetry broken phase we continued the flow of the effective potential including bosonic and fermionic fluctuations. The renormalization group flow is calculated within one consistent scheme of flow equations with which we also calculated the flow in the symmetric regime. The renormalization group flows of the wave function renormalizations Z_m, A_m, Z_d, A_d and the flows of the Yukawa couplings h_m, h_d are not continued. All flow equations are projected on commensurate antifer-

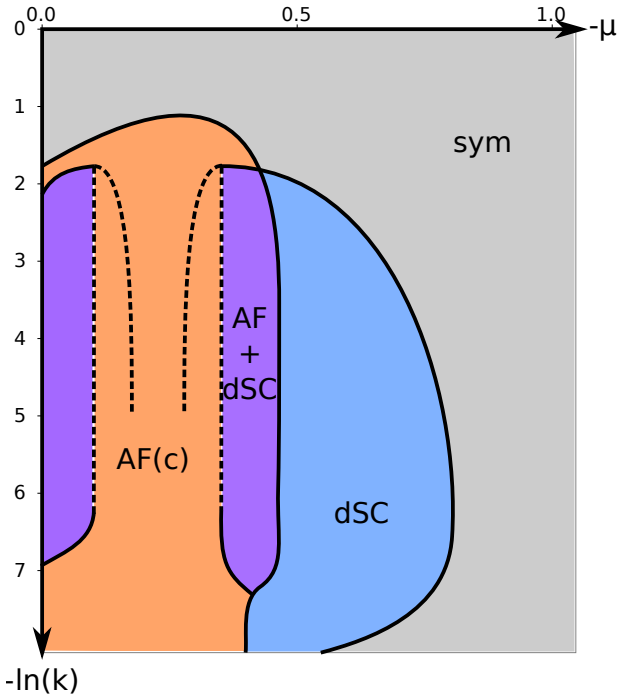

Figure 7: Flow diagram at $t' = -0.1$. This diagram collects solutions for the flow equations for each μ and shows the regimes of finite condensates in dependence of the scale k. In the limit $k \to 0$ or $\ln(k) \to -\infty$, the diagram shows the phases of the full effective action. Other scales $0 < k < 2\pi$ describe the physics in finite physical samples or local ordering. A coexistence phase is only present at intermediate scales.

romagnetism in the symmetry broken phases. We solve the flow equations from $t = \ln(k) = 12$ to $t = \ln(k) = -8$, where the numerical solution of the flow equation starts to break down.

In fig. 7 we summarize the solutions of the flow equations for fixed next-to-nearest neighbor hopping $t' = -0.1$ for various values of the chemical potential $-\mu \in [0, 1]$. This diagram corresponds to a collection of μ-phase diagrams for different sizes of physical samples. This flow diagram contains, in addition to antiferromagnetism and d-wave superconductivity, regions of coexistence of these two phases. However, the coexistence vanishes in the limit of small momentum scales k and thus large physical sample sizes. In other words, a coexistence phase in macroscopic physical systems can only be observed locally. The corresponding renormalization flow is shown in the flow of ρ_{m0} and ρ_{d0} of fig. 3. Here we see that both condensates emerge, however at scales $\ln(k) < -4$ they start to repel each other. Between the two coexistence regions in fig. 7 there is a region where only antiferromagnetism is present. While in the middle at $t' \approx -0.025$ the d-wave mass λ_{01} stays positive, there is a regime incapsulated by the dashed lines, in the vicinity of the coexistence phases, where this is not the case. Here the mass of the d-wave boson reaches a value of zero, however, the d-wave condensate is repelled immediately, such that the d-wave minimum ρ_d of effective potential always stays zero. Even though our calculation is not quantitatively accurate in the symmetry broken phase, the mechanism responsible for the mutual repellence is independent of the exact sizes of both condensates. That is why we expect the eradication of the coexistence phase at macroscopic scales to be a rather robust result.

La_2CuO_4 cuprates exhibit an orthorombic crystal structure with a very small asymmetry in lattice parameters in the CuO_4 plane $a \approx b \approx 5.4$Å and $c \approx 13.2$Å [68]. It is possible to quantitatively relate the renormalization group scale k approximately to physical scales. A scale of $k = 2\pi$ corresponds to the size of a Brillouin zone and can thus be related to the physics of the size of one unit cell. Applying this deduction to fig. 7 we find that superconductivity is most prominent at $\ln(k) \approx -7$ corresponding to a size of roughly 4×10^4Å. Furthermore, one can see that it is eradicated by allowing bosonic fluctuations at $\ln(k) < -9$ corresponding to $> 3 \times 10^5$Å.

C. On the Vanishing of Superconductivity

From fig. 7 one can infer that in the limit of $k \to 0$ the superconductivity vanishes. The question arises if this observation bears any physical relevance or whether it is a pitfall of our limited truncation in the symmetry broken phases. We believe the true physical phase diagram shows superconductivity for large sample sizes. There are two possible scenarios to explain our results. First, our model could be incomplete in a sense that there would be the need for an effect limiting the scale of physical fluctuations to a size of roughly 10^5Å in order to still obtain a d-wave superconductivity at macroscopic scales.

A second and much more likely scenario is that our limited truncation in the symmetry broken phases limits the emergence of d-wave superconductivity. The d-wave condensate emerges quadratically from a growing Yukawa coupling h_d, see eq. (B32). In the symmetry broken phases we do however not continue the flow of the Yukawa couplings. It is likely that h_d grows larger as a natural continuation of the flow in the symmetric phase, see fig. 3. As a consequence, the d-wave minimum ρ_{d0} would grow much larger than in our calculation and in turn could not be so easily destroyed by bosonic fluctuations.

While we cannot continue the renormalization flow of the d-wave Yukawa coupling, we examined if it would in principle be possible to obtain a non-vanishing d-wave condensate at even smaller scales k. We observe that the fermionic contributions, in the superconducting regime, always enhance the d-wave minimum ρ_{d0}, while only the bosonic fluctuations can reduce it.

If we, for the moment, consider a sub-theory containing only bosonic fields and neglecting all fermionic contributions, this theory would, in the limit $k \to 0$, become an effective three-dimensional statistical field theory. In accordance with the Mermin-Wagner theorem this the-

ory can have a phase transition. Since fermionic contributions can only enhance the d-wave minimum ρ_{d0}, we conclude that the full theory containing bosonic and fermionic fields could in principle exhibit a phase transition.

The main effect of h_d on the effective potential is the enhancement of the d-wave minimum of the effective potential ρ_{d0}. In fig. 8 we compare the flow of the minimum ρ_{d0} for different initial values at the symmetry breaking scale. One can see that in principle there exists a parameter range such that it is possible to be in the symmetry broken phase for an arbitrary finite scale k. However, the bosonic fluctuations tend to strongly reduce the d-wave condensate at any scale $\ln(k) < -7$.

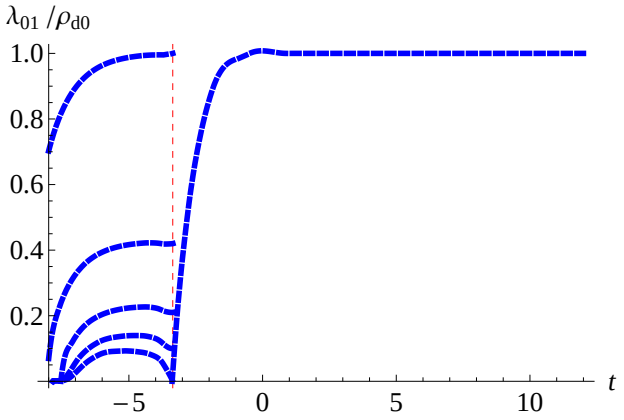

Figure 8: Renormalization flow at $\mu = -0.66$, $t' = -0.1$ of the superconducting mass λ_{01}, which is continued with the flow of the superconducting minimum ρ_{d0} below the symmetry breaking scale k_{SB}. The initial values of ρ_{d0} are artificially enhanced at k_{SB}, from bottom to top: $\rho_{d0} = 0$, $\rho_{d0} = 0.001$, $\rho_{d0} = 0.002$, $\rho_{d0} = 0.004$, $\rho_{d0} = 0.01$. The superconducting minimum of the effective potential ρ_{d0} is rescaled $\times 100$.

V. CONCLUSION

In this article we explored the $\mu - t'$−phase diagram of the Hubbard model on the square lattice at zero temperature. A calculation in the symmetric phase revealed phases of commensurate antiferromagnetism, incommensurate antiferromagnetism and d-wave superconductivity, see fig. 1. By comparing different truncations of the effective potential we examined the robustness of the results, see fig. 2. Furthermore, our truncation allows for the continuation of the renormalization flow into the symmetry broken regimes, see fig. 3. We examined the transition between commensurate and incommensurate antiferromagnetism in detail. We find that it coincides with the Van-Hove line and is of first order, see fig. 4. A nonzero antiferromagnetic condensate induces a continuous deformation of the Fermi surface, see fig. 5. Our regularization scheme (25) allows us to properly include fluctuations around the changing Fermi surface. Calculations in the symmetry broken regimes reveal a coexistence of antiferromagnetism and d-wave superconductity only on intermediate scales, see fig. 7. Beyond that, our results suggest that these phases have a tendency to repel each other. This mechanism leads to an eradication of the coexistence phase at macroscopic scales. A weakness of our calculations is the inability to decide the fate of the d-wave superconductivity at macroscopic scales, this can be resolved in the future by continuing the flow of the Yukawa couplings in the symmetry broken phases.

Acknowledgments - We would like to thank Jan M. Pawlowski, Manfred Salmhofer, Michael M. Scherer, Kambis Veschgini and Christof Wetterich for useful discussions. We further thank Mathias Neidig and Shirin Nkongolo for proofreading the manuscript. S.W. acknowledges support by the Heidelberg Graduate School of Fundamental Physics.

Appendix A: DEFINITIONS

1. Pauli Matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{A1}$$

2. Antisymmetric Matrix

$$\epsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \ , \ \epsilon^2 = -\mathbb{1} \tag{A2}$$

3. Dual Matrix

If a matrix A can be written in terms of Pauli Matrices $A = a\mathbb{1} + \vec{b}\vec{\sigma}$ and four real parameters a, b_1, b_2, b_3, its determinant evaluates to

$$\begin{aligned} \det(a\mathbb{1} + \vec{b}\vec{\sigma}) &= \begin{pmatrix} a+b_3 & b_1 - ib_2 \\ b_1 + ib_2 & a - b_3 \end{pmatrix} = a^2 - b_1^2 - b_2^2 - b_3^2 \\ &= \begin{pmatrix} a-b_3 & -b_1 + ib_2 \\ -b_1 - ib_2 & a + b_3 \end{pmatrix} = \det(a\mathbb{1} - \vec{b}\vec{\sigma}) \ , \end{aligned} \tag{A3}$$

such that we can define a dual matrix $A' = a\mathbb{1} - \vec{b}\vec{\sigma}$ with the same determinant. In the special case $\det A = 1$, it follows $A \in SU(2)$.

Appendix B: FLOW EQUATIONS

The derivation of the flow equation employs the flow equation of the effective average action, eq. (3).

1. Bosonic Contributions to the Effective Potential

The flow equation of the effective potential can be decomposed into bosonic and fermionic contributions $\partial_k U = (\partial_k U)_F + (\partial_k U)_B$. The bosonic contributions correspond to the flow equations of a bosonic $\mathrm{O}(2) \times \mathrm{O}(3)$ model. They can be obtained in a straightforward manner by calculating the second field derivative of the effective action

$$\Gamma^{(2)}_{BB} + R_{BB} = \begin{pmatrix} R_m & 0 & 0 & I_{md} & 0 \\ 0 & G_m & 0 & 0 & 0 \\ 0 & 0 & G_m & 0 & 0 \\ I_{md} & 0 & 0 & R_d & 0 \\ 0 & 0 & 0 & 0 & G_d \end{pmatrix} , \tag{B1}$$

which contains radial modes R_m, R_d, Goldstone modes G_m, G_d and exchange modes I_{md}

$$\begin{aligned} R_m &= P_m + 2\rho_m U^{2,0} + U^{(1,0)} + R_B \\ G_m &= P_m + U^{(1,0)} + R_B \\ R_d &= P_d + 2\rho_d U^{0,2} + U^{(0,1)} + R_B \\ G_d &= P_d + U^{(0,1)} + R_B \\ I_{md} &= 2\sqrt{\rho_m \rho_d} U^{(1,1)} \\ \det{}_{BB} &= R_m R_d - I_{md}^2 \ . \end{aligned} \tag{B2}$$

Inverting this matrix

$$(\Gamma^{(2)}_{BB} + R_{BB})^{-1} = \begin{pmatrix} \frac{R_d}{\det_{BB}} & 0 & 0 & -\frac{I_{md}}{\det_{BB}} & 0 \\ 0 & G_m^{-1} & 0 & 0 & 0 \\ 0 & 0 & G_m^{-1} & 0 & 0 \\ -\frac{I_{md}}{\det_{BB}} & 0 & 0 & \frac{R_m}{\det_{BB}} & 0 \\ 0 & 0 & 0 & 0 & G_d^{-1} \end{pmatrix} , \tag{B3}$$

immediately leads to the flow equation for the effective potential of an $\mathrm{O}(2) \times \mathrm{O}(3)$ model

$$\begin{aligned} &(\partial_k U_k)_B = \\ &\frac{1}{2} \sum_Q \left(\frac{R_m}{det_{BB}} + 2G_m^{-1} + \frac{R_d}{det_{BB}} + G_d^{-1} \right) \partial_k R_B \ . \end{aligned} \tag{B4}$$

2. Fermionic Contributions to the Effective Potential

The fermionic contribution to the effective potential

$$\begin{aligned} (\partial_k U)_F &= \frac{1}{2} \tilde{\partial}_k \mathrm{STr} \ln \Gamma^{(2)}_{FF,k}(Q, Q') \\ &= -\frac{1}{2} \tilde{\partial}_k \sum_{Q,Q'} \ln \det \Gamma^{(2)}_{FF,k}(Q, Q') \ , \end{aligned} \tag{B5}$$

is much harder to obtain due to the momentum shift by $\Pi = (0, \pi, \pi)$ in the antiferromagnetic channel. The fermionic part of the second field derivative of the effective action

$$\begin{aligned} (\Gamma^{(2)}_{FF} + R_{FF})(Q, Q') &= \begin{pmatrix} -h_d \epsilon (d_1 - i d_2) & h_m(\Pi) \vec{\sigma} \cdot \vec{m} \\ -h_m(\Pi) \vec{\sigma} \cdot \vec{m} & h_d \epsilon (d_1 + i d_2) \end{pmatrix} \\ + \begin{pmatrix} 0 & -\mathbb{1}(P_F(-Q) + R_F(-Q)) \\ \mathbb{1}(P_F(-Q) + R_F(-Q)) & 0 \end{pmatrix} \end{aligned} \tag{B6}$$

contains contributions from the fermionic propagator and the Yukawa couplings $h_m(\Pi), h_d$. It can be simplified by identifying similar terms

$$\begin{aligned} \Gamma^{(2)}_{FF,k}(Q, Q') &= (\Gamma^{(2)}_{FF} + R_{FF})(Q, Q') = \\ &\begin{pmatrix} \tilde{B}(Q) \delta(Q - Q') & -A^T(-Q, Q') \\ A(Q, Q') & B(Q) \delta(Q - Q') \end{pmatrix} , \end{aligned} \tag{B7}$$

where

$$\begin{aligned} A(Q, Q') &= Z_F (i\omega_Q + \xi_k(Q)) \delta(Q - Q') \\ &\quad - \underbrace{h_m(\Pi) \, \vec{m}}_{\vec{M}} \cdot \vec{\sigma} \, \delta(\Pi - Q + Q') \\ B(Q) &= h_d \, \epsilon \, (d_1 + i d_2) \, f_d(Q) = D(Q) \epsilon \\ \tilde{B}(Q) &= -h_d \, \epsilon \, (d_1 - i d_2) \, f_d(Q) = -\tilde{D}(Q) \epsilon \ . \end{aligned} \tag{B8}$$

A straightforward calculation that employs doubling the matrix in the argument of the determinant and employ-

ing its dual matrix (A3) leads to

$$\begin{aligned}
&\det \Gamma^{(2)}_{FF,k}(Q,Q') \\
&= \det\left(\Gamma^{(2)}_{FF,k}(Q,Q'')\begin{pmatrix}0 & \mathbb{1}\\ \mathbb{1} & 0\end{pmatrix}\Gamma^{(2)}_{FF,k}(-Q'',-Q')\begin{pmatrix}0 & \mathbb{1}\\ \mathbb{1} & 0\end{pmatrix}\right)^{1/2} \\
&= \det\left(\tilde{B}(Q)B(-Q)\delta(Q-Q') + A(Q,Q'')A(-Q'',-Q')\right) \\
&= \det\Big(\underbrace{\left(\tilde{D}(Q)D(Q) + Z_F^2\omega^2 + \xi_k(Q)^2 + \vec{M}^2\right)}_{a(Q)}\delta(Q-Q') \\
&\quad + \underbrace{(\xi_k(Q)+\xi_k(Q+\Pi))\vec{M}}_{\vec{b}(Q)}\cdot\vec{\sigma}\delta(Q-Q'+\Pi)\Big) \\
&= \det\Big(\left(a(Q)\delta(Q-Q'') - \vec{b}(Q)\cdot\vec{\sigma}\delta(Q-Q''+\Pi)\right) \\
&\quad \times \left(a(Q'')\delta(Q''-Q') + \vec{b}(Q'')\cdot\vec{\sigma}\delta(Q''-Q'+\Pi)\right)\Big)^{1/2} \\
&= \det\left((a(Q)a(Q+\Pi) - \vec{b}(Q)\cdot\vec{b}(Q))\delta(Q-Q')\right)^{1/2} ,
\end{aligned} \tag{B9}$$

where we have used symmetry properties of the fermionic kinetic term $\xi_k(Q) = \xi_k(-Q)$ and the d-wave form factor $f_d(Q) = \frac{1}{2}(\cos(q_x) - \cos(q_y)) = -f_d(Q+\Pi)$. Plugging all together and solving for ω yields

$$\det \Gamma^{(2)}_{FF,k}(Q,Q') = \det\left(J_+\, J_-\,\delta(Q-Q')\right)^{1/2} , \tag{B10}$$

where

$$\begin{aligned}
J_\pm &= Z_F^2\omega^2 + \tilde{D}(Q)D(Q) + \Big(\frac{Z_F}{2}\big((\xi_k(Q)+\xi_k(Q+\Pi)\big) \\
&\qquad \pm \sqrt{\vec{M}^2 + \frac{Z_F^2}{4}(\xi_k(Q)-\xi_k(Q+\Pi))^2}\ \Big)^2 \\
&= Z_F^2\omega^2 + 2h_d^2 f_d(Q)\rho_d + \Big(\frac{Z_F}{2}\big((\xi_k(Q)+\xi_k(Q+\Pi)\big) \\
&\qquad \pm \sqrt{2h_m(\Pi)^2\rho_m + \frac{Z_F^2}{4}(\xi_k(Q)-\xi_k(Q+\Pi))^2}\ \Big)^2 .
\end{aligned} \tag{B11}$$

The sum over Q' can be performed trivially

$$\sum_{Q,Q'} \ln\det \Gamma^{(2)}_{FF,k}(Q,Q') = \sum_Q \frac{1}{2}(\ln J_+ + \ln J_-) , \tag{B12}$$

which concludes the calculation of the fermionic contribution to the flow equation of the effective potential

$$(\partial_k U)_F = -\frac{1}{2}\tilde{\partial}_k \sum_Q \frac{1}{2}(\ln J_+ + \ln J_-) . \tag{B13}$$

3. On Matsubara Integrals

While the Matsubara integrals for higher order couplings can be integrated in a straightforward manner, the Matsubara integral for the effective potential itself evaluates to ∞. Since we are not interested in finite shifts of the overall energy minimum, we extract the finite part which depends on ρ_m, ρ_d via

$$\begin{aligned}
&\int_{-\infty}^{\infty} d\omega \ln(Z_F^2\omega^2 + A^2) \\
&= \int_{-\infty}^{\infty} d\omega \underbrace{\ln(Z_F^2\omega^2)}_{\infty} + \ln(1 + \frac{A^2}{Z_F^2\omega^2}) .
\end{aligned} \tag{B14}$$

The finite contribution can then be safely integrated

$$\begin{aligned}
&= \int_{-\infty}^{\infty} d\omega \ln(1 + \frac{A^2}{Z_F^2\omega^2}) \\
&= \omega \ln(1 + \frac{A^2}{Z_F^2\omega^2}) + \frac{2A}{Z_F}\arctan(\frac{Z_F\omega}{A})\Big|_{-\infty}^{\infty} \\
&= 2\pi\frac{A}{Z_F} .
\end{aligned} \tag{B15}$$

4. Flow Equations for the Bosonic Couplings

We expand the effective potential (15) around its minimum, thus we need to adjust the flow equations for a change in the minimum $\partial_k\rho_m, \partial_k\rho_d$.

$$\begin{aligned}
\partial_k\lambda_{i,j} &= \partial^i_{\rho_m}\partial^j_{\rho_d}\partial_k U(\rho_m,\rho_d)\big|_{\rho_{m0},\rho_{d0}} \\
&= \left(\partial^i_{\rho_m}\partial^j_{\rho_d}\partial_k U(\rho_m,\rho_d)\right)\big|_{\rho_{m0},\rho_{d0}} \\
&\quad + \left(\partial^{i+1}_{\rho_m}\partial^j_{\rho_d}U(\rho_m,\rho_d)\right)|_{\rho_{m0},\rho_{d0}}\partial_k\rho_{m0} \\
&\quad + \left(\partial^i_{\rho_m}\partial^{j+1}_{\rho_d}U(\rho_m,\rho_d)\right)|_{\rho_{m0},\rho_{d0}}\partial_k\rho_{d0}
\end{aligned} \tag{B16}$$

The flow equations of the minima of the effective potential are in the magnetic broken phase

$$\partial_k\rho_{m0} = -\frac{\partial_{\rho_m}(\partial_k U)}{\lambda_{20}}|_{\rho_{m0},\rho_{d0}} , \tag{B17}$$

in the d-wave broken phase

$$\partial_k\rho_{d0} = -\frac{\partial_{\rho_d}(\partial_k U)}{\lambda_{02}}|_{\rho_{m0},\rho_{d0}} , \tag{B18}$$

and in the regimes where both symmetries are broken

$$\begin{aligned}
\partial_k\rho_{m0} &= -\frac{\partial_{\rho_m}(\partial_k U)\lambda_{02} - \partial_{\rho_d}(\partial_k U)\lambda_{11}}{\lambda_{20}\lambda_{02} - \lambda_{11}^2}|_{\rho_{m0},\rho_{d0}} \\
\partial_k\rho_{d0} &= -\frac{\partial_{\rho_d}(\partial_k U)\lambda_{20} - \partial_{\rho_m}(\partial_k U)\lambda_{11}}{\lambda_{20}\lambda_{02} - \lambda_{11}^2}|_{\rho_{m0},\rho_{d0}} .
\end{aligned} \tag{B19}$$

5. Flowing Bosonization

Flowing bosonization induces a scale-dependent Hubbard-Stratonovich transformation [56, 57]. In the present work we translate all diverging momentum channels of the Hubbard action U to Yukawa interactions me-

diated by magnetic and d-wave bosons. To this purpose we derive exemplary the flowing bosonization in magnetic channel. In $\Gamma^{\mathrm{FHS}}_{m,k}$ we collect all couplings involved in the flowing bosonization.

$$\begin{aligned}
&\Gamma^{\mathrm{FHS}}_{m,k}\\
=&\frac{1}{2}\sum_Q \vec{m}^T(-Q)\left(P_m(Q)+m_m^2\right)\vec{m}(Q)\\
&-\sum_{Q_1,Q_2,Q_3} h_m(Q)\vec{m}(Q_1)\cdot\left(\psi^\dagger(Q_2)\vec{\sigma}\psi(Q_3)\right)\\
&\quad\times\delta(Q_1-Q_2+Q_3)\\
&-\frac{1}{2}\sum_{Q_1,\dots,Q_4}\lambda_F^m(Q_1-Q_2)\delta(Q_1-Q_2+Q_3-Q_4)\\
&\quad\times\left(\psi^\dagger(Q_1)\vec{\sigma}\psi(Q_2)\right)\left(\psi^\dagger(Q_3)\vec{\sigma}\psi(Q_4)\right) \qquad \text{(B20)}
\end{aligned}$$

We introduce an artificial scale-dependent bilinear field

$$\tilde{\vec{m}}(P)_k=\sum_Q\left(\psi^\dagger(Q)\vec{\sigma}\psi(P+Q)\right)\ , \qquad \text{(B21)}$$

whose scale-dependence is chosen such that

$$\partial_k\vec{m}(Q)=\alpha_k^m(Q)\tilde{\vec{m}}(Q)\ . \qquad \text{(B22)}$$

The function $\alpha_k^m(Q)$ will be specified later. Then we can rewrite $\Gamma^{\mathrm{FHS}}_{m,k}$ in terms of the artificial bilinear

$$\begin{aligned}
&\Gamma^{\mathrm{FHS}}_{m,k}\\
=&\frac{1}{2}\sum_Q \vec{m}^T(-Q)\left(P_m(Q)+m_m^2\right)\vec{m}(Q)\\
&-\sum_Q h_m(Q)\vec{m}^T(-Q)\tilde{\vec{m}}(Q)\\
&-\frac{1}{2}\sum_Q\lambda_F^m(Q)\tilde{\vec{m}}^T(-Q)\tilde{\vec{m}}(Q)\ . \qquad \text{(B23)}
\end{aligned}$$

The flow equation of the effective action obtains additional terms which arise from the scale-dependent bilinear

$$\begin{aligned}
\partial_k\Gamma_k=&\partial_k\Gamma_k|_{m_k}\\
&+\sum_Q\alpha_k^m(Q)\vec{m}^T(-Q)\left(P_m(Q)+m_m^2\right)\tilde{\vec{m}}(Q)\\
&-\sum_Q\alpha_k^m(Q)h_m(Q)\tilde{\vec{m}}^T(-Q)\tilde{\vec{m}}(Q)\ . \qquad \text{(B24)}
\end{aligned}$$

The flow equations for the couplings h_m and λ_m^F obtain additional terms

$$\begin{aligned}
\partial_k h_m(Q)&=\partial_k h_m(Q)|_{m_k}-\alpha_k^m(Q)\left(P_m(Q)+m_m^2\right)\\
\partial_k\lambda_F^m(Q)&=\partial_k\lambda_F^m(Q)|_{m_k}+2\alpha_m(Q)h_m(Q)\overset{!}{=}0\ . \qquad \text{(B25)}
\end{aligned}$$

By choosing $\alpha_k^m(Q)$ such that the flow of λ_m^F becomes zero, we induce an additional contribution to the flow of the Yukawa coupling h_m. A similar deduction can be done for the d-wave channel. Thus we arrive at modified flow equations

$$\begin{aligned}
\partial_k h_m(Q)&=\partial_k h_m(Q)|_{m_k}+\frac{P_m(Q)+m_m^2}{2h_m(Q)}\partial_k\lambda_F^m(Q)|_{m_k}\\
\partial_k h_d(Q)&=\partial_k h_d(Q)|_{d_k}+\frac{P_d(Q)+m_d^2}{2h_d(Q)}\partial_k\lambda_F^d(Q)|_{d_k}\ .
\end{aligned} \qquad \text{(B26)}$$

6. Flow Equations for Yukawa Couplings

The flow equations for h_m and h_d consist of a direct contribution and an indirect contribution arising from flowing bosonization, eq. (B26). In the latter case the flow of λ_F^m,λ_F^d is transformed into a contribution to the corresponding Yukawa coupling. When deriving the flow equations for λ_F^m and λ_F^d, there arises an ambiguity in choosing the external momenta, we choose $L=(0,\pi,0)$ and $L'(0,0,\pi)$ in order to evaluate our contributions close to the Fermi surface.

$$\begin{aligned}
\partial_k\lambda_F^m&=\frac{1}{3}\partial_k\Gamma^{(4)}_{F,ph}(L,L',-L,-L')\\
\partial_k\lambda_F^d&=\frac{1}{2}\Big(\partial_k\Gamma^{(4)}_{F,pp}(L,L,-L,-L)\\
&\qquad-\partial_k\Gamma^{(4)}_{F,pp}(L,L',-L,-L')\Big)
\end{aligned} \qquad \text{(B27)}$$

As a real boson the magnetic channel collects all contributions arising from particle-particle diagrams. The d-wave boson describes Cooper pairs, which is why we collect the particle-particle diagrams in this channel. The prescription of how to extract the contributions to the d-wave coupling λ_F^d was motivated in [50].

a. Magnetic Yukawa Coupling

$$
\begin{aligned}
\partial_k h_m(\Pi)^2 = \tilde{\partial}_k \sum_Q \\
&\frac{1}{6}\lambda_{10}\Big(\frac{8h_d^2 F_{d,k}(L+Q/2)F_{d,k}(L'+Q/2)U(\omega^2-\xi_k(L+Q)\xi_k(L'+Q))}{Z_F^2(Z_d\omega^2+F_{d,k}(Q))(\omega^2+\xi_k(L+Q)^2)(\omega^2+\xi_k(L'+Q)^2)} \\
&\quad + \frac{2U^2(\omega^2-\xi_k(Q)\xi_k(Q+\Pi))}{Z_F^2(\omega^2+\xi_k(Q)^2)(\omega^2+\xi_k(Q+\Pi)^2)} \\
&\quad + \frac{8h_d^4 F_{d,k}(L+Q/2)F_{d,k}(L'+Q/2)F_{d,k}(\Pi/2+Q/2)(\omega^2-\xi_k(L+Q)\xi_k(L'+Q))}{Z_F^2(Z_d\omega^2+F_{d,k}(Q))(Z_d\omega^2 F_{d,k}(\Pi+Q))(\omega^2+\xi_k(L+Q)^2)(\omega^2+\xi_k(L'+Q)^2)} \\
&\quad - \frac{6h_m(Q+L)^2 h_m(Q+L')^2(\omega^2-\xi_k(Q)\xi_k(Q+\Pi))}{Z_F^2(Z_m\omega^2+F_{m,k}(L+Q))(Z_m\omega^2+F_{m,k}(L'+Q))(\omega^2+\xi_k(Q)^2)(\omega^2+\xi_k(Q+\Pi)^2)}\Big) \\
&+ \frac{4h_m(\Pi)^2 h_d^2 F_{d,k}(L+Q/2)F_{d,k}(L'+Q/2)(\omega^2-\xi_k(L+Q)\xi_k(L'+Q))}{Z_F^2(Z_d\omega^2+F_{d,k}(Q))(\omega^2+\xi_k(L+Q)^2)(\omega^2+\xi_k(L'+Q)^2)} \\
&- \frac{2h_m(\Pi)^2 h_m(Q)^2(\omega^2-\xi_k(L+Q)\xi_k(L'+Q))}{Z_F^2(Z_m\omega^2+F_{m,k}(Q))(\omega^2+\xi_k(L+Q)^2)(\omega^2+\xi_k(L'+Q)^2)} \\
&+ \frac{2h_m(\Pi)U(\omega^2-\xi_k(Q)\xi_k(Q+\Pi))}{Z_F^2(\omega^2+\xi_k(Q)^2)(\omega^2+\xi_k(Q+\Pi)^2)}
\end{aligned}
\tag{B28}
$$

b. D-Wave Yukawa Coupling

$$
\begin{aligned}
\partial_k h_d^2 = \tilde{\partial}_k \sum_Q \\
&\frac{6h_d^2 F_{d,k}(Q)h_m(Q+L)^2}{Z_F^2(Z_m\omega^2+F_{m,k}(L+Q)(\omega^2+\xi_k(Q)^2))} \\
&+ \frac{1}{4}\lambda_{01}\frac{9h_m(Q+L)^4}{Z_F^2(Z_m\omega^2+F_{m,k}(L+Q))^2(\omega^2+\xi_k(Q)^2)} \\
&- \frac{1}{4}\lambda_{01}\frac{9h_m(Q+L)^2 h_m(Q+L')^2}{Z_F^2(Z_m\omega^2+F_{m,k}(L+Q))(Z_m\omega^2+F_{m,k}(L'+Q))(\omega^2+\xi_k(Q)^2)}
\end{aligned}
\tag{B29}
$$

7. Flow Equation for the Momentum-Dependent Propagators

The spatial momentum curvature at the minimum of the propagator A_m, A_d can be deduced by solving the flow equation for the momentum dependent propagators P_m, P_d. A straightforward method to obtain the flow of A_m, A_d is a derivative projection. However, we employ a finite difference projection, because it turned out to enhance the stability in our numerical calculations. We chose to evaluate the propagator at a finite distance from the minimum $P = (0, p, 0)$, with $p = 0.5$. The results are practically independent of the choice of p as long as it is $0 \ll p \ll \pi$.

$$
\begin{aligned}
A_m &= \left(P_m((0,\pi+p,\pi)) - P_m((0,\pi,\pi))\right)/p^2 \\
A_d &= \left(P_d((0,p,0)) - P_d((0,0,0))\right)/p^2
\end{aligned}
\tag{B30}
$$

$$
\begin{aligned}
\partial_k P_m(P) = \tilde{\partial}_k \sum_Q \frac{1}{4\pi} \\
&\Big(\frac{5\lambda_{20}}{Z_m\omega^2+F_{m,k}(Q)} + \frac{2\lambda_{11}}{Z_d\omega^2+F_{d,k}(Q)} \\
&- \frac{4h_m^2(P)(\omega^2-\xi_k(Q)\xi_k(Q+P))}{Z_F^2(\omega^2+\xi_k(Q)^2)(\omega^2+\xi_k(Q+P)^2)}\Big)
\end{aligned}
\tag{B31}
$$

$$\partial_k P_d(P) = \tilde{\partial}_k \sum_Q \frac{1}{4\pi} \Big(\frac{3\lambda_{11}}{Z_m\omega^2 + F_{m,k}(Q)} + \frac{4\lambda_{02}}{Z_d\omega^2 + F_{d,k}(Q)} - \frac{4h_d^2 F_{d,k}(Q+P/2)^2(\omega^2 + \xi_k(Q)\xi_k(Q+P))}{Z_F^2(\omega^2+\xi_k(Q)^2)(\omega^2+\xi_k(Q+P)^2)} \Big) \quad \text{(B32)}$$

$$\partial_k Z_F = \tilde{\partial}_k \sum_Q \frac{1}{8\pi\omega_p} \Big(\frac{6h_m(Q)^2(\omega+\omega_p)}{Z_F(Z_m\omega^2 + F_{m,k}(Q))((\omega+\omega_p)^2 + \xi_k(Q))} - \frac{6h_m(Q)^2(\omega-\omega_p)}{Z_F(Z_m\omega^2 + F_{m,k}(Q))((\omega-\omega_p)^2 + \xi_k(Q))} + \frac{4h_d^2 f(Q/2)^2(\omega+\omega_p)}{Z_F(Z_d\omega^2 + F_{d,k}(Q))((\omega+\omega_p)^2 + \xi_k(Q)^2)} - \frac{4h_d^2 f(Q/2)^2(\omega-\omega_p)}{Z_F(Z_d\omega^2 + F_{d,k}(Q))((\omega-\omega_p)^2 + \xi_k(Q)^2)} \Big) \quad \text{(B36)}$$

8. Flow Equations for Wave Function Renormalizations

The flow equations for the wave function renormalizations Z_m, Z_d, Z_F of the bosons and fermions can be extracted from the flow equations of the corresponding propagator. For this purpose we evaluate the flow of the propagators at a finite momentum $P = (\omega_p, 0, 0)$ from the minimum of the corresponding propagator. We choose $\omega_p = 0.5$.

$$\begin{aligned} Z_m &= \left(P_m((\omega_p,\pi,\pi)) - P_m((0,\pi,\pi))\right)/\omega_p^2 \\ Z_d &= \left(P_d((\omega_p,0,0)) - P_d((0,0,0))\right)/\omega_p^2 \\ Z_F &= \left(P_F((\omega_p,0,0)) - P_F((0,0,0))\right)/(i\omega_p) \end{aligned} \quad \text{(B33)}$$

In the incommensurate case the flow equations are evaluated at the minimum of the magnetic propagator $(0, \pi + \delta_{ic}, \pi)$. This definition leads to the corresponding flow equations

$$\partial_k Z_m = \tilde{\partial}_k \sum_Q \frac{1}{4\pi\omega_p^2} \Big(\frac{4h_m(\Pi)^2(\omega(\omega+\omega_p) - \xi_k(Q)\xi_k(Q+\Pi))}{Z_F^2(\omega^2+\xi_k(Q)^2)((\omega+\omega_p)^2+\xi_k(Q+\Pi)^2)} - \frac{4h_m(\Pi)^2(\omega^2 - \xi_k(Q)\xi_k(Q+\Pi))}{Z_F^2(\omega^2+\xi_k(Q)^2)(\omega^2+\xi_k(Q+\Pi)^2)} \Big) , \quad \text{(B34)}$$

where in the incommensurate case all fermionic kinetic terms $\xi_k(Q+\Pi)$ are shifted to $\xi_k(Q+\Pi+(0,\delta_{ic},0))$.

$$\partial_k Z_d = \tilde{\partial}_k \sum_Q \frac{1}{4\pi\omega_p^2} \Big(\frac{4h_d^2 F_{d,k}(Q)(\omega(\omega+\omega_p) + \xi_k(Q)^2)}{Z_F^2(\omega^2+\xi_k(Q)^2)((\omega+\omega_p)^2+\xi_k(Q)^2)} - \frac{4h_d^2 F_{d,k}(Q)}{Z_F^2(\omega^2+\xi_k(Q)^2)} \Big) \quad \text{(B35)}$$

Appendix C: REGULATOR

The fermionic propagator diverges at the Fermi surface

$$Z_F^2\xi(Q)\xi(Q+\Pi) - \Delta_a = 0 \ , \Delta_a = 2h_m(\Pi)^2\rho_a \ . \quad \text{(C1)}$$

Thus it is imperative to regularize the fermionic propagator at the Fermi surface. In the presence of a nonzero antiferromagnetic condensate the geometry of the Fermi surface changes, which requires the regulator to capture this deformation. Exemplary we demonstrate here how the regulator (25) removes the divergence at the Fermi surface. $\xi(\Pi), \xi(\Pi+Q)$ occur interchangeably as a product, thus at fixed Q the Fermi surface is regularized by

$$Z_F^2\xi_k(Q)\xi_k(Q+\Pi) - \Delta_a = 0 \ , \quad \text{(C2)}$$

where the kinetic terms $Z_F\xi_k(Q) = Z_F\xi(Q) + R_F(\xi(Q))$ and $Z_F\xi_k(Q+\Pi) = Z_F\xi(Q+\Pi) + R_F(\xi(Q+\Pi))$ are regularized by different terms of the regulator, which are without loss of generality

$$\begin{aligned} R_F(\xi(Q)) &= Z_F \,\mathrm{sign}(\xi(Q) - \frac{\Delta_a}{Z_F^2\xi_k(Q+\Pi)}) \\ &\quad \times (k - |\xi(Q) - \frac{\Delta_a}{Z_F^2\xi_k(Q+\Pi)}|)\Theta(k - |\xi(Q) - \frac{\Delta_a}{Z_F^2\xi_k(Q+\Pi)}|) \\ R_F(\xi(Q+\Pi)) &= Z_F \,\mathrm{sign}(\xi(Q+\Pi)) \\ &\quad \times (k - |\xi(Q+\Pi)|)\Theta(k - |\xi(Q+\Pi)|) \ . \end{aligned} \quad \text{(C3)}$$

We now examine what happens if the fermionic kinetic terms get too small and come too close to the Fermi surface. The first regulator in eq. (C3) is responsible for introducing a gap at the Fermi surface

$$\begin{aligned} &Z_F^2\xi_k(Q)\xi_k(Q+\Pi) - \Delta_a = \\ &Z_F^2\xi_k(Q+\Pi)\,\mathrm{sign}(\xi(Q) - \frac{\Delta_a}{Z_F^2\xi_k(Q+\Pi)})k \ . \end{aligned} \quad \text{(C4)}$$

The second regulator in eq. (C3) acts if the kinetic terms get too close to zero

$$Z_F^2 \operatorname{sign}(\xi(Q+\Pi))k \operatorname{sign}(\xi(Q) - \frac{\Delta_a}{Z_F^2 \xi_k(Q+\Pi)})k$$
$$= Z_F^2 \operatorname{sign}(Z_F^2 \xi(Q+\Pi)\xi(Q) - \Delta_a)k^2 \ . \qquad \text{(C5)}$$

We have used here that $\operatorname{sign}(\xi(Q+\Pi)) = \operatorname{sign}(\xi_k(Q+\Pi))$. An important property is that the Fermi surface is always approached with the correct sign.

[1] J. Hubbard, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **276**, 238 (1963).
[2] M. C. Gutzwiller, Physical Review Letters **10**, 159 (1963).
[3] J. Kanamori, Progress of Theoretical Physics **30**, 275 (1963).
[4] P. W. Anderson, Science **235**, 1196 (1987).
[5] K. Miyake, S. Schmitt-Rink, and C. M. Varma, Physical Review B **34**, 6554 (1986).
[6] D. J. Scalapino, E. Loh, and J. E. Hirsch, Physical Review B **34**, 8190 (1986).
[7] N. Bickers, D. Scalapino, and R. Scalettar, International Journal of Modern Physics B **01**, 687 (1987).
[8] P. A. Lee and N. Read, Physical Review Letters **58**, 2691 (1987).
[9] A. J. Millis, H. Monien, and D. Pines, Physical Review B **42**, 167 (1990).
[10] P. Monthoux, A. V. Balatsky, and D. Pines, Physical Review Letters **67**, 3448 (1991).
[11] D. Scalapino, Physics Reports **250**, 329 (1995).
[12] N. E. Bickers, D. J. Scalapino, and S. R. White, Physical Review Letters **62**, 961 (1989).
[13] N. Bulut, D. J. Scalapino, and S. R. White, Physical Review B **47**, 6157 (1993).
[14] T. Maier, M. Jarrell, T. Pruschke, and J. Keller, Physical Review Letters **85**, 1524 (2000).
[15] I. Grote, E. Körding, and F. Wegner, Journal of Low Temperature Physics **126**, 1385 (2002).
[16] V. Hankevych, I. Grote, and F. Wegner, Physical Review B **66** (2002), 10.1103/physrevb.66.094516.
[17] D. Sénéchal, P.-L. Lavertu, M.-A. Marois, and A.-M. S. Tremblay, Physical Review Letters **94** (2005), 10.1103/physrevlett.94.156404.
[18] T. A. Maier, M. S. Jarrell, and D. J. Scalapino, Physical Review Letters **96** (2006), 10.1103/physrevlett.96.047005.
[19] T. A. Maier, A. Macridin, M. Jarrell, and D. J. Scalapino, Physical Review B **76** (2007), 10.1103/physrevb.76.144516.
[20] D. J. Scalapino, Reviews of Modern Physics **84**, 1383 (2012).
[21] H. J. Schulz, Europhysics Letters (EPL) **4**, 609 (1987).
[22] I. E. Dzialoshinskii, Zhurnal Eksperimentalnoi i Teoreticheskoi Fiziki **93**, 1487 (1987).
[23] P. Lederer, G. Montambaux, and D. Poilblanc, Journal de Physique **48**, 1613 (1987).
[24] D. Zanchi and H. J. Schulz, Zeitschrift für Physik B Condensed Matter **103**, 339 (1996).
[25] D. Zanchi and H. J. Schulz, Europhysics Letters (EPL) **44**, 235 (1998).
[26] C. J. Halboth and W. Metzner, Physical Review Letters **85**, 5162 (2000).
[27] C. J. Halboth and W. Metzner, Physical Review B **61**, 7364 (2000).
[28] M. Salmhofer and C. Honerkamp, Progress of Theoretical Physics **105**, 1 (2001).
[29] C. Honerkamp, M. Salmhofer, N. Furukawa, and T. M. Rice, Physical Review B **63** (2001), 10.1103/physrevb.63.035109.
[30] C. Honerkamp and M. Salmhofer, Physical Review B **64** (2001), 10.1103/physrevb.64.184516.
[31] A. A. Katanin and A. P. Kampf, Physical Review B **72** (2005), 10.1103/physrevb.72.205128.
[32] W. Metzner, M. Salmhofer, C. Honerkamp, V. Meden, and K. Schönhammer, Reviews of Modern Physics **84**, 299 (2012).
[33] C. Husemann and M. Salmhofer, Phys. Rev. B **79**, 195125 (2009).
[34] C. Husemann, K.-U. Giering, and M. Salmhofer, Physical Review B **85** (2012), 10.1103/physrevb.85.075121.
[35] S. Uebelacker and C. Honerkamp, Physical Review B **86** (2012), 10.1103/physrevb.86.235140.
[36] A. Eberlein, Physical Review B **92** (2015), 10.1103/physrevb.92.235146.
[37] K.-U. Giering and M. Salmhofer, Physical Review B **86** (2012), 10.1103/physrevb.86.245122.
[38] J. Lichtenstein, D. S. de la Peña, D. Rohe, E. D. Napoli, C. Honerkamp, and S. Maier, Computer Physics Communications **213**, 100 (2017).
[39] K. Veschgini and M. Salmhofer, To appear soon (2017).
[40] W. Metzner, J. Reiss, and D. Rohe, physica status solidi (b) **243**, 46 (2006).
[41] J. Reiss, D. Rohe, and W. Metzner, Physical Review B **75** (2007), 10.1103/physrevb.75.075110.
[42] A. Eberlein and W. Metzner, Physical Review B **89** (2014), 10.1103/physrevb.89.035126.
[43] J. Wang, A. Eberlein, and W. Metzner, Physical Review B **89** (2014), 10.1103/physrevb.89.121116.
[44] H. Yamase, A. Eberlein, and W. Metzner, Physical Review Letters **116** (2016), 10.1103/physrevlett.116.096402.
[45] K. Veschgini, PhD Thesis, Heidelberg University Library (2017), 10.11588/heidok.00022476.
[46] T. Baier, E. Bick, and C. Wetterich, Physical Review B **62**, 15471 (2000).
[47] T. Baier, E. Bick, and C. Wetterich, Physical Review B **70** (2004), 10.1103/physrevb.70.125111.
[48] T. Baier, E. Bick, and C. Wetterich, Physics Letters B **605**, 144 (2005).
[49] H. Krahl and C. Wetterich, Physics Letters A **367**, 263 (2007).
[50] H. C. Krahl, J. A. Müller, and C. Wetterich, Physical Review B **79** (2009), 10.1103/physrevb.79.094526.

[51] H. C. Krahl, S. Friederich, and C. Wetterich, Physical Review B **80** (2009), 10.1103/physrevb.80.014436.
[52] S. Friederich, H. C. Krahl, and C. Wetterich, Physical Review B **81** (2010), 10.1103/physrevb.81.235108.
[53] S. Friederich, H. C. Krahl, and C. Wetterich, Physical Review B **83** (2011), 10.1103/physrevb.83.155125.
[54] J. Hubbard, Physical Review Letters **3**, 77 (1959).
[55] R. L. Stratonovich, Soviet Physics Doklady **2**, 416 (1957).
[56] H. Gies and C. Wetterich, Physical Review D **65** (2002), 10.1103/physrevd.65.065001.
[57] S. Floerchinger and C. Wetterich, Physics Letters B **680**, 371 (2009).
[58] C. Wetterich, Physics Letters B **301**, 90 (1993).
[59] D. F. Litim, Physical Review D **64** (2001), 10.1103/physrevd.64.105007.
[60] R. Coldea, S. M. Hayden, G. Aeppli, T. G. Perring, C. D. Frost, T. E. Mason, S.-W. Cheong, and Z. Fisk, Physical Review Letters **86**, 5377 (2001).
[61] N. M. R. Peres and M. A. N. Araújo, Physical Review B **65** (2002), 10.1103/physrevb.65.132404.
[62] S.-L. Yang, J. A. Sobota, Y. He, Y. Wang, D. Leuenberger, H. Soifer, M. Hashimoto, D. H. Lu, H. Eisaki, B. Moritz, T. P. Devereaux, P. S. Kirchmann, and Z.-X. Shen, Physical Review B **96** (2017), 10.1103/physrevb.96.245112.
[63] E. Pavarini, I. Dasgupta, T. Saha-Dasgupta, O. Jepsen, and O. K. Andersen, Physical Review Letters **87** (2001), 10.1103/physrevlett.87.047003.
[64] J.-Y. P. Delannoy, M. J. P. Gingras, P. C. W. Holdsworth, and A.-M. S. Tremblay, Physical Review B **79** (2009), 10.1103/physrevb.79.235130.
[65] C. Honerkamp and M. Salmhofer, Physical Review Letters **87** (2001), 10.1103/physrevlett.87.187004.
[66] H. J. Schulz, Physical Review Letters **64**, 1445 (1990).
[67] J. M. Pawlowski, M. M. Scherer, R. Schmidt, and S. J. Wetzel, Annals of Physics **384**, 165 (2017).
[68] J. Longo and P. Raccah, Journal of Solid State Chemistry **6**, 526 (1973).

7 Discussion

The different topics in this thesis unite under the purpose to calculate phase diagrams of various physical systems. We have resolved problems in existing methods, invented new methods and applied all these methods to examine phase diagrams.

7.1 Functional Renormalization Group Flows

The first part of this thesis, *Physics and the Choice of Regulators in Functional Renormalisation Group Flows*, addresses the reliability of the results of functional renormalization group calculations [23] and their robustness with respect to changes in the regularization scheme. Even though the functional renormalization group is formulated in an exact functional differential equation, it can only be solved by truncating it to a finite number of coupled differential equations. The results of the full functional differential equation are independent of the explicit choice of the regulator. However, the results of a truncated system of differential equation depend on the regulator choice. In this case reliable physical results can only be obtained by a suitable regulator. An obvious question arises: is there an optimal regulator which yields the results that agree best with the underlying physical system? Until now, it has been shown that a functional *a priori* criterion can be employed to deduce the best regulator if the physical system is described by a model in local potential approximation [103]. However, many physical systems cannot be formulated in this approximation.

In order to address the search for an optimized regulator practically, we construct a new *a priori* criterion to identify the best regulator from a given set of regulators. It is based on comparing the length of the renormalization flow trajectories in theory space. When applying our criterion to models in local potential approximation, our optimal regulator coincides with the regulator obtained in [103]. Our criterion is however more practicable to implement, because it circumvents the difficult functional optimization step. Thus our optimization procedure might be suitable for finding the optimal regulator for more complicated systems, where a direct functional optimization is not feasible.

In our work we discuss how the results of previous functional renormalization group calculations are influenced by different choices of regulators. In a two-boson system we show that a change in the regulator can influence the critical behavior so far that it induces a change of the universality class. Moreover, in this model we measure explicitly the robustness of the critical exponents in various truncations with respect to a change in the regulator. We confirm that a better truncation is much more robust with respect to unsuitable regulator choices. At the example of the Fermi polaron problem we explain the difficulty of formulating a regularization scheme which treats fermions and bosons on an equal basis.

In the Hubbard model in the antiferromagnetic regime the Fermi surface is deformed by the antiferromagnetic condensate. At that point conventional regulators cannot renormalize around a changing Fermi surface geometry. For this reason, this work was crucial in determining a suitable regularization scheme for the flow equations of the Hubbard model.

7.2 High-Temperature Superconductivity

High-temperature superconductivity was already found in cuprate ceramics in 1986 [55]. Now, more than 30 years of efforts of many condensed matter physicists later, there is still no comprehensive understanding of the phase diagram of high-temperature superconductors. The phase diagram of high-temperature superconductors displays many unconventional phases: at low temperature and close to half-filling, cuprates are in an insulating antiferromagnetic state. Depending on the level of doping, which is defined by the chemical potential, and the hopping parameters, this antiferromagnetic state can either be commensurate or incommensurate. Higher levels of doping lead to d-wave superconductivity. Even higher levels of doping induce a ferromagnetic phase. At higher temperatures there is an elusive precursor phase called the pseudogap phase.

The most promising model to describe the physics of cuprate superconductors is the Hubbard model on the square lattice [60–62]. It describes electrons moving in two-dimensional copper oxide planes. Most of the phases of the Hubbard model cannot be theoretically accessed by many traditional methods, thus the mechanisms of their emergence are still not fully understood. It is difficult to apply perturbation theory, because there is no suitable small expansion parameter. Hence, by perturbation theory, the phase diagram can only be accessed from the limits of very small or very strong interaction strengths [104, 105]. Monte-Carlo simulations can successfully describe the Hubbard model at half-filling. However, these simulations

fail in regions with electron or hole doping, since away from half-filling they suffer from the fermion sign problem [106].

We address the problem of examining the phase diagram of the $t - t'$ Hubbard model on the square lattice with the functional renormalization group [23]. Here, the flow equation interpolates between the microscopic Hubbard action in Matsubara formalism and the full quantum effective action. The first applications of this method were formulated in a purely fermionic language [63–78]. They were successful in identifying the dominant channels leading to a diverging four-Fermi interaction at some symmetry breaking scale or temperature. At this divergence, the flow equations signal the onset of magnetic or d-wave superconducting symmetry broken regimes. In the purely fermionic formalism it is very challenging to circumvent the divergence to continue the flow equations into the symmetry broken regimes. In order to calculate the full quantum effective action it is important to follow the renormalization flow into the symmetry broken phases. So far researchers have treated this problem by switching to other methods like mean field theory [79–84]. One concept to address this problem within one consistent renormalization calculation is the bosonization of the diverging momentum channels of the four-Fermi interaction.

In the article *Exploring the Hubbard Model on the Square Lattice at Zero Temperature with a Bosonized Functional Renormalization Approach* we build upon previous work [85–92] and translate the diverging momentum channels into bosons via a scale dependent Hubbard-Stratonovich transformation, called flowing bosonization. In contrast to traditional Hubbard-Stratonovich transformations, where the fermionic action is formulated in a bosonic model right from the start, flowing bosonization only generates the bosonic channels if their corresponding fermionic momentum channels would enhance the four-Fermi interaction. Any divergence of the four-Fermi interaction is transmitted via Yukawa interactions into bosonic sectors. There, a vanishing bosonic mass term then signals the instability in the respective momentum channel. Flowing bosonization makes it possible to continue the flow equations into the symmetry broken regimes and allow for a calculation of the full quantum effective action. Furthermore, the nature of the bosons yields a direct interpretation of the underlying phase. As long as all possible divergencies of the four-Fermi interaction are covered by the flowing bosonization, this procedure gives rise to an unbiased treatment of all possible phases on an equal basis.

Previous work has established and improved this scheme and explained many important parts of the Hubbard model phase diagram at finite temperature. Still using the standard Hubbard-Stratanovich transformation, an early examination revealed

antiferromagnetism at half filling and close to it [86, 87]. The introduction of flowing bosonization allowed for the examination of d-wave superconductivity at higher doping [89]. And finally, the inclusion of charge density and s-wave bosons allowed for an accurate calculation of the phase boundaries [92].

In this thesis we examine the zero temperature phase diagram of the Hubbard model on the square lattice dependent on the internal properties of the underlying cuprates. These are the next-to-nearest neighbor hopping t' and the chemical potential μ, which indicates the level of doping. The nearest neighbor hopping t sets the energy scale and the Coulomb interaction is fixed at $U = 3$ [1].

At zero temperature we face extra challenges not present in finite temperature calculations. The Matsubara frequencies are discrete at nonzero temperature, however, in the limit of zero temperature, they take on continuous values. Thus at finite temperature, at small momentum scales k, the flow equations are dominated by only one Matsubara frequency in the denominator of the bosonic propagator $\omega = 0$ and two Matsubara frequencies in the fermionic propagator $\omega = \pm\pi T$. At zero temperature one needs to integrate over all Matsubara frequencies. Because of this, the fermionic propagator is no longer gapped by the lowest Matsubara frequencies $\omega = \pm\pi T$. In other words, while at finite temperature the macroscopic bosonic degrees of freedom decouple from the fermions, at zero temperature, the fermionic fluctuations contribute even on macroscopic scales. The absent Matsubara gap at zero temperature causes the fermionic propagator to diverge at the Fermi surface. Thus special care needs to be taken to renormalize around the correct Fermi surface, which has thereby also an increased effect on the flow equations compared to finite temperature calculations. Another simplifying observation at finite temperature is the dimensional reduction in the bosonic sector at macroscopic scales. Since in this scenario only the Matsubara zero mode contributes to the flow equations, the theory can be influenced only by spatial fluctuations.

In this work we carefully treat these problems in order to be able to examine the phase diagram of the Hubbard model at zero temperature. For this purpose we formulate a consistent system of flow equations valid in all symmetric and symmetry broken regimes. These flow equations are regularized within a scheme that considers the changing Fermi surface geometry and topology arising from an antiferromagnetic condensate. By solving the flow equations in the symmetric regime, we examine the leading instabilities of the four-Fermi interaction. We identify three different channels in which symmetry breaking occurs. These correspond to commensurate antiferromagnetism, incommensurate antiferromagnetism and d-wave superconduc-

[1] For a definition of the Hubbard model refer to eq. (2.2)

tivity. Compared to the purely fermionic calculations, our framework can much more easily distinguish between commensurate and incommensurate antiferromagnetism. This is because the minimum of the magnetic inverse propagator is directly related to the type of antiferromagnetism. This property allows for an accurate identification of the first order transition between commensurate and incommensurate order. We are able to pin down the location of this transition and find that it coincides with the line of Van Hove singularities.

By continuing the flow equations into the regimes of symmetry breaking, we are able to examine the interplay of the magnetic and d-wave order parameters. At intermediate scales there is a regime of coexisting antiferromagnetism and d-wave superconductivity. However, these two phases have a tendency to repel each other. This property leads to the eradication of the coexistence phase at macroscopic scales.

In conclusion, we have studied the emergence of three different phases in the Hubbard model at moderate doping and electron hopping: commensurate antiferromagnetism, incommensurate antiferromagnetism and d-wave superconductivity. Furthermore, we have identified the transitions between these phases and examined the interplay between different orders while including fluctuations even deep in the symmetry broken regimes. Our work leaves open the fate of d-wave superconductivity at scales of physical samples. Although very demanding, this question can in principle be answered by continuing the flow of the Yukawa interactions into the symmetry broken regimes. The major technical achievement of this thesis is the proper treatment of the Fermi surface, which changes its geometry and topology in the symmetry broken phases. The most important physical results are the emergence and the vanishing of the coexistence phase and the nature and location of the first order transition between commensurate and incommensurate antiferromagnetism.

7.3 Artificial Neural Networks and Phase Transitions

While many methods to examine the physics of phase diagrams have been applied to their limits, there are still open questions. For example the pseudogap phase in the phase diagram of high-temperature superconductors or the critical end point in quantum chromodynamics are still not fully understood. Also, many new types of materials and phases of matter exhibit features in phase diagrams where conventional methods have not been successful. We expect to solve some of these problems with radically different methods. In this thesis we develop machine learning algorithms as methods for examining phase diagrams. They are not bound by limits of human intuition or intelligence and could thus lead to some new and unexpected

results.

Since 2016 machine learning methods have been transitioning from a data science and image recognition to the physics of phase diagrams. Physicists were able to simulate physical systems on restricted Boltzmann machines [35]. Their method is successful in finding the ground state and describing the unitary time evolution of one- and two-dimensional spin models. Standard feed-forward neural networks were employed to calculate accurate phase boundaries [46]. They were trained on Monte-Carlo configurations, without giving the algorithm any knowledge of the Hamiltonian or of the occurring orders. This is achieved by training the neural network in a supervised manner in different regimes where the nature of the phases is known. In many physical systems, the algorithm can then successfully predict the phases on configurations where the phases are *a priori* unknown. This algorithm has demonstrated the best overall performance among machine learning algorithms on difficult physical systems. Until now it has been employed to find phase transitions in systems with topological phases [47], in strongly correlated electron systems [48] and in lattice gauge theory [3][2]. Furthermore, they are a good candidate to approximately circumvent the sign problem [107]. Phases can also be found by linear unsupervised learning algorithms, i.e. principal component analysis [49], without even knowing of the existence of different phases. Because of its linear nature, principal component analysis has trouble finding the phases in more complicated models. Therefore, the need for a nonlinear unsupervised machine learning algorithm arises. It was suggested to employ supervised learning methods to do unsupervised learning [50] in such cases. In *Unsupervised Learning of Phase Transitions: From Principal Component Analysis to Variational Autoencoders* we propose the first direct nonlinear unsupervised learning algorithm, the neural network based autoencoder, to detect phase transitions.

At the time of developing the autoencoder approach, very little was known about unsupervised machine learning of phase transitions. Hence, we had the freedom to examine the results of many different algorithms without prejudice. In our article we compare principal component analysis, kernel principal component analysis, clustering methods, manifold methods, autoencoders and variational autoencoders. All of these algorithms were exemplary trained on Monte-Carlo samples of the ferromagnetic Ising model in two dimensions, the simplest model to exhibit a phase transition. They did not receive any further information of the Hamiltonian or of the underlying orders. It turns out that most of these algorithms are successful in

[2]For more details see *Machine Learning of Explicit Order Parameters: From the Ising Model to SU(2) Lattice Gauge Theory*.

discovering some structures on the Monte-Carlo samples. However, most of the results of the machine learning algorithms can only be recognized if one already knows them. Hence, we focus on the algorithms which can potentially give insights which are not known before. The only algorithms satisfying this request are (kernel) principal component analysis and (variational) autoencoders. It turns out, that these algorithms learn the order parameter, the magnetization, in a latent representation, in order to discriminate between different samples. The latent representation takes on different values in different phases of the Ising model. Hence, we propose to employ the average absolute value of the latent representation as an indicator to estimate a phase transition between different regimes. Starting from the insight that the ordered phase contains more structure than the disordered phase, we introduce the average reconstruction loss function as a universal identifier for phase transitions. We further demonstrate the ability of (kernel) principal component analysis and (variational) autoencoders to successfully find the correct phase transition in the antiferromagnetic Ising model in two dimensions and in the three-dimensional XY model. Until now, the autoencoder framework has been applied to different systems including the Ising model [51], the BKT phase in the two dimensional XY model [108] and spin glasses [109]. Furthermore, it inspired progress in molecular simulations [110]. In order to find the BKT phase the variational autoencoder was improved by physical constraints like energy conservation. This enabled the autoencoder to be employed to sample meaningful physical configurations of the unordered phase and the BKT phase. The spin glass transitons were found by relating the cross entropy reconstruction loss of autoencoders to the entropy of the physical system [109]. For extracting information out of molecular simulations, the autoencoder was improved by calculating a weighted average reconstruction loss. This new autoencoder was trained on a supercomputer to find an optimal and physically interpretable reaction coordinate and an optimum probability distribution which can be used for sampling. All these results were obtained within one year from our proposal, which highlights the potential of autoencoders as a tool to discover physical phases.

Over the last year it turned out that the most powerful algorithm for exploring phase diagrams of complex physical systems is the artificial feed-forward neural network in the supervised context. These artificial neural networks suffer from a huge drawback. There is still no comprehensive theory of what they actually learn while training in order to make their predictions. This unpredictability is a serious issue in many disciplines where neural networks are employed. In physics, as a quantitative research area, it is imperative to explain research results on physical ground to gain reliable insights. So far these machine learning algorithms have only

been applied to uncover phase diagrams of systems which are already known, in these cases one can judge the reliability of the results by comparing them to existing calculations. However, when applying artificial neural networks to new systems one has no reason to trust the results.

One way to deal with this problem is to resort to other machine learning algorithms, which can be more easily interpreted. One such machine learning algorithm is the support vector machine. It is less powerful in learning hierarchical structures, however the decision boundary can be read off directly from the learned coefficients. A first application to the two-dimensional Ising model and the two-dimensional Ising gauge theory demonstrates the ability of support vector machines to correctly predict phases in these models. In the Ising model this algorithm distinguishes between phases via the magnetization, the underlying order parameter [111].

In *Machine Learning of Explicit Order Parameters: From the Ising Model to SU(2) Lattice Gauge Theory* we successfully address the problem of interpreting artificial neural networks classifying physical phases. We develop a new way to interpret neural networks. It is based on a discrete interpolation between a general neural network and a minimal neural network which still performs equally well on the training data. At the heart of this interpretation method is the parameter-dependent *correlation probing neural network*. It consists of several hierarchically stacked neural network building blocks. Each of these building blocks has one or several parameters which can reduce the expressibility of the neural network. By training a sequence of neural networks to predict phases in a physical system and by stepwise lowering the expressibility of the whole neural network, one eventually finds the optimal minimal neural network. This minimal neural network still performs equally well as the original full neural network. However, reducing the expressibility below the minimal neural network causes a drastic drop in the classification performance. In general there can be more than one optimal minimal neural networks, corresponding to different independent decision functions. If the decision function of these minimal neural networks is sufficiently symmetric, it can be expressed as a simple function of the input data. In this form one can easily perform a regression analysis on very few parameters compared to the millions of parameters of the original neural network. If artificial neural networks are tasked with classifying phases of physical models, it turns out that the decision functions of these minimal neural networks all correspond to physical quantities. In the case of the two-dimensional Ising model this is the magnetization, as also learned by the support vector machine. However, another quantity which is learned by neural networks is the expected energy per spin. In conclusion, we have explained the results of neural networks classifying phases of

the Ising model by physical quantities. Thus, we have justified the results, even if the phase transition of the Ising model would have been unknown.

These findings bear another significance. Our method of interpreting neural networks is constructive and it does not rely on the knowledge of any order parameter beforehand. In other words, we have found the order parameter of the Ising model, just by interpreting a neural network, without any physical deduction. This provides the last component needed for an automatic determination of phase diagram: (i) We use unsupervised learning to infer the existence of different phases on the basis of Monte-Carlo samples. (ii) We employ supervised feed-forward neural networks to find the accurate phase boundary. (iii) By interpreting what the neural network has learned, we identify the underlying order parameter, which determines the nature of the ordered phase.

This three-step determination of phase diagrams can be applied to more complicated models than spin models. In our article we automatically detect and interpret the confinement-deconfinement transition in four-dimensional Euclidean SU(2) lattice gauge theory. Unsupervised learning successfully detects the existence of two different phases. Supervised learning estimates the phase boundary close to the results from direct lattice calculations. Finally, by interpreting the decision function of the supervised learning neural network, we find that it distinguishes between phases by the Polyakov loop. This quantity is the order parameter corresponding to the confinement-deconfinement phase transiton in lattice gauge theory.

To summarize to impact on physical phase diagrams: using machine learning, we performed an automatic evaluation of the phase diagrams of the Ising model and SU(2) lattice gauge theory, including the determination of the order parameter, which is in agreement with previous calculations. Since this method of determining the nature of the order is radically different from other existing methods, it might help uncover the nature of phases where conventional methods have failed.

From a different perspective, we have developed a successful method to interpret artificial neural networks. Perhaps this work can contribute to the understanding of artificial neural networks in general, such that, some time in the near future, they will not be considered as black box algorithms any more.

Based on the developments in this thesis we plan to apply artificial neural networks to the Hubbard Model in the near future.

8 Summary

In this thesis we have improved existing techniques and developed new methods to explore phase diagrams. These methods were applied to several systems including the Ising model, the Hubbard model and SU(2) lattice gauge theory.

In *Physics and the Choice of Regulators in Functional Renormalisation Group Flows* we examined how much one can trust results obtained by functional renormalization group calculations. Further, we devised a criterion for finding an optimal regulator in renormalization group flows.

We analyzed the zero temperature phase diagram of the Hubbard model on the square lattice in *Exploring the Hubbard Model on the Square Lattice at Zero Temperature with a Bosonized Functional Renormalization Approach*. Our calculation was able to find leading instabilities which correspond to the phases of d-wave superconductivity and commensurate and incommensurate antiferromagnetism. We examined the transition between each of these phases and found a coexistence of superconductivity and antiferromagnetism, but only on intermediate energy scales. We pinned down the location of the commensurate-incommensurate transition, which coincides with the line of Van Hove singularities and found that it is of first order. The major technical achievement is the continuation of the renormalization group flow into regimes of antiferromagnetism where the geometry of the Fermi surface changes.

The article *Unsupervised Learning of Phase Transitions: From Principal Component Analysis to Variational Autoencoders* contains the proposal of a machine learning algorithm to find phases in physical systems in an unsupervised manner. While we introduced it at simple spin models, this algorithm has so far been employed to explore systems like spin glasses, the two-dimensional XY model and molecular configurations.

In *Machine Learning of Explicit Order Parameters: From the Ising Model to SU(2) Lattice Gauge Theory* we developed a method to interpret artificial neural networks. The procedure is tailored to neural networks classifying phases of matter, where many physical quantities can be formulated in equations obeying many symmetries. We found that neural networks base their decisions on order parameters and other

thermodynamic quantities. We further determined the phase diagram of SU(2) lattice gauge theory, including the existence of different phases, the nature of the ordered phase and the phase boundary. The benefit of this article is twofold: on the one hand it justifies the application of neural networks to classify phases of matter on a physical basis. On the other hand our procedure can be used to determine the order parameter of the underlying system.

9 Contributions to the Articles

1. **Physics and the Choice of Regulators in Functional Renormalisation Group Flows**

   Sebastian J. Wetzel was the corresponding author of this article, however in this discipline it is conventional to order the authors alphabetically. Sebastian J. Wetzel contributed to the theoretical basics and performed the examination of the regulator loops and the multi-field models. Jan M. Pawlowski had the idea for this work and devised the theoretical framework and the optimization criterion. Michael Scherer contributed substantially to the introduction and the theoretical basics and was a direct supervisor to Sebastian J. Wetzel during all stages of the work. Richard Schmidt was responsible for the content of the Fermi-Polaron chapter.

2. **Unsupervised Learning of Phase Transitions: From Principal Component Analysis to Variational Autoencoders**

   Sebastian J. Wetzel concepted and conducted all of the work for this article.

3. **Machine Learning of Explicit Order Parameters: From the Ising Model to SU(2) Lattice Gauge Theory**

   Sebastian J. Wetzel concepted and conducted all of the work for this article, except the Monte-Carlo simulations which were used as input in the SU(2) case. Those were produced by Manuel Scherzer.

4. **Exploring the Hubbard Model on the Square Lattice at Zero Temperature with a Bosonized Functional Renormalization Approach**

   Sebastian J. Wetzel concepted and conducted all of the work for this article. Christof Wetterich helped with advice.

10 Bibliography

[1] J. M. Pawlowski, M. M. Scherer, R. Schmidt, and S. J. Wetzel. Physics and the choice of regulators in functional renormalisation group flows. *Annals of Physics*, 384:165–197, September 2017. doi:10.1016/j.aop.2017.06.017.

[2] S. J. Wetzel. Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders. *Physical Review E*, 96(2), August 2017. doi:10.1103/physreve.96.022140.

[3] S. J. Wetzel and M. Scherzer. Machine learning of explicit order parameters: From the ising model to SU(2) lattice gauge theory. *Physical Review B*, 96 (18), November 2017. doi:10.1103/physrevb.96.184410.

[4] S. J. Wetzel. Exploring the hubbard model on the square lattice at zero temperature with a bosonized functional renormalization approach. *ArXiv e-prints*, December 2017. URL `http://adsabs.harvard.edu/abs/2017arXiv171204297W`.

[5] G. Jaeger. The ehrenfest classification of phase transitions: Introduction and evolution. *Archive for History of Exact Sciences*, 53(1):51–81, May 1998. doi:10.1007/s004070050021.

[6] L. Landau. On the theory of phase transitions. *Zh. Eksp. Teor.*, 7:19–32, 1969.

[7] E. Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, 31 (1):253–258, February 1925. doi:10.1007/bf02980577.

[8] L. Onsager. Crystal statistics. i. a two-dimensional model with an order-disorder transition. *Physical Review*, 65(3-4):117–149, February 1944. doi:10.1103/physrev.65.117.

[9] V. L. Berezinskiĭ. Destruction of long-range order in one-dimensional and two-dimensional systems having a continuous symmetry group i. classical systems. *Soviet Journal of Experimental and Theoretical Physics*, 32:493, 1971. URL `http://adsabs.harvard.edu/abs/1971JETP...32..493B`.

[10] J. M. Kosterlitz and D. J. Thouless. Ordering, metastability and phase transitions in two-dimensional systems. *Journal of Physics C Solid State Physics*, 6:1181–1203, April 1973. doi:10.1088/0022-3719/6/7/010. URL `http://adsabs.harvard.edu/abs/1973JPhC....6.1181K`.

[11] N. D. Mermin and H. Wagner. Absence of ferromagnetism or antiferromagnetism in one- or two-dimensional isotropic heisenberg models. *Physical Review Letters*, 17(22):1133–1136, November 1966. doi:10.1103/physrevlett.17.1133.

[12] P. C. Hohenberg. Existence of long-range order in one and two dimensions. *Physical Review*, 158(2):383–386, June 1967. doi:10.1103/physrev.158.383.

[13] D. C. Tsui, H. L. Stormer, and A. C. Gossard. Two-dimensional magnetotransport in the extreme quantum limit. *Physical Review Letters*, 48(22): 1559–1562, May 1982. doi:10.1103/physrevlett.48.1559.

[14] R. B. Laughlin. Anomalous quantum hall effect: An incompressible quantum fluid with fractionally charged excitations. *Physical Review Letters*, 50(18): 1395–1398, May 1983. doi:10.1103/physrevlett.50.1395.

[15] A. Petermann and E. C. G. Stueckelberg. *La normalisation des constantes dans la théorie des quanta*. Birkhäuser, 1953. doi:10.5169/seals-112426.

[16] M. Gell-Mann and F. E. Low. Quantum electrodynamics at small distances. *Physical Review*, 95(5):1300–1312, September 1954. doi:10.1103/physrev.95.1300.

[17] C. G. Callan. Broken scale invariance in scalar field theory. *Physical Review D*, 2(8):1541–1547, October 1970. doi:10.1103/physrevd.2.1541.

[18] L. P. Kadanoff. Scaling laws for Ising models near T(c). *Physics*, 2:263–272, 1966.

[19] K. G. Wilson. Renormalization group and critical phenomena. i. renormalization group and the kadanoff scaling picture. *Physical Review B*, 4(9):3174–3183, November 1971. doi:10.1103/physrevb.4.3174.

[20] K. G. Wilson. Renormalization group and critical phenomena. II. phase-space cell analysis of critical behavior. *Physical Review B*, 4(9):3184–3205, November 1971. doi:10.1103/physrevb.4.3184.

[21] F. J. Wegner and A. Houghton. Renormalization group equation for critical phenomena. *Physical Review A*, 8(1):401–412, July 1973. doi:10.1103/physreva.8.401.

[22] J. Polchinski. Renormalization and effective lagrangians. *Nuclear Physics B*, 231(2):269–295, January 1984. doi:10.1016/0550-3213(84)90287-6.

[23] C. Wetterich. Exact evolution equation for the effective potential. *Physics Letters B*, 301(1):90–94, February 1993. doi:10.1016/0370-2693(93)90726-x.

[24] T. R. Morris. The exact renormalisation group and approximate solutions. *International Journal of Modern Physics A*, 09(14):2411–2449, June 1994. doi:10.1142/s0217751x94000972.

[25] W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4):115–133, December 1943. doi:10.1007/bf02478259.

[26] D. O. Hebb. Organization of behavior. *Journal of Clinical Psychology*, 6(3):307–307, July 1950. doi:10.1002/1097-4679(195007)6:3<307::aid-jclp2270060338>3.0.co;2-k.

[27] F. Rosenblatt. *The Perceptron: A Perceiving and Recognizing Automaton*. Cornell Aeronautical Laboratory, New York, January 1957.

[28] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536, October 1986. doi:10.1038/323533a0.

[29] P. J. Werbos. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. PhD thesis, Harvard University, 1974. Department of Applied Mathematics.

[30] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[31] G. Hinton et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6):82–97, November 2012. doi:10.1109/msp.2012.2205597.

[32] D. Silver et al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, January 2016. doi:10.1038/nature16961.

[33] G. Kasieczka, T. Plehn, M. Russell, and T. Schell. Deep-learning top taggers or the end of QCD? *Journal of High Energy Physics*, 2017(5), may 2017. doi:10.1007/jhep05(2017)006.

[34] B. E. Husic and V. S. Pande. Unsupervised learning of dynamical and molecular similarity using variance minimization. *ArXiv e-prints*, December 2017. URL `http://adsabs.harvard.edu/abs/2017arXiv171207704H`.

[35] G. Carleo and M. Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, February 2017. doi:10.1126/science.aag2302.

[36] G. Torlai and R. G. Melko. Learning thermodynamics with boltzmann machines. *Physical Review B*, 94(16), October 2016. doi:10.1103/physrevb.94.165134.

[37] D.-L. Deng, X. Li, and S. D. Sarma. Machine learning topological states. *Physical Review B*, 96(19), November 2017. doi:10.1103/physrevb.96.195145.

[38] D.-L. Deng, X. Li, and S. D. Sarma. Quantum entanglement in neural network states. *Physical Review X*, 7(2), May 2017. doi:10.1103/physrevx.7.021021.

[39] X. Gao and L.-M. Duan. Efficient representation of quantum many-body states with deep neural networks. *Nature Communications*, 8(1), September 2017. doi:10.1038/s41467-017-00705-2.

[40] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo. Many-body quantum state tomography with neural networks. *ArXiv e-prints*, March 2017. URL `http://adsabs.harvard.edu/abs/2017arXiv170305334T`.

[41] K.-I. Aoki and T. Kobayashi. Restricted boltzmann machines for the long range ising models. *Modern Physics Letters B*, 30(34):1650401, December 2016. doi:10.1142/s0217984916504017.

[42] L. Huang and L. Wang. Accelerated monte carlo simulations with restricted boltzmann machines. *Physical Review B*, 95(3), January 2017. doi:10.1103/physrevb.95.035105.

[43] J. Liu, Y. Qi, Z. Y. Meng, and L. Fu. Self-learning monte carlo method. *Physical Review B*, 95(4), January 2017. doi:10.1103/physrevb.95.041101.

[44] N. Portman and I. Tamblyn. Sampling algorithms for validation of supervised learning models for ising-like systems. *Journal of Computational Physics*, 350: 871–890, December 2017. doi:10.1016/j.jcp.2017.06.045.

[45] L. Li, T. E. Baker, S. R. White, and K. Burke. Pure density functional for strong correlation and the thermodynamic limit from machine learning. *Physical Review B*, 94(24), December 2016. doi:10.1103/physrevb.94.245129.

[46] J. Carrasquilla and R. G. Melko. Machine learning phases of matter. *Nature Physics*, February 2017. doi:10.1038/nphys4035.

[47] Y. Zhang and E.-A. Kim. Quantum loop topography for machine learning. *Physical Review Letters*, 118(21), May 2017. doi:10.1103/physrevlett.118.216401.

[48] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami. Machine learning phases of strongly correlated fermions. *Physical Review X*, 7(3), August 2017. doi:10.1103/physrevx.7.031038.

[49] L. Wang. Discovering phase transitions with unsupervised learning. *Physical Review B*, 94(19), November 2016. doi:10.1103/PhysRevB.94.195105.

[50] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber. Learning phase transitions by confusion. *Nature Physics*, February 2017. doi:10.1038/nphys4037.

[51] W. Hu, R. R. P. Singh, and R. T. Scalettar. Discovering phases, phase transitions, and crossovers through unsupervised machine learning: A critical examination. *Physical Review E*, 95(6), June 2017. doi:10.1103/physreve.95.062122.

[52] H. Motzkau. Schematic doping phase diagram of cuprate hight-temperature superconductors, August 2013. URL `https://commons.wikimedia.org/wiki/File:Cuprates_phasedigagram_en.svg`. wikimedia.org.

[53] H. Kamerlingh Onnes. The resistance of pure mercury at helium temperatures. *Comm. Phys. Lab. Univ. Leiden*, 122 and 124, 1911.

[54] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Theory of superconductivity. *Physical Review*, 108(5):1175–1204, December 1957. doi:10.1103/physrev.108.1175.

[55] J. G. Bednorz and K. A. Mueller. Possible highT c superconductivity in the balacuo system. *Zeitschrift für Physik B Condensed Matter*, 64(2):189–193, June 1986. doi:10.1007/bf01303701.

[56] Y. Kamihara, H. Hiramatsu, M. Hirano, R. Kawamura, H. Yanagi, T. Kamiya, and H. Hosono. Iron-based layered superconductor: LaOFeP. *Journal of the American Chemical Society*, 128(31):10012–10013, August 2006. doi:10.1021/ja063355c.

[57] Y. Kamihara, T. Watanabe, M. Hirano, and H. Hosono. Iron-based layered superconductor la[o1-xFx]FeAs (x= 0.05-0.12) withTc= 26 k. *Journal of the American Chemical Society*, 130(11):3296–3297, March 2008. doi:10.1021/ja800073m.

[58] A. P. Drozdov, M. I. Eremets, I. A. Troyan, V. Ksenofontov, and S. I. Shylin. Conventional superconductivity at 203 kelvin at high pressures in the sulfur hydride system. *Nature*, 525(7567):73–76, August 2015. doi:10.1038/nature14964.

[59] Y. Ge, F. Zhang, and Y. Yao. First-principles demonstration of superconductivity at 280 k in hydrogen sulfide with low phosphorus substitution. *Physical Review B*, 93(22), June 2016. doi:10.1103/physrevb.93.224513.

[60] J. Hubbard. Electron correlations in narrow energy bands. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 276(1365): 238–257, November 1963. doi:10.1098/rspa.1963.0204.

[61] M. C. Gutzwiller. Effect of correlation on the ferromagnetism of transition metals. *Physical Review Letters*, 10(5):159–162, March 1963. doi:10.1103/physrevlett.10.159.

[62] J. Kanamori. Electron correlation and ferromagnetism of transition metals. *Progress of Theoretical Physics*, 30(3):275–289, September 1963. doi:10.1143/ptp.30.275.

[63] D. Zanchi and H. J. Schulz. Instabilities of weakly correlated electronic gas on a two dimensional lattice. *Zeitschrift für Physik B Condensed Matter*, 103 (2):339–342, June 1996. doi:10.1007/s002570050385.

[64] D. Zanchi and H. J. Schulz. Weakly correlated electrons on a square lattice: A renormalization group theory. *Europhysics Letters (EPL)*, 44(2):235–241, October 1998. doi:10.1209/epl/i1998-00462-x.

[65] C. J. Halboth and W. Metzner. d-wave superconductivity and pomeranchuk instability in the two-dimensional hubbard model. *Physical Review Letters*, 85 (24):5162–5165, December 2000. doi:10.1103/physrevlett.85.5162.

[66] C. J. Halboth and W. Metzner. Renormalization-group analysis of the two-dimensional hubbard model. *Physical Review B*, 61(11):7364–7377, March 2000. doi:10.1103/physrevb.61.7364.

[67] M. Salmhofer and C. Honerkamp. Fermionic renormalization group flows: Technique and theory. *Progress of Theoretical Physics*, 105(1):1–35, January 2001. doi:10.1143/ptp.105.1.

[68] C. Honerkamp, M. Salmhofer, N. Furukawa, and T. M. Rice. Breakdown of the landau-fermi liquid in two dimensions due to umklapp scattering. *Physical Review B*, 63(3), January 2001. doi:10.1103/physrevb.63.035109.

[69] C. Honerkamp and M. Salmhofer. Temperature-flow renormalization group and the competition between superconductivity and ferromagnetism. *Physical Review B*, 64(18), October 2001. doi:10.1103/physrevb.64.184516.

[70] A. A. Katanin and A. P. Kampf. Order-parameter symmetries for magnetic and superconducting instabilities: Bethe-salpeter analysis of functional renormalization-group solutions. *Physical Review B*, 72(20), November 2005. doi:10.1103/physrevb.72.205128.

[71] W. Metzner, M. Salmhofer, C. Honerkamp, V. Meden, and K. Schönhammer. Functional renormalization group approach to correlated fermion systems. *Reviews of Modern Physics*, 84(1):299–352, March 2012. doi:10.1103/revmodphys.84.299.

[72] C. Husemann and M. Salmhofer. Efficient parametrization of the vertex function, Ω scheme, and the t, t' hubbard model at van hove filling. *Phys. Rev. B*, 79:195125, May 2009. doi:10.1103/PhysRevB.79.195125. URL `https://link.aps.org/doi/10.1103/PhysRevB.79.195125`.

[73] C. Husemann, K.-U. Giering, and M. Salmhofer. Frequency-dependent vertex functions of the (t,t′) hubbard model at weak coupling. *Physical Review B*, 85(7), February 2012. doi:10.1103/physrevb.85.075121.

[74] S. Uebelacker and C. Honerkamp. Self-energy feedback and frequency-dependent interactions in the functional renormalization group flow for the two-dimensional hubbard model. *Physical Review B*, 86(23), December 2012. doi:10.1103/physrevb.86.235140.

[75] A. Eberlein. Self-energy effects in functional renormalization group flows of the two-dimensionalt-t′hubbard model away from van hove filling. *Physical Review B*, 92(23), December 2015. doi:10.1103/physrevb.92.235146.

[76] K.-U. Giering and M. Salmhofer. Self-energy flows in the two-dimensional repulsive hubbard model. *Physical Review B*, 86(24), December 2012. doi:10.1103/physrevb.86.245122.

[77] J. Lichtenstein, D. S. de la Peña, D. Rohe, E. D. Napoli, C. Honerkamp, and S. A. Maier. High-performance functional renormalization group calculations for interacting fermions. *Computer Physics Communications*, 213:100–110, April 2017. doi:10.1016/j.cpc.2016.12.013.

[78] K. Veschgini and M. Salmhofer. *To appear soon*, 2018.

[79] W. Metzner, J. Reiss, and D. Rohe. Magnetic and superconducting correlations in the two-dimensional hubbard model. *physica status solidi (b)*, 243(1): 46–56, January 2006. doi:10.1002/pssb.200562445.

[80] J. Reiss, D. Rohe, and W. Metzner. Renormalized mean-field analysis of antiferromagnetism andd-wave superconductivity in the two-dimensional hubbard model. *Physical Review B*, 75(7), February 2007. doi:10.1103/physrevb.75.075110.

[81] A. Eberlein and W. Metzner. Superconductivity in the two-dimensionalt-t′-hubbard model. *Physical Review B*, 89(3), January 2014. doi:10.1103/physrevb.89.035126.

[82] J. Wang, A. Eberlein, and W. Metzner. Competing order in correlated electron systems made simple: Consistent fusion of functional renormalization and mean-field theory. *Physical Review B*, 89(12), March 2014. doi:10.1103/physrevb.89.121116.

[83] H. Yamase, A. Eberlein, and W. Metzner. Coexistence of incommensurate magnetism and superconductivity in the two-dimensional hubbard model. *Physical Review Letters*, 116(9), March 2016. doi:10.1103/physrevlett.116.096402.

[84] K. Veschgini. A combined functional renormalization and mean-field study of competing orders in the two-dimensional hubbard model. *PhD Thesis, Heidelberg University Library*, 2017. doi:10.11588/heidok.00022476.

[85] T. Baier, E. Bick, and C. Wetterich. Spontaneous symmetry breaking in the colored hubbard model. *Physical Review B*, 62(23):15471–15479, December 2000. doi:10.1103/physrevb.62.15471.

[86] T. Baier, E. Bick, and C. Wetterich. Temperature dependence of antiferromagnetic order in the hubbard model. *Physical Review B*, 70(12), September 2004. doi:10.1103/physrevb.70.125111.

[87] T. Baier, E. Bick, and C. Wetterich. Antiferromagnetic gap in the hubbard model. *Physics Letters B*, 605(1-2):144–150, January 2005. doi:10.1016/j.physletb.2004.11.022.

[88] H. C. Krahl and C. Wetterich. Functional renormalization group for d-wave superconductivity. *Physics Letters A*, 367(4-5):263–267, July 2007. doi:10.1016/j.physleta.2007.03.028.

[89] H. C. Krahl, J. A. Müller, and C. Wetterich. Generation ofd-wave coupling in the two-dimensional hubbard model from functional renormalization. *Physical Review B*, 79(9), March 2009. doi:10.1103/physrevb.79.094526.

[90] H. C. Krahl, S. Friederich, and C. Wetterich. Incommensurate antiferromagnetic fluctuations in the two-dimensional hubbard model. *Physical Review B*, 80(1), July 2009. doi:10.1103/physrevb.80.014436.

[91] S. Friederich, H. C. Krahl, and C. Wetterich. Four-point vertex in the hubbard model and partial bosonization. *Physical Review B*, 81(23), June 2010. doi:10.1103/physrevb.81.235108.

[92] S. Friederich, H. C. Krahl, and C. Wetterich. Functional renormalization for spontaneous symmetry breaking in the hubbard model. *Physical Review B*, 83 (15), April 2011. doi:10.1103/physrevb.83.155125.

[93] M. Gell-Mann. Symmetries of baryons and mesons. *Physical Review*, 125(3): 1067–1084, feb 1962. doi:10.1103/physrev.125.1067.

[94] M. Gell-Mann. A schematic model of baryons and mesons. *Physics Letters*, 8 (3):214–215, February 1964. doi:10.1016/s0031-9163(64)92001-3.

[95] G. Zweig. An SU_3 model for strong interaction symmetry and its breaking; Version 1. Technical Report CERN-TH-401, CERN, Geneva, January 1964. URL `http://cds.cern.ch/record/352337`.

[96] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. Advantages of the color octet gluon picture. *Physics Letters B*, 47(4):365–368, November 1973. doi:10.1016/0370-2693(73)90625-4.

[97] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Physical Review*, 96(1):191–195, October 1954. doi:10.1103/physrev.96.191.

[98] D. J. Gross and F. Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Physical Review Letters*, 30(26):1343–1346, June 1973. doi:10.1103/physrevlett.30.1343.

[99] H. D. Politzer. Reliable perturbative results for strong interactions? *Physical Review Letters*, 30(26):1346–1349, June 1973. doi:10.1103/physrevlett.30.1346.

[100] F. Karsch and E. Laermann. Thermodynamics and in-medium hadron properties from lattice qcd. In R. C. Hwa and X.-N. Wang, editors, *Quark-Gluon Plasma*, page 1, 2004. URL `http://adsabs.harvard.edu/abs/2004qgp..conf....1K`.

[101] A. Maire. Phase diagram of qcd matter : Quark-gluon plasma, June 2015. URL `https://cds.cern.ch/record/2025215`. CERN.

[102] K. G. Wilson. Confinement of quarks. *Physical Review D*, 10(8):2445–2459, October 1974. doi:10.1103/physrevd.10.2445.

[103] D. F. Litim. Optimized renormalization group flows. *Physical Review D*, 64 (10), October 2001. doi:10.1103/physrevd.64.105007.

[104] S. Pairault, D. Sénéchal, and A.-M. S. Tremblay. Strong-coupling expansion for the hubbard model. *Physical Review Letters*, 80(24):5389–5392, June 1998. doi:10.1103/physrevlett.80.5389.

[105] S. Mahlert, F. Gebhard, E. Jeckelmann, R. M. Noack, and S. Nishimoto. Fourth-order perturbation theory for the half-filled hubbard model in infinite dimensions. *The European Physical Journal B - Condensed Matter*, 36(4): 491–509, December 2003. doi:10.1140/epjb/e2004-00005-5.

[106] D. J. Scalapino. Numerical studies of the 2d hubbard model. In *Handbook of High-Temperature Superconductivity*, pages 495–526. Springer New York, 2007. doi:10.1007/978-0-387-68734-6_13.

[107] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst. Machine learning quantum phases of matter beyond the fermion sign problem. *Scientific Reports*, 7(1), August 2017. doi:10.1038/s41598-017-09098-0.

[108] M. Cristoforetti, G. Jurman, A. I. Nardelli, and C. Furlanello. Towards meaningful physics from generative models. *ArXiv e-prints*, May 2017. URL `http://adsabs.harvard.edu/abs/2017arXiv170509524C`.

[109] K. Huang. Connecting phase transition theory with unsupervised learning. *ArXiv e-prints*, December 2017. URL `http://adsabs.harvard.edu/abs/2017arXiv171205704H`.

[110] J. M. Lamim Ribeiro, P. Bravo Collado, Y. Wang, and P. Tiwary. Reweighted autoencoded variational bayes for enhanced sampling (rave). *ArXiv e-prints*, February 2018. URL `http://adsabs.harvard.edu/abs/2018arXiv180203420L`.

[111] P. Ponte and R. G. Melko. Kernel methods for interpretable machine learning of order parameters. *Physical Review B*, 96(20), November 2017. doi:10.1103/physrevb.96.205146.

Danksagung

An dieser Stelle bedanke ich mich bei Herrn Prof. Christof Wetterich für die Möglichkeit an dieser Doktorarbeit zu forschen, für die nützlichen Diskussionen über Physik und für die für die inspirierenden Ratschläge, von denen diese Doktorarbeit profitierte. Ich bedanke mich besonders für die Möglichkeit meine eigenen Leidenschaften zu verfolgen, und diese in diese Doktorarbeit einfließen zu lassen.

Bei Herrn Prof. Jan Martin Pawlowski bedanke ich mich für die interessanten Gespräche über Physik, die funktionale Renormierungsgruppe und auch Machine Learning, und außerdem für die Zweitbetreuung meiner Doktorarbeit.

Bei Herrn Prof. Thomas Gasenzer bedanke ich mich für sein Interesse und die Zweitbegutachtung dieser Arbeit.

An Herrn Prof. Selim Jochim und Herrn Prof. Björn Malte Schäfer vielen Dank für die Bereiterklärung mich in der Disputation zu prüfen.

Vielen Dank an die Kollegen und Koautoren für eine fruchtbare Zusammenarbeit bei den Forschungsartikeln, die Teil dieser Doktorarbeit sind. Dazu gehören neben den bereits genannten: Michael M. Scherer, Manuel Scherzer und Richard Schmidt.

Ich bedanke mich bei Timo Milbich, Björn Ommer, Manfred Salmhofer, Ion-Olimpiu Stamatescu, Kambis Veschgini für die interessanten Diskussionen, die für den Erfolg dieser Doktorarbeit unerlässlich waren.

Vielen Dank an Carsten Kempf, dem ich die Ehre verdanke einen sehr engagierten und interessierten Studenten für einige Monate seiner Masterarbeit betreut haben zu dürfen, und dessen Projekt eine inspirierende Grundlage für mein Interesse an neuronalen Netzen bildete.

Für die tolle Arbeitsatmosphäre bedanke ich mich bei den Kollegen des Institus

für theoretische Physik.

Vielen Dank an Mathias Neidig für die Freundschaft über das ganze Studium hinweg und für das Probelesen dieser Arbeit.

Ich bedanke mich bei Shirin Nkongolo für die liebevolle Unterstützung und für das Probelesen dieser Arbeit.

Vielen Dank an die vielen netten Menschen und Freunde, die mich während dieser Doktorarbeit begleiteten.

Mein ganz besonderer Dank gehört meiner Familie, Karl, Maria-Luise und Johannes Wetzel, die mich während dieser Doktorarbeit und des gesamten Studiums unterstützt haben.

Für die finanzielle Unterstützung bedanke ich mich bei der Heidelberg Graduate School of Fundamental Physics. Für eine ideelle Förderung möchte ich mich bei der International Max Planck Reseach School for Quantum Dynamics bedanken.

Erklärung

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 22.02.2018

.....

(Unterschrift)