Quantum Scattering and Machine Learning in Dirac Materials

by

Chendi Han

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved October 2022 by the
Graduate Supervisory Committee:

Ying-Cheng Lai, Chair
Hongbin Yu
Gautam Dasarathy
Jae-sun Seo

ARIZONA STATE UNIVERSITY

December 2022

ABSTRACT

A remarkable phenomenon in contemporary physics is quantum scarring in classically chaotic systems, where the wave functions tend to concentrate on classical periodic orbits. Quantum scarring has been studied for more than four decades, but the problem of efficiently detecting quantum scars has remained to be challenging, relying mostly on human visualization of wave function patterns. This paper develops a machine learning approach to detecting quantum scars in an automated and highly efficient manner. In particular, this paper exploits Meta learning. The first step is to construct a few-shot classification algorithm, under the requirement that the one-shot classification accuracy be larger than 90%. Then propose a scheme based on a combination of neural networks to improve the accuracy. This paper shows that the machine learning scheme can find the correct quantum scars from thousands images of wave functions, without any human intervention, regardless of the symmetry of the underlying classical system. This will be the first application of Meta learning to quantum systems.

Interacting spin networks are fundamental to quantum computing. Data-based tomography of time-independent spin networks has been achieved, but an open challenge is to ascertain the structures of time-dependent spin networks using time series measurements taken locally from a small subset of the spins. Physically, the dynamical evolution of a spin network under time-dependent driving or perturbation is described by the Heisenberg equation of motion. Motivated by this basic fact, this paper articulates a physics-enhanced machine learning framework whose core is Heisenberg neural networks. This paper demonstrates that, from local measurements, not only the local Hamiltonian can be recovered but the Hamiltonian reflecting the interacting structure of the whole system can also be faithfully reconstructed. Using Heisenberg neural machine on spin networks of a variety of structures. In the extreme

case where measurements are taken from only one spin, the achieved tomography fidelity values can reach about 90%. The developed machine learning framework is applicable to any time-dependent systems whose quantum dynamical evolution is governed by the Heisenberg equation of motion.

To my family, Yongjun Han, Aihua Zhou and Chenyang Han.

ACKNOWLEDGMENTS

First, I am full of gratitude to my advisor, Prof. Ying-Cheng Lai, who gave me guidance about the research direction, discussed problems I had met, and continuously supported me over the past five years. Without Prof. Lai's help, I could not achieve this. His deep insights and enthusiasm for physics have influenced and inspired me. His regular schedule also inspires me to form a good habits. I cannot forget his care for me in everyday life and his help whenever I met difficulties at ASU. And he also gives me so much advice and helps in my future career. I have learned a lot and feel very happy under his supervision.

I am also grateful to Prof. Liang Huang from Lanzhou University, who guide me to this area and recommended me this great opportunity at ASU. He is very knowledgeable about the new research and I learned a lot from the discussions with him.

I have a good time in our Chaos group. We discuss with and learn from each other and have many great memories. I especially want to thank Prof. Hongya Xu and Dr. Chengzhen Wang for their discussion and collaboration in the past five years. I thank students in our group, Prof. Junjie Jiang, Dr. Zhidan Zhao, Dr. Shirin Panahi, Mr. Lingwei Kong, Ms. Lili Ye, Mr. Zhengmeng Zhai, Mr. Amin Moradi, and visiting students and scholars Prof. Chun Zhang, Prof. Rui Xiao, Prof. Huawei Fan, Dr. Yu Meng, Dr. Shijie Zhou and others.

While job seeking, I received lots of help from ASU professors, previous students in our group, and my friends. I especially want to thank Prof. Oswald Chong for continuously recommending me opportunities. I also received suggestions and recommendations from Dr. Lezhi Wang, Dr. Xiaocong Du, Ms. Bowen Wang, Dr. Riqi Su, Dr. Yuzhong Chen, Dr. Rui Yang, Dr. Guanglei Wang. Thanks for their time.

I have traveled to lots of places in America in the past five years. Including San Diego, Los Angeles, San Francisco, Las Vegas, Death Valley National Park, Joshua Tree National Park, Zion National Park, Bryce Canyon National Park, Grand Canyon National Park, Rocky Mountain National Park, Washington DC and New York. I received lots of help from my friends, especially Mr. Zhongxing Zeng and my classmates from Lanzhou University. Thanks for them and I have a great experience here.

I want to give special thanks to my committee members, Prof. Hongbin Yu, Prof. Gautam Dasarathy and Prof. Jae-sun Seo, as well as the previous committee member Prof. Visar Berisha and Prof. Dragica Vasileska for their precious time as well as invaluable comments and advice.

Finally, I would like to express my deepest thankfulness to my family, especially to my beloved parents and my brother, who give all their love to me and support me in all my pursuits.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Chapter 1

INTRODUCTION

1.1 Classical Perspective of Relativistic Quantum Chaos

When chaos meets with (nonrelativistic) quantum mechanics, a set of signatures is left behind [1, 2, 3]. For example, in a closed chaotic Hamiltonian system, the probability for an infinitesimal separation between two neighboring energy levels diminishes [4, 5, 6], in contrast to an integrable system where this probability is most pronounced. In a closed chaotic billiard system, the eigen-wavefunctions can concentrate about the unstable periodic orbits that are the fundamental building blocks of classical chaos - the phenomenon of quantum scarring [7, 8]. In electronic transport through a quantum dot structure whose geometry generates chaos in the classical limit, the fluctuations in the quantum transmission or conductance can be greatly smoothed out in comparison with the situation where the dot geometry produces classically integrable dynamics [9]. Classical chaos can also lead to localization of the wavefunctions [10]. The study of the manifestations of classical chaos in the corresponding (nonrelativistic) quantum system constitutes the field of quantum chaos [1, 2, 3], which has been active for four decades.

Recently, *relativistic quantum chaos* (RQC) [11, 12] has emerged as a field to study the signatures of classical chaos in relativistic quantum systems. The main motivation for RQC came from the tremendous interest in and development of two-dimensional (2D) Dirac materials [13, 14] such as graphene [15, 16, 17, 18, 19, 20, 21] and topological insulators [22]. The energy bands of these materials typically contain a Dirac cone structure, stipulating a linear energy-momentum dispersion relation near a Dirac

point. In this energy regime, the quasiparticles of the materials are pseudospin-1/2 fermions and are described by the massless Dirac equation for a two-component pseudospinor. Studies of RQC have revealed a relatively stronger suppression of chaos in relativistic quantum than in nonrelativistic quantum systems. For example, in a study [23] on electronic transport through quantum dots that exhibit different types of classical scattering dynamics (e.g., integrable or chaotic), in the Schrödinger system, classical chaos can dramatically smooth out the sharp fluctuations associated with resonances in quantum transmission or conductance that would be present if the classical dynamics were integrable, a phenomenon that can be exploited for chaos based modulation or control of quantum transport characteristics [24, 25]. However, in a relativistic quantum graphene dot, sharp resonances are still present in the transmission even when the corresponding classical dynamics become fully chaotic [23]. Another example is persistent currents [26, 27, 28, 29], permanent currents in the absence of any external power source, in a ring domain with a central magnetic flux that breaks the time reversal symmetry. The currents have been observed in a variety of nonrelativistic quantum material systems [30, 31, 32, 33, 34, 35, 36, 37]. Random impurities in metallic or semiconductor systems tend to diminish the persistent currents [38, 39, 40, 41, 42, 43, 44, 45], where they decay exponentially to zero as the disorder strength is increased. In fact, in a Schrödinger ring system, boundary deformations leading to classical chaos typically destroy persistent currents [46, 47] as effectively as random disorder. However, in graphene or other Dirac materials, persistent currents were found to be robust [48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60]. A theoretical study demonstrated that, in Dirac ring systems, even when there are substantial boundary deformations leading to fully developed chaos in the classical limit, persistent currents of comparable magnitude with that in the integrable ring system can arise (henceforth the term "superpersistent currents") [46, 47]. From the

point of view of the quantum signatures of classical chaos, it can be understood that the signatures of chaos are less pronounced in the Dirac ring system as compared with the corresponding Schrödinger ring.

To establish the weakening of chaos in relativistic quantum systems on a comprehensive analysis of the out-of-time-order correlator (OTOC) [61]. Historically, OTOC was proposed in the field of superconductivity [62]. In basic quantum mechanics, OTOC measures the degree of commutation between two operators, which can be used to characterize diverse physical phenomena such as operator spreading [63, 64], growth in quantum entanglement [65, 66], and nuclear magnetic resonances [67]. Recently, the study of OTOC has attracted a great deal of attention due to its relevance to a large number of fields in physics. For example, OTOC has been employed to search for the quantum butterfly effect in many body systems such as those described by the Ising model [68, 69] and the Sachdev-Ye-Kitaev (SYK) model [70, 71]. In high energy physics, OTOC has been exploited to establish the correspondence between conformal field theory (CFT) and holography [72]. More recently, OTOC has been introduced into the field of quantum chaos [72, 65, 70, 73, 74, 75, 76, 77, 68, 71, 66, 67, 69]. To our knowledge, all existing studies on OTOC to uncover and characterize the signatures of classical chaos were for nonrelativistic quantum systems described by the Schrödinger equation.

We focus on the semiclassical regime in which the plane wave approximation is valid and Fresnel's law is applicable [78]. Intuitively, due to the total absence of backscattering of massless spin-1/2 particles [79, 80] and the purely relativistic quantum phenomenon of Klein tunneling [81, 82, 83], the decay of the spinor wave would be enhanced when comparing with that of classical electromagnetic waves from the same cavity, so trapping of the former would seem impossible. Indeed, a detailed scaling analysis of the ratio of the mean escape time of an electromagnetic wave to that

of a spin-1/2 wave reveals that, for both integrable and chaotic cavities, in the regime of large effective refractive index values ($n \gg 1$), the ratio is proportional to n but in the regime of $n \ll 1$, the ratio is inversely proportional to n. This means that, in these two asymptotic regimes, the averaging lifetime of the spin-1/2 wave is indeed much smaller than that of the electromagnetic wave. The surprising phenomenon is that, in between the two asymptotic regimes, an interval in n emerges, in which the ratio is about one, indicating that the spin-1/2 wave can live as long as the electromagnetic counterpart. This means that, high Q can be achieved for spin-1/2 particles. Since the constant ratio also holds for classical chaotic cavities, non-isotropic coherent emission can be expected. The finding suggests strongly that the two microlasing conditions for photons can be fulfilled for spin-1/2 particles. Our analysis provides insights into Dirac electronics and photonics, and has potential applications in developing unconventional cavity laser designs based on Dirac photonic crystals, and optical-like electronics with 2D Dirac materials.

1.2 Scattering and Transport in Dirac Materials

We investigate atomic collapse in pseudospin-1 Dirac material systems whose energy band structure constitutes a pair of Dirac cones and a flat band through the conic intersecting point [84, 85]. We obtain analytic solutions of the Dirac-Weyl equation for the three-component spinor in the presence of a Coulomb impurity and derive a general criterion for the occurrence of atomic collapse in terms of the normalized strength of Coulomb interaction and the angular momentum quantum number. In particular, for the lowest angular momentum state, the solution coincides with that for pseudospin-1/2 systems, but with a reduction in the density of resonance peaks. For higher angular momentum states, the underlying pseudospin-1 wave functions exhibit a singularity at the point of zero kinetic energy. Divergence of the local density

of states associated with the flat band leads to an inverse square type of singularity in the conductivity. These results provide insights into the physics of the two-body problem for relativistic quantum pseudospin-1 quasiparticle systems.

In Chapter 2, we study confinement of quasiparticles in a cavity in an α-T_3 lattice, which can be created by applying a step electrostatic potential, the boundary of which divides the lattice system into two regions, as shown in Fig. 2.1(b) [86]. Experimentally, such a structure can be generated by a STM-tip induced potential [87, 88, 89, 90] or through the method of doping [91]. The geometric shape of the cavity can be chosen to generate integrable (e.g., a circle), mixed (e.g., an ellipse), or chaotic (e.g., a stadium) dynamics in the classical limit. In order to confine an electron inside the cavity, its energy should be far away from the Klein tunneling regime that occurs for $E \approx U/2$ for graphene [83, 92, 93, 94, 95, 96, 89] and pseudospin-1 materials [96]. For the confinement problem to have physical and applied significance, we focus on the quantum-dot regime where the effect of Klein tunneling is weak [97, 78, 98]. Firstly, we choose the incident energy E such that it is much smaller than the electric potential: $|E| \ll |U|$, as shown by the energy band structure in Fig. 2.1(c), so as to avoid the Klein tunneling regime. The wavevector inside the cavity is thus much larger than that outside. Secondly, we choose the size of the cavity such that the system in the quantum-dot regime defined by $k_0 R \ll 1 \ll VR$, where k_0 is the wavevector outside, R is the effective size (radius) of the cavity, and $V = U/v_g$ is some normalized potential strength (with v_g being the group velocity). In this regime, outside of the cavity the wave characteristics of the quasiparticles dominate but inside the cavity the particle nature becomes important - there are then carriers in the cavity. The linear dispersion relation associated with the Dirac cones is $E = v_g k_0$. The electron motion inside the cavity can be studied in terms of Dirac electron optics [99, 100, 101, 92, 102, 103, 104, 105, 106, 107, 108, 109, 110, 87, 111, 112, 113,

114, 115, 89, 116, 117, 118, 119, 120, 121, 122, 123]. As the value of the pseudospin parameter α increases from zero to one, the nature of the quasiparticles of the system changes from pseudospin-1/2 to hybrid and finally to pseudospin-1. For $\alpha \neq 0, 1$, the time reversal symmetry (T-symmetry) is broken [124].

We uncover a chaos based mechanism to modulate pseudospin [125]. Focusing on graphene, we exploit the principle of mutual coupling between a waveguide and a nearby cavity. The idea is originated from optics, where the asymmetric configuration of coupled waveguide and microcavity is commonly used to induce Fano resonances in various optical devices [126, 127, 128, 129]. Figure illustrates such a transport system, where the graphene waveguide, a channel-like structure through which pseudospin-1/2 quasiparticles propagate, can be fabricated by applying some properly designed electrical potential [130, 131, 132, 133] and has been experimentally realized [105, 111]. Experimentally, a dielectric cavity in graphene can also be created through an external gate potential [87, 88, 89, 91, 90], as in scattering [80, 108, 110, 134]. In our work, the shape of the cavity is chosen such that the corresponding classical dynamics are fully chaotic ((e.g., the shape of Africa) [6, 135]. Due to the proximity of the waveguide and cavity, there is interaction between them, forming bonding and anti-bonding states and effectively generating an artificial molecular structure on graphene [136, 137].

We investigate the "complete" optical responses of 2D Dirac materials with a flat band [138], in the sense that both the real and imaginary parts of the optical conductivity are derived, using the α-$\mathcal{T}_3$ lattice as a paradigmatic model system of such materials. This lattice is formed by adding an atom at the center of each unit cell of the honeycomb graphene lattice [139], where the low energy excitations can be described by the pseudospin-1 Dirac-Weyl equation. The parameter α characterizes the interaction strength between the central atom and any of its nearest neighbors, relative to that between two neighboring atoms at the vertices of the hexagonal cell.

For $\alpha = 0$ there is no coupling between the central atom and a vertex , so the lattice degenerates to graphene with pseudospin-1/2 quasiparticles. As the value of α increases from zero to one, a flat band through the conic interaction of the two Dirac cones emerges and its physical influences become progressively pronounced [140, 141]. For $\alpha = 1$, the lattice generates pseudospin-1 quasiparticles. The flat band can lead to physical phenomena such as the divergence of conductivity [142, 143, 84]. Under a continuous approximation, an α-$\mathcal{T}_3$ lattice can be treated as a thin layer with certain surface conductivity. Unlike graphene, here the surface conductivity is contributed to by three types of transitions between the bands: intraband transition, cone-to-cone transition, and flat-band-to-cone transition. Previously, the optical conductivity of α-$\mathcal{T}_3$ lattice was "partially" studied in the sense that only the real part of the conductivity has been derived [141, 144, 145, 146]. Considering that the imaginary part can affect the optical response significantly and is therefore important for developing α-$\mathcal{T}_3$ lattice based optical devices, we are led to derive a complete conductivity formula with both the real and imaginary parts. The formula is verified through two independent approaches and leads to two previously uncovered phenomena. First, while the intraband transition leads to TM polarized waves at low frequencies (1-10 THz), TE polarized waves can emerge at high frequencies (100-300 THz), due to the two interband transitions. Second, the unique flat-band-to-cone transition generates multi-frequency TE propagating waves for $\alpha \in (0.4, 0.6)$ and a strong optical response for $\alpha = 1$. These phenomena are confirmed through studying the behaviors of propagating surface wave and scattering.

1.3 Physics Enhanced Machine Learning and Few Shot Classification

In Chapter 3, we address adaptability, a fundamental issue in machine learning, of Hamiltonian neural networks [147]. More precisely, we consider the situation where

a target Hamiltonian system can experience slow drift or sudden changes in some parameters. Slow environmental variations can lead to adiabatic parameter drifting, while external disturbances can lead to sudden parameter changes. We ask if it is possible to design HNNs, which are trained with data from a small number of parameter values of the target system, to have the predictive power for parameter values that are not in the training set. Inspired by the recent work on predicting critical transitions and collapse in dissipative dynamical systems based on reservoir computing [148, 149, 150, 151, 152], we articulate a class of HNNs whose input layer contains a set of channels that are specifically used for inputting the values of the distinct parameters of interest to the neural network. The number of the parameter channels is equal to the number of freely varying parameters in the target Hamiltonian system. The simplest case is where the target system has a single bifurcation or control parameter so only one input parameter channel to the neural network is necessary. We demonstrate that, by incorporating such a parameter channel into a feed-forward type of HNNs and conducting training using time series data from a small number of bifurcation parameter values (e.g., four), we effectively make the HNN adaptable to parameter variations. That is, the so-trained HNN has inherited the rules governing the dynamical evolution of the target Hamiltonian system. When a parameter value of interest, which is not in the training parameter set, is fed into the HNN through the parameter channel, the machine is capable of generating dynamical behaviors that statistically match those of the target system at this particular parameter value. The HNN has thus become adaptable because it has never been exposed to any information or data from the target system at this parameter value, yet the neural machine can reproduce the dynamical behavior. Using the Hénon-Heiles model as a prototypical target Hamiltonian system, we demonstrate that our adaptable HNN can successfully predict the dynamical behaviors, integrable or chaotic, for

any parameter values that are reasonably close to those in the training parameter set. Remarkably, by feeding a systematically varying set of bifurcation parameter values into the parameter channel, the HNN can successfully predict the transition to chaos in the target Hamiltonian system, which we characterize using two measures: the ensemble maximum Lyapunov exponent and the alignment index. It is worth emphasizing that, in the existing literature on HNNs [153, 154, 155, 156, 157], training and prediction are done at the same set of parameter values of the target Hamiltonian system, but our work goes beyond by making the HNN significantly more powerful with enhanced and expanded predictability.

In this Chapter 4, we solve a general class of inverse problems in spins systems by exploiting machine learning [158, 159]. Our work was partly inspired by the recent work on the classical Hamiltonian Neural Networks (HNNs) [160, 153, 155, 156], where the basic idea is to introduce a physics-based, customized loss function to "force" the dynamical evolution of the system to follow that stipulated by the classical Hamilton's equations. However, the existing HNNs are not directly applicable to quantum Hamiltonians, thereby requiring new approaches. Our idea originates from the basic physical consideration that the dynamical evolution of spins systems is governed by the Heisenberg equations of motion. We are thus motivated to develop a class of Heisenberg Neural Networks (HENNs) by exploiting deep learning to predict the Hamiltonian but under the constraint of the Heisenberg equations of motion. The HENNs have the advantage of guaranteeing that the underlying quantum evolution possesses the Hermitian structure. Our assumptions for the quantum Hamiltonian are listed as follows. (i) Quantum Hamiltonian varies continuously with time, a situation that can be expected to hold in experiments in general. (ii) The measured spins are coupled with the rest of the spins directly or indirectly. (iii) There are no Pauli terms in the Hamiltonian commute with each other.

In this Chapter 5, we exploit Meta learning [161] to develop an automated and efficient quantum scar detector. The essence of Meta machine learning is to encode "previous experience" into a pretrained neural network so that it can quickly adapt to new input. Meta learning has found broad applications in regression, classification and reinforcement learning. Our idea is to exploit some existing data sets to train the neural network, which can be completely unrelated to quantum systems but with images similar to scarring patterns. In this regard, the Omniglot dataset [162] widely used in the field of image classification stands out as an excellent choice. Our strategy is to train a "preliminary" version of the neural network that has the ability to distinguish different classes of noisy images from the Omniglot dataset and then perform few-shot classification to further train the neural network with a small number of quantum scars. A pertinent issue is that standard Meta learning algorithms can only be used to classify but not detect patterns. Our solution is to use an ensemble of neural networks. In particular, non-scarring states as input to the neural networks can be used as references with significantly different statistical features than those associated with scarring states, thereby accomplishing the detection task. We demonstrate that the neural network so trained can detect and classify quantum scars with remarkable efficiency and accuracy even with limited training data of actual quantum scars. Our Meta-learning based approach not only solves a long-standing problem in the field of quantum chaos, but can also be generalized to address challenging image detection and classification problems in other fields.

1.4 Outline of This Thesis

In Chapter 1, we give the introduction and background for the systems and phenomena we investigate. In Chapter 2, we study the quantum scattering of an electrically generated potential cavity in the quantum-dot regime, we find that the regime

of small α values offers the best confinement possible among the spectrum of $\alpha - T_3$ materials. In Chapter 3, we introduce a class of HNNs capable of adaptable prediction of nonlinear physical systems: by training the neural network based on time series from a small number of bifurcation-parameter values of the target Hamiltonian system, the HNN can predict the dynamical states at other parameter values. In Chapter 4, we articulate a physics-enhanced machine-learning framework whose core is Heisenberg neural networks. In particular, we develop a deep learning algorithm according to some physics-motivated loss function based on the Heisenberg equation. In Chapter 5, we show that the Meta learning scheme can find the correct quantum scars from thousands images of wavefunctions without any human intervention.

Chapter 2

ELECTRICAL CONFINEMENT IN A SPECTRUM OF DIRAC MATERIALS WITH CLASSICALLY INTEGRABLE, MIXED, AND CHAOTIC DYNAMICS

An emergent class of two-dimensional Dirac materials is α-T_3 lattices that can be realized by adding an atom at the center of each unit cell of a lattice with T_3 symmetry. The interaction strength α between this atom and any of its nearest neighbors is a parameter that can be continuously tuned between zero and one to generate a spectrum of materials. We investigate the fundamentally important and practically relevant issue of quasiparticle confinement for the entire spectrum of α-T_3 materials. Except for the two end points, i.e., $\alpha = 0, 1$, which correspond to the graphene and pseudospin-1 lattices, respectively, the time-reversal symmetry is broken, leading to the removal of level degeneracy and facilitating confinement. Taking the approach of quantum scattering off an electrically generated potential cavity in the quantum-dot regime, we characterize confinement by identifying and examining the scattering resonances. We study a number of cavities with characteristically distinct classical dynamics: circular, annular, elliptical and stadium cavities. For the circular and annular cavities with classically integrable and mixed dynamics, respectively, the scattering matrix can be analytically obtained, so the scattering cross sections and the Wigner-Smith time delay associated with the resonances can be calculated to quantify confinement. For the elliptical and stadium cavities with mixed and chaotic dynamics in the classical limit, respectively, the scattering-matrix approach is infeasible, so we adopt an efficient numerical method to calculate the scattering wavefunctions and experimentally accessible measures of confinement such as the magnetic moment. The main finding is that, for all the cases, the regime of small α values offers the

best confinement possible among the spectrum of α-T_3 materials, which is general and holds regardless of the nature of the corresponding classical dynamics.

2.1 Introduction

There has been tremendous development of research on two-dimensional (2D) Dirac materials since the experimental realization of graphene [15, 17, 19, 163]. A distinct characteristic of the low-energy excitations in 2D Dirac materials is the pseudospin degree of freedom. A variety of such materials have been studied, suggesting that a continuous spectrum of pseudospin quasiparticles may exist in these materials and may be experimentally realized. At the lower end of the spectrum is graphene, whose energy band structure constitutes a pair of Dirac cones with the corresponding low-energy excitations being pseudospin-1/2 particles, which are described by a two-component spinor wavefunction governed by the standard 2D Dirac equation [19]. At the high end of the spectrum are Dirac materials with a T_3 symmetry, whose energy band contains a pair of Dirac cones and a flat band through the conic intersecting point [139]. Because of the existence of three distinct bands, the low-energy excitations need to be described by a spinor wavefunction of three components, corresponding effectively to pseudospin-1 quasiparticles that obey the Dirac-Weyl equation. In between the pseudospin-1/2 and pseudospin-1 extremes lies a spectrum of pseudospin quasiparticles that can be generated by the corresponding spectrum of α-T_3 lattices [140, 141, 164, 165, 145, 124, 166, 167, 168, 98, 169].

The relativistic quantum behaviors of α-T_3 particles are described by the generalized Dirac-Weyl equation with a three-component spinor, where the original 2×2 Pauli matrices (for pseudospin-1/2 particles) are replaced by a set of 3×3 matrices - the set of generalized Pauli matrices. Figure 2.1(a) illustrates the structure of an α-T_3 lattice with a T_3 symmetry, where the unit cell is essentially that of the graphene

honeycomb lattice with an additional atom at the center of the hexagon. There are then three nonequivalent atoms in the α-T_3 unit cell, where α is a parameter characterizing the coupling strength between the central atom and any of the six atoms at the vertices of the hexagon [140]. Especially, if in the graphene lattice the nearest-neighbor interaction energy is t, then the interaction energy between the central atom and one on the hexagon is αt.

The possible values of α range from zero to one. For $\alpha = 0$, the presence of the central atom has no effect on the hexagonal lattice, so the whole lattice effectively reduces to that of graphene with pseudospin-1/2 quasiparticles. For $\alpha = 1$, the interaction energy between the central atom and one on the hexagon is identical to that of the nearest neighbor interaction in the graphene lattice, so the α-T_3 lattice possesses the full T_3 symmetry with pseudospin-1 particles. For $\alpha \in (0, 1)$, the α-T_3 lattice is essentially a hybrid between the graphene and pseudospin-1 lattices [164]. For convenience, we call α the pseudospin parameter.

Experimentally, one way to realize α-T_3 hybrid and pseudospin-1 systems through photonic crystals [170, 171, 172]. The three nonequivalent atoms can be simulated by using coupled waveguides generated by laser inscription [172]. Different lattice constants can be used to build a quantum dot structure [173]. Electronic materials can also be exploited to generate pseudospin-1 lattice systems such as transition-metal oxide $SrTiO_3/SrIrO_3/SrTiO_3$ trilayer heterostructures [174], $SrCu_2(BO_3)_2$ [175] or graphene-In_2Te_2 [176].

In potential device applications of α-T_3 materials, a fundamental issue is how to effectively confine the quasiparticles in an enclosure. Similar to optics [177], such a structure can be exploited for storage and energy transfer in spintronics and valleytronics [178, 179]. However, even in a perfectly circular cavity generated by, e.g., an electrostatic potential, confinement of pseudospin-1/2 particles in

graphene is already a non-trivial issue [78], due to the phenomenon of Klein tunneling [83, 92, 93, 94, 95, 96, 89]. To confine pseudospin-1 particles is also difficult, due to super-Klein tunneling [96] in which particles can penetrate through a high and wide potential barrier at any angle. Confinement becomes even more difficult in realistic situations where geometric deformations from the circular shape are inevitable, which can lead to chaotic dynamics in the classical limit [180, 181, 129]. What are the general features of confining pseudospin quasiparticles of the α-T_3 lattice in comparison with the confinement of pseudospin-1/2 and pseudospin-1 particles? The purpose of this paper is to address this question that not only is useful for gaining understanding into the fundamental physics of the quasiparticles in the exotic 2D materials but also has implications to developing future α-T_3 material based electronic/spintronic devices.

To be concrete, we study confinement of quasiparticles in a cavity in an α-T_3 lattice, which can be created by applying a step electrostatic potential, the boundary of which divides the lattice system into two regions, as shown in Fig. 2.1(b). Experimentally, such a structure can be generated by a STM-tip induced potential [87, 88, 89, 90] or through the method of doping [91]. The geometric shape of the cavity can be chosen to generate integrable (e.g., a circle), mixed (e.g., an ellipse), or chaotic (e.g., a stadium) dynamics in the classical limit. In order to confine an electron inside the cavity, its energy should be far away from the Klein tunneling regime that occurs for $E \approx U/2$ for graphene [83, 92, 93, 94, 95, 96, 89] and pseudospin-1 materials [96]. For the confinement problem to have physical and applied significance, we focus on the quantum-dot regime where the effect of Klein tunneling is weak [97, 78, 98]. Firstly, we choose the incident energy E such that it is much smaller than the electric potential: $|E| \ll |U|$, as shown by the energy band structure in Fig. 2.1(c), so as to avoid the Klein tunneling regime. The wavevector inside the cavity is thus much larger

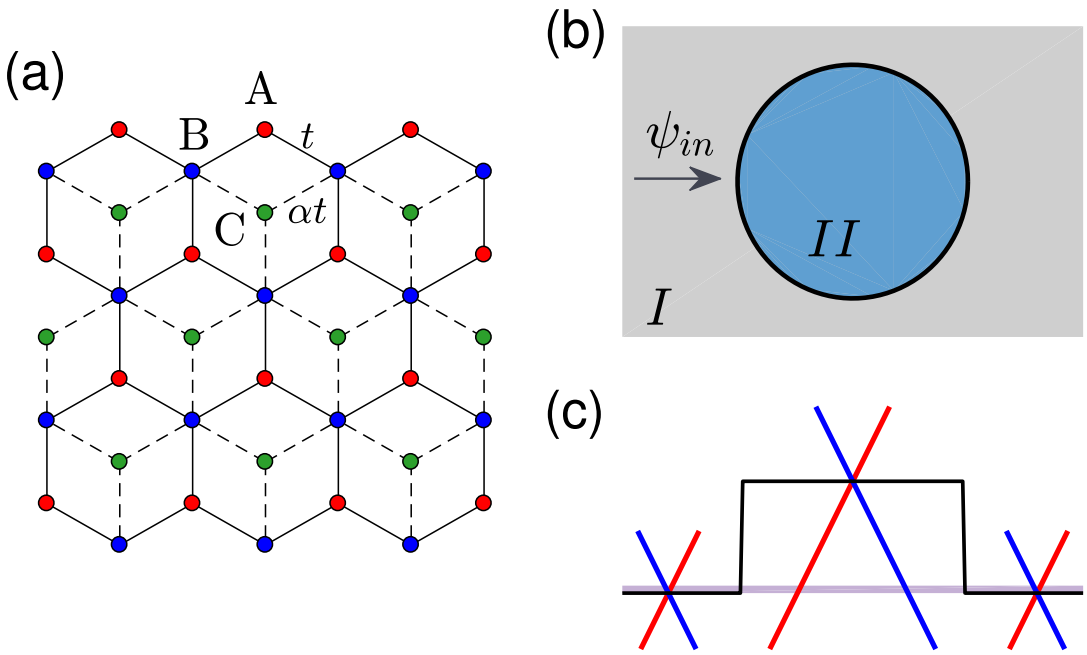

Figure 2.1: Illustration of an α-T_3 lattice, confinement cavity, and the energy band structure. (a) An α-T_3 lattice, where A, B and C represent the three non-equivalent atoms, labeled as red, blue and green circles. The hopping energy between A and B (solid line) is t, and that between B and C (dashed line) is αt. (b) A circular cavity (blue) in an α-T_3 sheet (gray), which can be realized through an externally applied electric potential. Two geometric regions are specified: one outside and another inside the cavity. (c) For pseudospin-1/2 system ($\alpha = 0$), energy-band structure for scattering, where carriers dominate inside the cavity, forming a quantum-dot structure. Blue and red lines correspond to the two different linear dispersion relationship which associate with two non-equivalent atoms [83]. The horizontal gray line illustrates the incident energy, which is much lower than the potential height.

than that outside. Secondly, we choose the size of the cavity such that the system in the quantum-dot regime defined by $k_0 R \ll 1 \ll VR$, where k_0 is the wavevector outside, R is the effective size (radius) of the cavity, and $V = U/v_g$ is some normalized potential strength (with v_g being the group velocity). In this regime, outside of the cavity the wave characteristics of the quasiparticles dominate but inside the cavity the particle nature becomes important - there are then carriers in the cavity. The linear dispersion relation associated with the Dirac cones is $E = v_g k_0$. The electron motion inside the cavity can be studied in terms of Dirac electron op-

tics [99, 100, 101, 92, 102, 103, 104, 105, 106, 107, 108, 109, 110, 87, 111, 112, 113, 114, 115, 89, 116, 117, 118, 119, 120, 121, 122, 123]. As the value of the pseudospin parameter α increases from zero to one, the nature of the quasiparticles of the system changes from pseudospin-1/2 to hybrid and finally to pseudospin-1. For $\alpha \neq 0, 1$, the time reversal symmetry (T-symmetry) is broken [124].

The main finding of this paper is that, in the quantum-dot regime, among the possible α-T_3 materials, the strongest or optimal confinement occurs for hybrid materials in between the pseudospin-1/2 and pseudospin-1 limits but near the graphene end, i.e., for some value of $\alpha \leq 0$. This result holds not only for the perfectly circular cavity with classical integrable dynamics, but also for deformed cavities with mixed or chaotic dynamics in the classical limit. Contributing factors to this phenomenon include T-symmetry breaking in the hybrid material and unconventional wavefunction behaviors as induced by the boundary conditions.

In Sec. 2.2, we analyze the scattering process from circular cavity for α-T_3 materials by calculating the quasibound states and the far-field behavior. Especially, because of the circular symmetry, in the quantum-dot regime there are quasibound states with nonzero angular momentum whose lifetime is infinite, and the confinement efficacy is determined the number of such states. We find that, for the hybrid system, because of the T-symmetry breaking, more such quasibound states can emerge as compared with the graphene and pseudospin-1 limits. In Sec. 2.3, we study the annular cavity for which the scattering process can still be understood analytically. Calculation of the Wigner-Smith time delay for different values of α and different potential profiles indicates that, for certain potential profile, maximum confinement occurs for a small but non-zero value of α. Finally, in Sec. 2.4, we study confinement in deformed cavities with classical mixed and chaotic dynamics, which find applications in enhancing directional emission. In this case, all the quasibound states have

a finite lifetime. However, practically, confinement is still possible, which can be characterized by physically measurable quantities such as the magnetic moment.

2.2 Confinement in a Circular Cavity

We consider the tight-binding model for a free α-T_3 particle and derive the corresponding continuum Hamiltonian. In the tight-binding framework, the Bloch Hamiltonian of the α-T_3 lattice system is given by

$$H(\mathbf{k}) = \begin{pmatrix} 0 & f_{\mathbf{k}} & 0 \\ f_{\mathbf{k}}^* & 0 & \alpha f_{\mathbf{k}} \\ 0 & \alpha f_{\mathbf{k}}^* & 0 \end{pmatrix} \tag{2.1}$$

where

$$f_{\mathbf{k}} = -t[e^{ik_y a_t} + 2e^{-ik_y a_t/2}\cos(\sqrt{3}/2 k_x a_t)]$$

and a_t is the lattice constant. Expanding $f_{\mathbf{k}}$ at zero yields $f_{\mathbf{k}} \approx v_g(\nu k_x - ik_y)$, where $v_g = 3a_t t/2$ and $\nu = \pm$ is the valley index. Imposing the normalization $t \to t/\sqrt{1+\alpha^2}$, we get the continuum model for low-energy excitations in the α-T_3 lattice as

$$H = v_g \boldsymbol{S}^\alpha \cdot \boldsymbol{p}, \tag{2.2}$$

where $\boldsymbol{S}^\alpha$ denotes the generalized Pauli matrices that depend on the material parameter α. The details for solving the corresponding eigenvalue problem are presented in Appendix 2.6.

Quantum scattering from a circular cavity has been studied for graphene ($\alpha = 0$) [80, 108, 110, 134] and pseudospin-1 ($\alpha = 1$) [96, 182] systems. There has also been a study of scattering from a centrally symmetric potential in α-T_3 materials [169]. Because of the circular symmetry in the potential profile, the scattering problem can be solved analytically.

The Hamiltonian for a general α-T_3 lattice with a circular cavity is

$$H = v_g \mathbf{S}^\alpha \cdot \mathbf{p} + v_g V \Theta(R - r), \tag{2.3}$$

where V is potential height, Θ is Heaviside step function, and $\mathbf{S}^\alpha$ are the generalized Pauli matrices that depend on the parameter α (detailed in Appendix 2.6). For this Hamiltonian, the three-component spinor wavefunction can be obtained analytically (Appendix 2.7).

Let q be the wavevector inside the cavity. The linear dispersion relation gives $qv_g = k_0 v_g - V$. In the quantum-dot regime $k_0 R \ll 1 \ll VR$, we have $|k_0| \ll |q|$. In the language of Dirac electron optics, waves inside the cavity will have a large relative refractive index, rendering existent a critical angle for total internal reflection [97, 78, 98], which makes confinement possible.

There are different ways to characterize confinement quantitatively. For example, we can use the total scattering cross section defined as

$$\sigma(\theta') = \oint |f(\theta, \theta')|^2 d\theta, \tag{2.4}$$

where θ' is the incident angle, θ specifies the direction of measurement, and $f(\theta, \theta')$ is determined by the far field behavior of the scattering wave (Appendix 2.6). A large total cross section corresponds to stronger scattering. Alternatively, the wavefunction inside the cavity can be used to quantify confinement. With the analytically obtained three-component spinor wavefunction $\Psi = (\psi_A, \psi_B, \psi_C)^T$ inside the cavity, we have the average density of state (DOS) as

$$\text{DOS} = \int_{\text{cavity}} \left(|\psi_A|^2 + |\psi_B|^2 + |\psi_C|^2 \right). \tag{2.5}$$

We can normalize the DOS to unit area after integration. In experiments, information about the average DOS can be obtained through conductivity measurement [87, 88,

89, 90]. Another useful quantity to characterize confinement is the Wigner-Smith time delay [183, 184] defined through the scattering matrix S:

$$\tau = -i\hbar \mathrm{Tr}\ \left(S^{\dagger} \frac{\partial S}{\partial E} \right). \tag{2.6}$$

A large positive value of τ is indicative of a confinement state [185].

We consider the quantum-dot regime $k_0 R \ll 1 \ll VR$ where the incident wavevector is infinitesimally small for a fixed potential height but the wavevector is large inside the cavity. (In terms of the small wavelength inside the cavity, this regime can be called the semiclassical regime.) Figure 2.2(a) shows, in the parameter plane (α,$k_0 R$), the total scattering cross section σ with color-coded values. We see that σ exhibits peak values, each corresponding to a quasibound state. The dominant peaks arise in the small $k_0 R$ regime and depend on the value of α as well. The result in Fig. 2.2(a) suggests that, in the quantum-dot regime, as the value of α is varied, there can be drastic changes in the quantum states from the point of view of confinement. This can be further seen in Fig. 2.2(b) that shows, for three different values of α ($\alpha = 0, 0.1, 1$), σ versus $k_0 R$. While there are peaks in $10^{-1} < k_0 R < 10^{1}$ for all three cases, in the small region $10^{-2} < k_0 R < 10^{-1}$, only the $\alpha = 0.1$ case has seen a peak. The existence of a strong quasibound state in $10^{-2} < k_0 R < 10^{-1}$ for $\alpha = 0.1$ (but not for $\alpha = 0$ or $\alpha = 1$) is further demonstrated by the average DOS and the Wigner-Smith time delay versus $k_0 R$, as shown in Figs. 2.2(c) and 2.2(d), respectively.

A heuristic understanding of the emergence of strong confinement states associated with α-T_3 scattering in the hybrid system can be gained by examining the far-field behavior such as backscattering. For graphene, due to the $\pi/2$ Berry phase, in the low energy regime backscattering is ruled out [79, 108]. For pseudospin-1 systems, scattering in the far field is isotropic [96]. A question is: what is the far-field behavior in the scattering of α-T_3 particles for $0 < \alpha < 1$? To address this question, we analyze

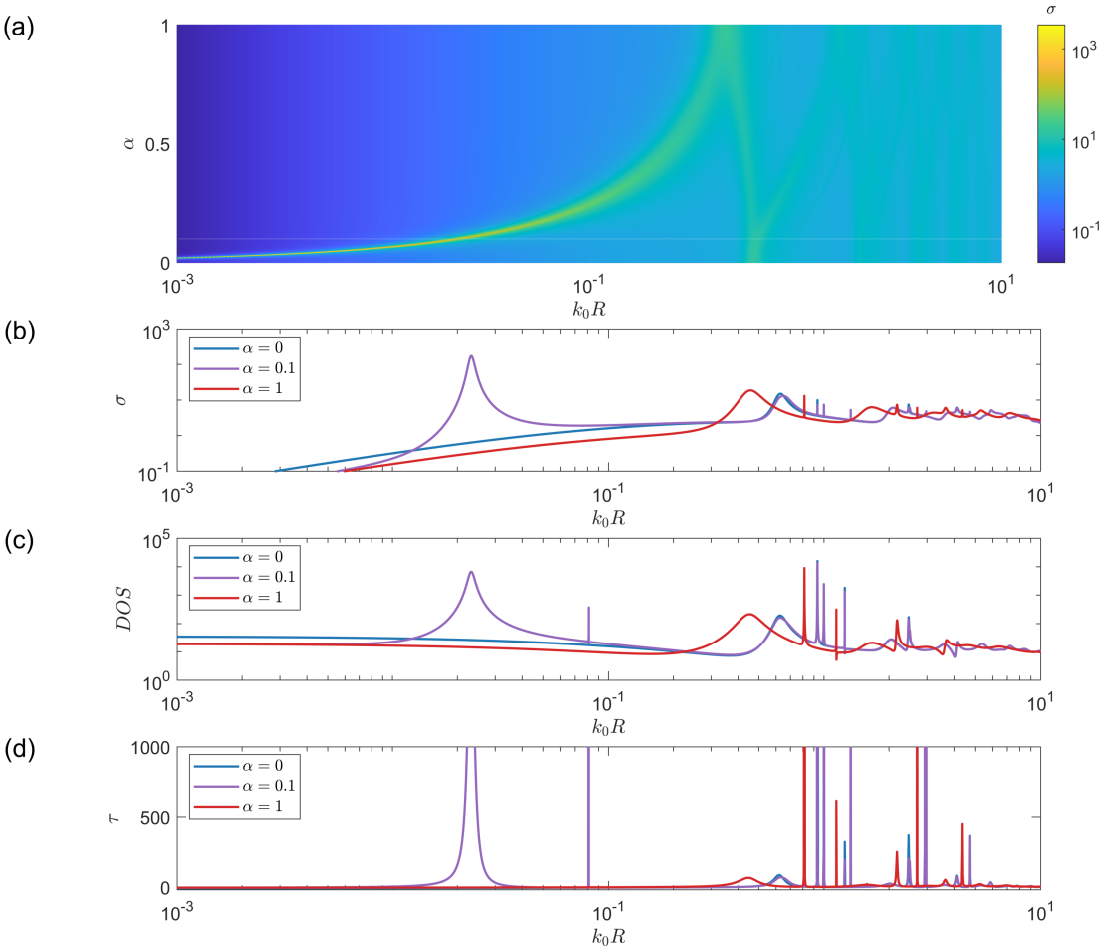

Figure 2.2: Scattering of α-T_3 particles in the regime of small incident wavevectors. The circular electrical cavity has $VR = 30$. (a) Color-coded values of the scattering cross section σ in the parameter plane (k_0R,α), where a local peak (resonance) in σ indicates the existence of a quasi-bound state and the inverse of the half width of the peak is proportional to the lifetime of the state. There is dominance of resonant peaks in the regime of small α values. (b-d) Behaviors of σ, average DOS, and Wigner-Smith time delay versus k_0R for three values of α: 0, 0.1, and 1. In the region of small k_0R values, only the $\alpha = 0.1$ case exhibits a strong peak, indicating a superior ability for α-T_3 material with $\alpha = 0.1$ to confine electrons to those of graphene and pseudospin-1 materials.

different scattering channels. For pseudospin-1/2 system (graphene), there are two lowest angular momentum states: $\pm 1/2$. The counterparts of these states in α-T_3 scattering in the hybrid system are $l = 1$ and $l = 0$. Let A_l be the scattering wave amplitude with angular momentum l. In the semiclassical regime, only the $l = 0$ and $l = 1$ channels contribute [80] to A_l. For pseudospin-1 scattering, because of T-

symmetry preservation, we have $A_0 \gg A_1 = A_{-1}$. For the hybrid system [$\alpha \in (0, 1)$], both A_0 and A_1 can be appreciable.

In the low energy regime, the angular scattering cross section is given by

$$\frac{d\sigma}{d\theta} \approx \frac{2}{\pi k_0} \left| A_0 + A_1 e^{i\theta} \right|^2 . \tag{2.7}$$

For graphene, the coefficients A_0 and A_1 are replaced by $A_{-1/2}$ and $A_{1/2}$ with $A_{-1/2} = A_{1/2}$, leading to vanishing scattering cross section for $\theta = \pi$. For pseudospin-1 scattering, the lowest angular momentum mode is $l = 0$. In this case, scattering is isotropic [96]. For the hybrid system, in the quantum-dot regime, the dominant scattering channels are still $l = 0$ and $l = 1$ but, differing from pseudospin-1/2 or pseudospin-1 scattering, the scattering channels are dependent upon each other. As a result, backscattering will be enhanced even when the value of α is small (but not equal to zero) (see Appendix 2.7 for details).

To characterize enhancement in backscattering, we define the ratio between the values of the angular scattering cross section at angles zero and π:

$$\delta = \frac{d\sigma/d\theta|_{\theta=\pi}}{d\sigma/d\theta|_{\theta=0}}, \tag{2.8}$$

as the enhancement factor, which can be calculated numerically from the scattering matrix. Figure 2.3 shows, for two values of k_0R, δ versus α. In each case, there exists a range of α values in which backscattering is enhanced ($\delta > 1$). Remarkably, in each case there is a peak in δ. For small values of α, the enhancement factor is given by

$$\delta \propto \alpha^4 (k_0 R)^{-2}. \tag{2.9}$$

At the peak, we have $\delta \gg 1$, so backscattering is greatly enhanced. This provides an explanation for the emergence of a quasibound state at the corresponding value of α.

We examine quasibound states in hybrid material systems in more detail. For plane wave scattering, while there is mixing among scattering associated with different angular-momentum channels as characterized by the scattering matrix, at a

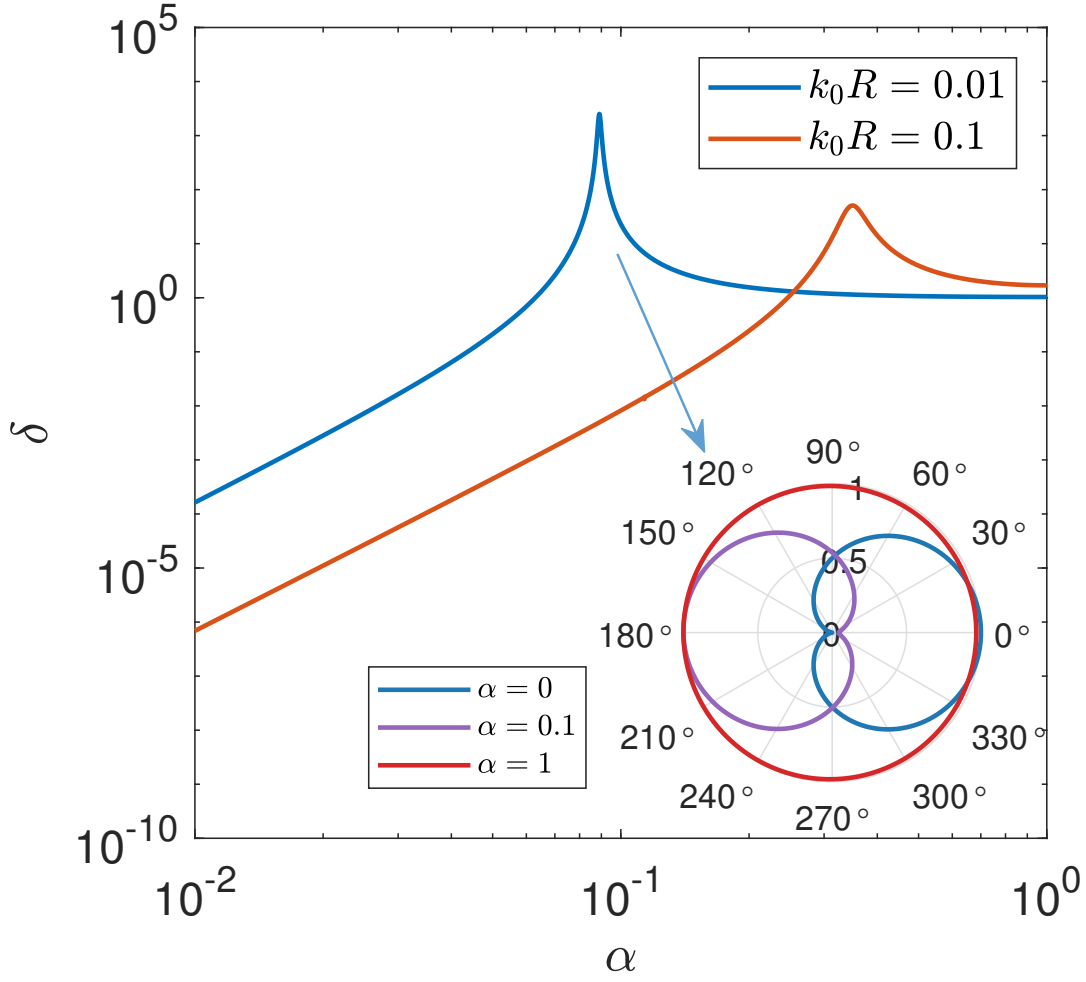

Figure 2.3: Enhancement of backscattering from a circular cavity in the hybrid systems. (a) Enhancement ratio δ versus α for two values of k_0R. In each case, a peak in δ arises, signifying strong enhancement of backscattering. For the case of $k_0R = 0.01$, the peak appears at $\alpha \approx 0.1$ with the enhancement ratio exceeding 10^3. As shown in the inset, a polar representation of the normalized differential cross section for $k_0R = 0.01$ and $VR = 30$, in the low energy regime, there is no backscattering for pseudospin-1/2 and pseudospin-1 scattering is isotropic. However, in certain hybrid systems, strong backscattering can arise with diminishing forward scattering.

resonance scattering depends strongly on the angular momentum [80, 186]. Let B_l be the coefficient of the spinor wavefunction inside the cavity associated with angular momentum l. Each quasibound state corresponds to a maximum in B_l. For scattering in an infinite plane, the lifetime of the quasibound states can be infinite. The confinement quality is thus determined by the number of quasibound states. To estimate this number for a hybrid system, we begin with the degenerate states for pseudospin-1/2 and pseudospin-1 systems as a result of T-symmetry preservation (Appendix 2.7), where the degeneracy is broken for $\alpha \neq 0, 1$. The number of quasibound states will then be larger for the hybrid system than for graphene or the pseudospin-1 system.

For $\alpha \to 0$ and $\alpha \to 1$, the behavior of broken degeneracy can be studied numerically and analytically. To facilitate numerical computations, we fix the incident angle and vary the potential height V. The existence of the confinement states for the first few angular momentum values in the (α, VR) parameter plane are shown in Fig. 2.4. As the value of α is increased from zero, there is fast separation of the originally degenerate states. For $\alpha \to 1$, restoration of the degeneracy occurs in a slower manner. Thus, for small values of α, we expect a more dramatic deviation in the behaviors of the quantum states from those of graphene.

In the quantum-dot regime, the asymptotic behavior of the scattering wavefunction associated with each resonance (or quasibound state) can be obtained analytically, e.g, through the method of level doubling [186, 187] (Appendix 2.7). The energy separation in the two limiting regimes is given by

$$\begin{aligned} &\Delta_l \propto \alpha^2 l/(k_0 R), && \text{for } \alpha \to 0 && \text{and} \quad l \neq 0, \\ &\Delta_l \propto (1-\alpha) k_0 R/l, && \text{for } \alpha \to 1 && \text{and} \quad l \neq 0. \end{aligned} \tag{2.10}$$

For $\alpha \to 0$, the amount of level separation follows a quadratic dependence on α with the quantity $k_0 R \ll 1$ in the denominator. As a result, the amount of separation

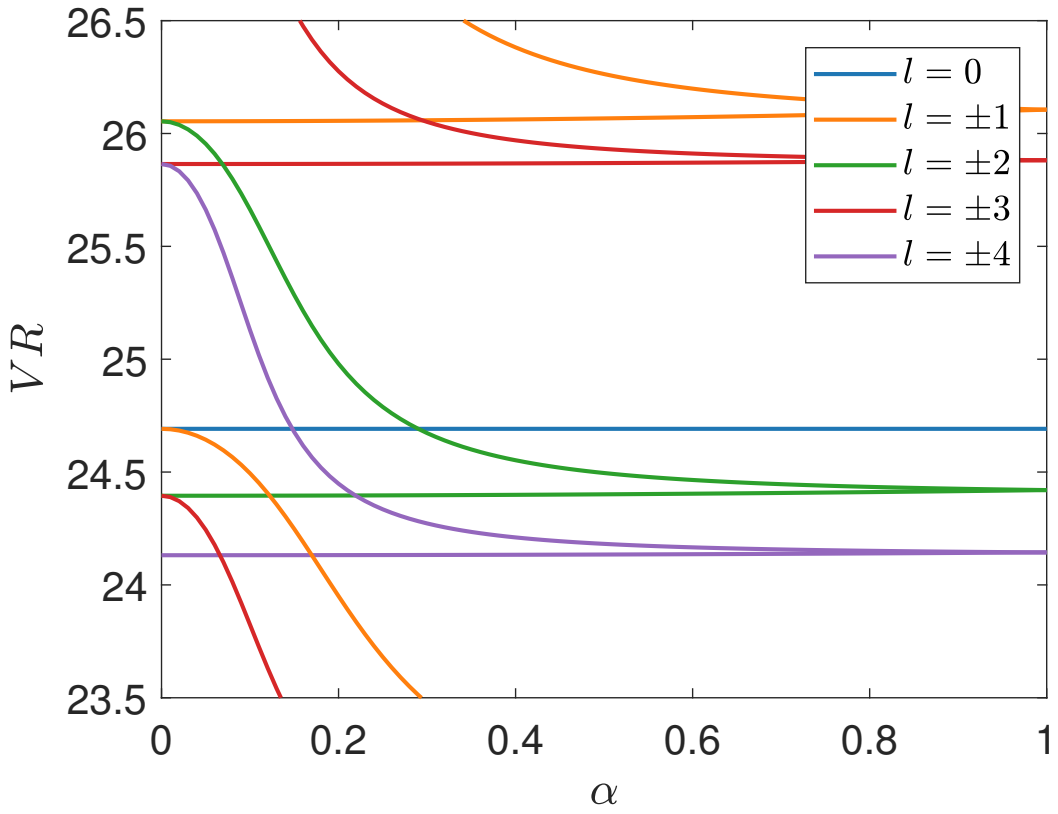

Figure 2.4: Quasibound states associated with the first few angular momentum channels in the parameter plane (α, VR). In the limit $\alpha \to 0$, level degeneracy occurs in the way of $l \to -l+1$. In the opposite limit $\alpha \to 1$, degeneracy follows the rule $l \to -l$. The amount of level separation depends quadratically on α as its value is increased from zero, but the dependence becomes linear as α approaches one. This indicates the emergence of more quasibound states in the small α regime, implying a strong confinement ability there.

grows quickly as α is increased from zero. Near the pseudospin-1 limit, the amount of the separation has a linear dependence on $k_0R(1-\alpha)$ and thus decreases slowly to zero as the value of α approaches one. Numerical support for these analytic estimates is given in Fig. 2.4.

The quick increase in the number of quasibound states as the value of α is increased from zero implies a strong confinement ability of α-T_3 materials near the graphene end.

2.3 Confinement in an Annular Cavity

In the fields of microcavity optics and quantum chaos, annular cavity is a commonly studied type of structures for confinement [188, 189, 123] because it offers a convenient and systematic way to generate the whole spectrum of classical behaviors ranging from integrable dynamics to chaos. In particular, the system is integrable when the two circles are concentric (ring cavity). Mixed classical dynamics with coexistence of Kolmogorov–Arnold–Moser (KAM) islands and chaos arise when there is a small displacement between the centers of the two circles. Fully developed chaotic dynamics occur when the displacement is sufficiently large. An appealing feature of the annular structure is that the quantum scattering matrix can be analytically calculated for all cases of classical dynamics through a proper coordinate transform [123].

To generate an annular structure on an α-T_3 sheet, we apply the following electrical potential:

$$V(\mathbf{r}) = V_1\Theta(R_1 - r)\Theta(|\mathbf{r} - \boldsymbol{\xi}|) + V_2\Theta(R_2 - |\mathbf{r} - \boldsymbol{\xi}|), \tag{2.11}$$

where the two circles have radius R_1 and R_2, respectively, and their centers are located at O and O' with the displacement vector $\boldsymbol{\xi} = O' - O$ between them. A simple change of coordinates gives $\boldsymbol{\xi} = (\xi, 0)$, which aligns the displacement in the x direction. Figures 2.5(a1-c1) illustrate three representative annular profiles corresponding to three different values of ξ. For the integrable dynamics case of $\xi = 0$, due to the perfect circular symmetry in the potential, there is no mixing among the states with different angular momenta. (Analytic formulas for the scattering matrix and cross sections for this case are given in Appendix 2.8.) For $\xi \neq 0$, the circular symmetry is broken and the classical dynamics contain a chaotic component. For example, for the potential profile in Fig. 2.5(b1), we have checked that the classical phase space contains both KAM tori and chaotic regions, but for a a larger value of ξ [e.g., the

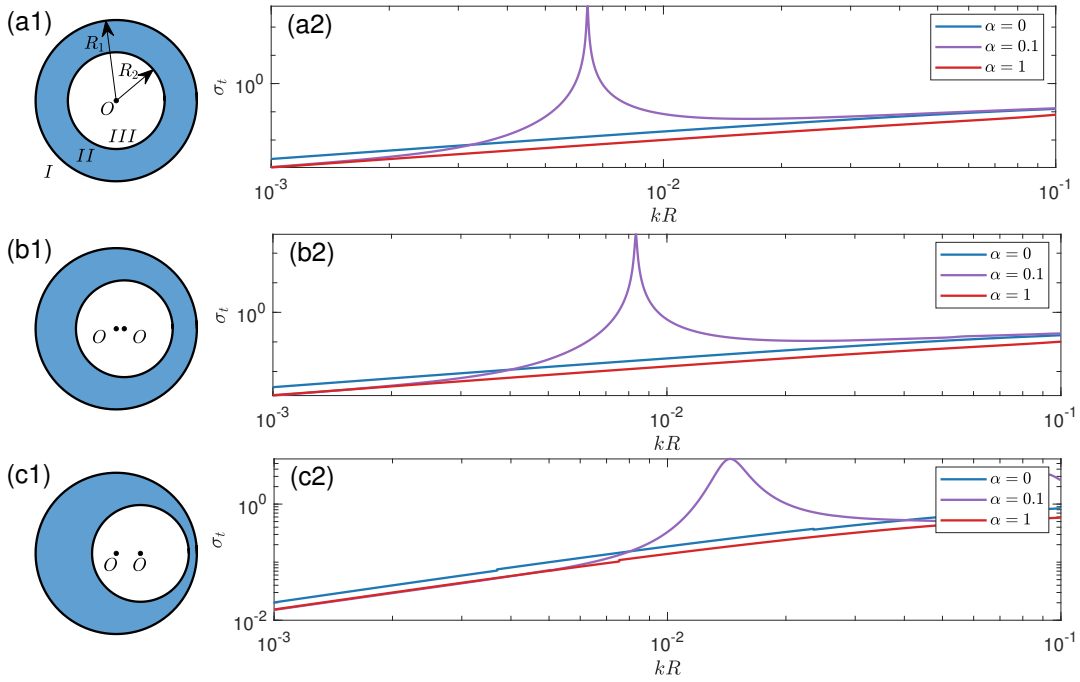

Figure 2.5: Three types of annular cavities and the behaviors of the corresponding total scattering cross section. (a1) A ring cavity ($\xi = 0$), where the two boundaries divide the lattice into three regions and the classical dynamics are integrable. (a2) The corresponding total cross section σ versus the dimensionless quantity $k_0 R_1$ for different values of α for the parameter setting $R_2/R_1 = 0.6$ and $\xi = 0$. (b1) Confinement geometry for $\xi R_1 = 0.1$, where O and O' are the centers of the outer and inner circles, respectively. In this case, the classical dynamics are mixed. (b2) The corresponding total scattering cross section. (c1) Potential profile for $\xi R_1 = 0.3$, where the classical dynamics are fully chaotic and (c2) the corresponding scattering cross section. For all three types of cavities, there is a resonant peak in the total cross section for $\alpha = 0.1$, but no such peak appears for $\alpha = 0$ or $\alpha = 1$, implying a much stronger ability to confine electrons for the α-T_3 ($\alpha = 0.1$) cavity than for the graphene or pseudospin-1 cavity. Fully developed classical chaos can smooth out the resonance to some extent, but it is still pronounced.

case of Fig. 2.5(c1)], all KAM tori have been destroyed, leading to fully developed chaos [188]. (Appendix 2.9 gives the analytic formulas of some key quantities characterizing relativistic quantum chaotic scattering in this case.)

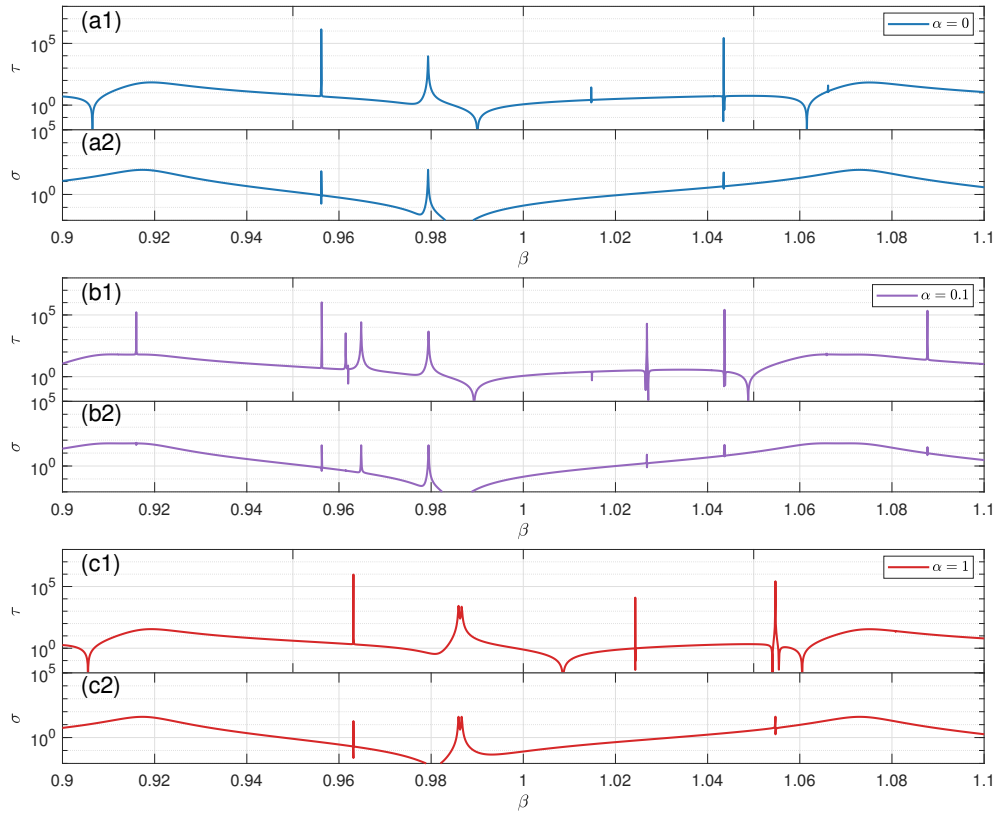

Figure 2.6: Comparison of confinement abilities among three material systems in an eccentric annular cavity. Shown are the total cross section and the Wigner-Smith time delay versus the normalized potential height β for a fixed incident energy. The parameter setting is $V_1R_1 = -10\beta$ and $V_2R_1 = 40\beta$, $k_0R_1 = 0.1$, and $\xi R_1 = 0.05$ (a1-c1) Total scattering cross section and (a2-c2) Wigner-Smith time delay versus β for $\alpha = 0$, $\alpha = 0.1$, and $\alpha = 1$, respectively. Among the three cases, the hybrid system ($\alpha = 0.1$) exhibits more resonant peaks than the other two cases, indicating a stronger confinement ability.

For convenience, we choose the radius R_1 of the outer circle as the characteristic length of the system. Figures 2.5(a2-c2) show the total scattering cross section versus the quantity k_0R_1 for the three cases in Figs. 2.5(a1-c1), respectively, where each panel contains results for three values of α: 0, 0.1, and 1. In all three cases of the potential profile, in the range of k_0R_1 values plotted, for $\alpha = 0$ or $\alpha = 1$, there is no

resonance. However, for $\alpha = 0.1$, there is a strong resonant peak about $k_0 R_1 = 10^{-2}$ for all three cases. The resonance is relatively sharp when the classical dynamics are integrable and mixed, as shown in Figs. 2.5(a2) and 2.5(b2), respectively. Fully developed chaos can smooth out the resonance to certain extent, but it is still quite pronounced, as shown in Fig. 2.5(c2).

To characterize the confinement quality for different cavity geometries and for different values of α, we fix the incident energy and increase the normalized potential height β from zero, which is defined through $V_1 R_1 = -10\beta$ and $V_2 R_1 = 40\beta$. Figure 2.6 shows, for $\xi R_1 = 0.05$ (mixed classical dynamics), the total cross section and the Wigner-Smith time delay versus β in the range $\beta \in [0.9, 1, 1]$ for three values of α. In all three cases, there are a number of resonant peaks, each corresponding to a quasibound state, where the inverse of the width of a resonant peak determines the lifetime of the state. Simply counting the number of resonant peaks in all three cases, we find that the hybrid material system has more as compared with either the graphene or the pseudospin-1 material. The reason, as in the case of a circular cavity, is that, for $\alpha \neq 0, 1$, the broken T-symmetry makes the originally degenerate states (for $\alpha = 0$ or 1) nondegenerate.

To further characterize the confinement, we integrate the curves of the Wigner-Smith time delay in Figs. 2.6(a2-c2) - an approach often employed to quantify confinement in optical cavities [190, 191, 185]:

$$\overline{\tau} = \frac{1}{\beta_2 - \beta_1} \int_{\beta_1}^{\beta_2} \tau(k_0, \beta) d\beta. \tag{2.12}$$

Roughly, the so obtained average time delay corresponds to the number of quasibound states contained in the integration interval. To be systematic, we calculate the integral in the parameter plane $(\alpha, \xi R_1)$, as shown in the color-coded graph in Fig. 2.7(a). We observe a region of relatively higher values of $\overline{\tau}$ (brighter color) for $0 < \alpha <$

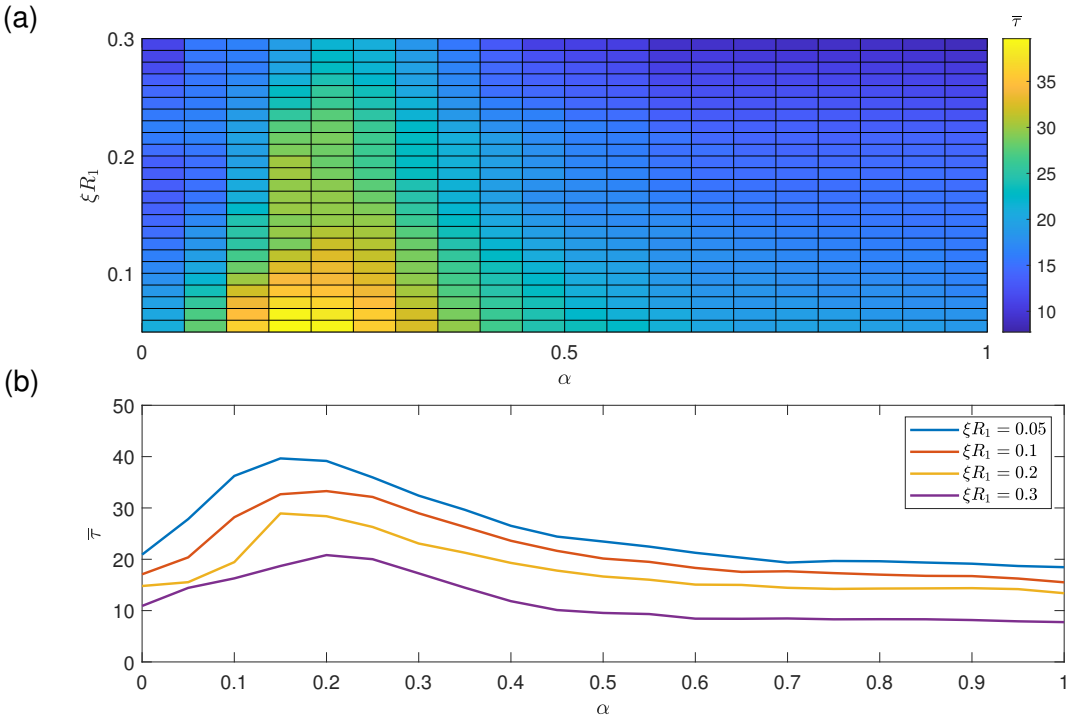

Figure 2.7: Confinement for different α-T_3 materials and different annular cavity geometries. (a) Color-coded value of the integrated Winger-Smith time delay $\overline{\tau}$ in the parameter plane $(\alpha, \xi R_1)$. For α in the vicinity of 0.2 and small values of ξ, $\overline{\tau}$ attains relatively large values. Computationally, the 2D parameter plane is represented by a grid. For each grid point, the integration is carried out with respect to β for $\beta \in [0.4, 1.6]$. Varying the integration interval does not change the result significantly. (b) The value of $\overline{\tau}$ versus α for four values of the geometric parameter ξR_1. Regardless of the geometry, best confinement occurs for α-T_3 materials with the value of α in the vicinity of 0.2.

0.5 and $\xi R_1 \leq 0.15$, indicating that a stronger confinement can be achieved for the corresponding lattice structure and cavity geometry. For a fixed geometry, the confinement ability exhibits a non-monotonous behavior as the value of α is increased from zero to one, as shown in Fig. 2.7(b) for four different annular cavities.

2.4 Effect of Geometric Deformations on Confinement

2.4.1 Confinement in Deformed Cavities with Distinct Classical Dynamics: Numerical Demonstration

We have demonstrated that the family of annular cavities, regardless of the corresponding classical dynamics, is able to confine α-T_3 particles. Can confinement be achieved in more general cavities with their geometrical shape deformed from the circular shape, such as the elliptical cavity with mixed classical dynamics or the stadium-shaped chaotic cavity (a paradigm in the field of quantum chaos [2, 11, 12])? (There is quantum chaotic scattering [192, 193, 194, 195, 196, 23] in this case.) For pseudospin-1/2 particles, previous studies based on the method of finite-domain scattering (by setting zero the wavevector outside of the cavity) revealed that confinement modes can exist in the stadium cavity [197, 198, 199, 200]. Geometrically, the elliptical and stadium cavities can be generated through continuous deformation of the circular cavity. To quantitatively assess the effect of geometric deformation on confinement, we carry out a comparative analysis of the confinement quality of α-T_3 particles in the circular, elliptical, and stadium cavities. During this section, chaos means the dynamics for classical billiard system, free particles moving inside a billiard with special geometry and spin degree is exclude. However, for quantum calculation of α-T_3 particles, we will naturally take spin into account.

While the scattering of α-T_3 particles in the circular and annular cavities can be solved analytically, for a deformed cavity without the circular symmetry, analytic solutions are not feasible. In fact, to our knowledge, there were no previous numerical methods for solving the generalized Dirac-Weyl equation for α-T_3 particles that are neither pseudospin-1/2 nor pseudospin-1. Taking advantage of a recently developed computational method [201] for pseudospin-1 particles based on the multiple

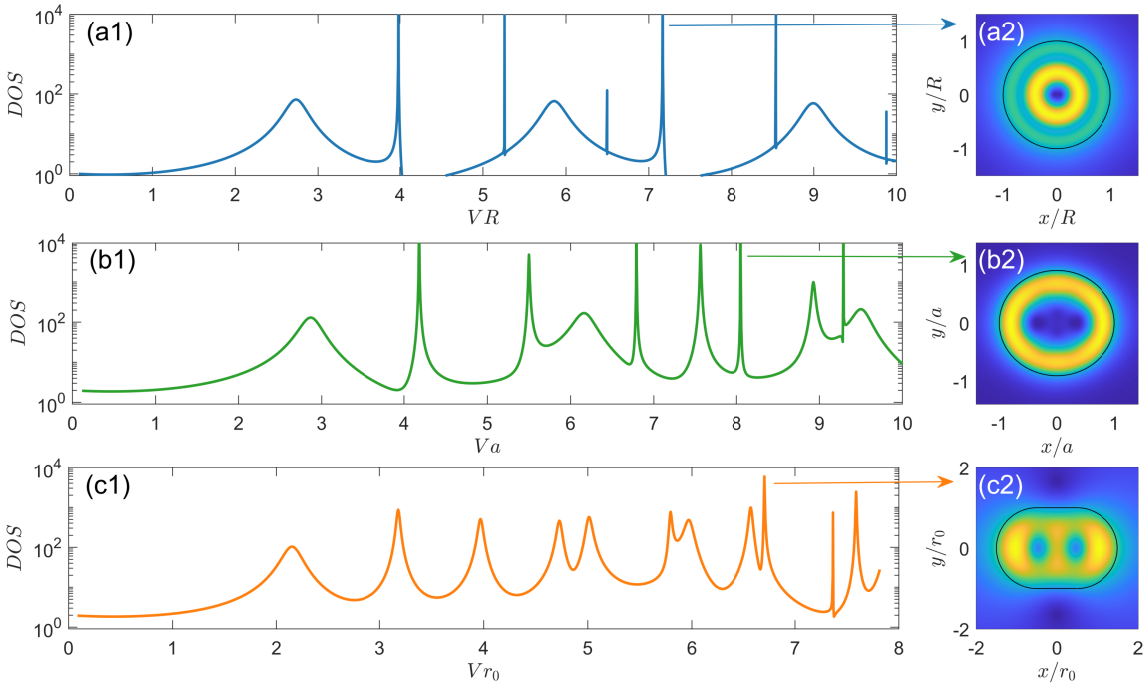

Figure 2.8: Resonances and DOS patterns calculated from the MMP method for three types of cavities for pseudospin-1/2 wave. The top, middle, and bottom rows correspond to the circular, elliptical, and stadium-shaped cavities with classically integrable, mixed, and chaotic dynamics, respectively. For a meaningful comparison, the areas of the cavities are set equal: $\pi ab = \pi R^2$ for the ellipse, where a and b are respectively the semimajor and semiminor axes (eccentricity $\gamma = a/b$), and $\pi r_0^2 + 2l_0 r_0 = \pi R^2$ for the stadium, where l_0 is the length of the straight segment and r_0 is the radius of the semicircle. In the left column, the normalized energy parameter is VR for the circular cavity, Va for the elliptical cavity, and Vr_0 for the stadium. (a1-c1) The DOS at an arbitrary point in the cavity versus the energy parameter for the three types of cavities, respectively. In (a1), the incident wavevector is $k_0 R = 1/10$ and the nearly periodic resonant peaks correspond to different angular momentum states. (a2) A representative DOS pattern associated with a resonant peak, where the square root of the DOS is color-coded. In (b1), the eccentricity of the elliptical cavity is $\gamma = 1.1$ and the incident wavevector is $k_0 a = \sqrt{1.1}/10$. (b2) Color-coded square root of the DOS pattern for a resonant state. In (c1), the stadium cavity has $A = \pi R^2$, $r_0/l_0 = 1$, and $L = 2\pi r_0 + 2l_0$, and the incident wavevector is $k_0 r_0 = \sqrt{\pi/(\pi+2)}/10$. (c2) Color-coded square root of the DOS pattern for a representative resonant state.

multipole (MMP) method in optics [202, 203, 204, 205, 206], we have developed an efficient computational method [201] to solve the spinor wavefunctions associated with the scattering of α-T_3 particles from an *arbitrary* geometric domain (Appendix 2.10). The basic idea is to place two sets of fictitious "poles," one inside the cavity and

another outside, which are regarded as the sources for generating the scattering wavefunction. The multiple set of poles (henceforth the term "multiple multipole") can be determined by matching the wavefunctions at the cavity boundary. From the so-calculated scattering wavefunctions, we get the density of states (DOS) as a function of some energy-related parameter of the system, where each peak in the DOS corresponds to a quasibound state. (The details of the MMP method for quantum scattering associated with the generalized Dirac-Weyl equation for α-T_3 particles are given in Appendix 2.10.)

Figure 2.8 illustrates, for $\alpha = 0$ and three types of cavities (circular, elliptical, and stadium-shaped), DOS versus some normalized energy parameter and the representative DOS patterns in the physical domain corresponding to a pronounced resonant state. The three cavities have the same area. Shown in the left column is the average DOS per unit area inside the cavity versus an energy parameter for the three cavities. In each case, in the interval of the energy parameter, there are a number of distinct peaks in the DOS plot, each corresponding to a quasibound state. The right column shows, for each case, the DOS pattern associated with a typical pronounced resonant peak. Note that, for $\alpha = 0$, both components of the spinor wavefunction are continuous across the cavity boundary. These results, which are obtained from our MMP method, agree with the previous results for scattering of pseudospin-1/2 particles in graphene [197, 198, 199, 200].

2.4.2 Recurrence of Period-2 Type of Quasibound Modes

In quantum confinement, a fundamental issue is recurrence where, as the wavevector varies, a particular resonant state can arise periodically at a distinct set of energy or wavevector values. In bounded systems, e.g., the stadium cavity with classical chaotic dynamics that has played a paradigmatic role in the study of the phenomenon

of scarring in nonrelativistic quantum mechanics [8], scars associated with certain periodic orbits can recur, which can be described by the Gutzwiller formula [2, 3]. Recurrence of quantum scars can also arise in graphene billiard systems [207]. In relativistic quantum billiard systems with T-symmetry breaking [6], chiral scars of massless spin-1/2 fermions for certain classes of periodic orbits can arise [135, 208, 209], which can recur with the energy or the wavevector. In open (scattering) systems with quasibound states, there are still relationships among the classical period orbits, the wavefunctions, and directional emission in nonrelativistic quantum systems, but the current understanding is that recurrence of the quasibound states is unlikely [210, 211, 129]. Can this conventional wisdom be applied to α-T_3 particles in a cavity?

Our answer to the above question is surprisingly a "no," as we have found a class of "period-2" for α-T_3 particles that can actually recur. In general, quantum scars are referred to as the unusually high concentrations of the wavefunction along the classical periodic orbits [8]. Period-2 means particle hit with one boundary and bounce back. Especially, for the elliptical cavity, we find that these modes correspond to the spinor wavefunctions concentrated along the minor axis of the cavity, as shown in Figs. 2.9(a) and 2.9(b) for two cavities with different values of eccentricity, where the insets illustrate the corresponding classical orbits. For convenience, we fix the incident energy E and vary the height V of the electric potential so as to systematically increase the wavevector q inside the cavity. We find that, regardless of the geometric shape of the cavity and of the value of α, the period-2 mode can repeatedly occur in a periodic fashion. Figure 2.9(a) shows, for eccentricity value $\gamma = 1.1$, $\alpha = 1/3$ (red asterisks) and $\alpha = 1$ (purple asterisks), the wavevector value at which the mode emerges versus the index of the recurring mode. A similar plot is displayed in Fig. 2.9(b) for $\gamma = 1.5$, $\alpha = 2/3$ (red asterisks) and $\alpha = 1$ (purple asterisks). The

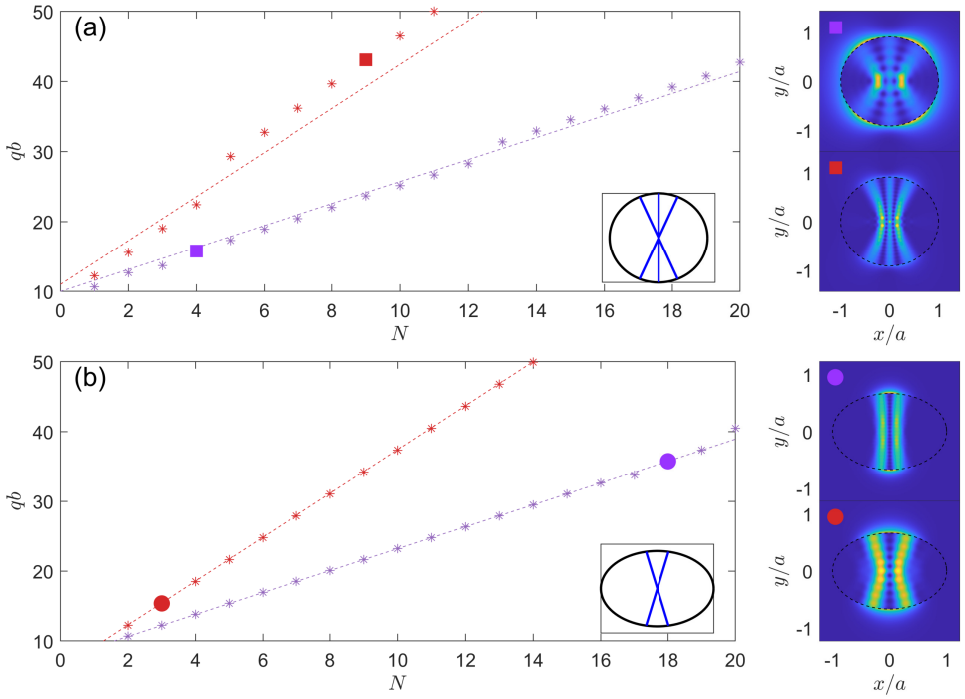

Figure 2.9: Recurrence of a class of quasibound state in elliptical cavities. The quasibound state concentrates near the minor axis of the cavity, corresponding to a period-2 type of classical orbits. (a) Dimensionless wavevector value qb at which such a quasibound state emerges versus the index of the resonant energy level for $\gamma = 1.1$, where two types of α-T_3 waves are shown: $\alpha = 1/3$ (purple asterisks) and $\alpha = 1$ (red asterisks). Insert: one type of period-2 orbit, with wavefunction localized along the blue line.(b) Similar plots but for $\gamma = 1.5$, $\alpha = 2/3$ (purple asterisks) and $\alpha = 1$ (red asterisks). The right panels show representative spinor wavefunction patterns for some specific values of q as indicated in the panels on the left.

spinor wavefunction patterns associated with some representative recurring modes are shown on the right-hand side of Figs. 2.9(a) and 2.9(b). We find that the spacing in the wavevector for recurrence follows certain rules:

$$\Delta q = 2\pi/L, \quad \text{for } \alpha = 1, \tag{2.13}$$

$$\Delta q = \pi/L, \quad \text{for } \alpha \in (0, 1), \tag{2.14}$$

where $L = 2b$ is the length of the minor axis of the elliptical cavity. These resemble the semiclassical rules for quantum scars in closed chaotic systems [2, 3]. For $\alpha \neq 1$ in which the T-symmetry is broken, the frequency in the wavevector for the quasibound mode to occur is twice of that for pseudospin-1 system ($\alpha = 1$). Beyond

T-symmetry breaking, we notice that such phenomena happens regardless of period-2 orbits. Similar phenomena arise in Schrödinger equation.

We remark that, in Fig. 2.8, a few relatively low excited states are presented. Some relatively high states are shown in Fig. 2.9. In particular, in Fig. 2.9(a), we have $qb \approx 40$, where b is the dimension of the cavity and q is the wavevector inside. The setting puts the system in the quantum dot regime, as the wavevector inside the cavity is large. As shown in the right panels of Fig. 2.9, there are many nodes inside the cavity and the period-2 mode corresponds to a highly exited state. We see that, because of the small wavelength, deformation and classical chaos can have an effect on the states, driving the wavefunction towards concentrating along some classical periodic orbits.

2.4.3 Characterization of Confinement by Magnetic Moment

For the circular and annular cavities, we have used the Wigner-Smith delay time and the total scattering cross section to quantify confinement, which can be calculated from the analytic scattering matrix. For a mixed or chaotic cavity, while the scattering wavefunctions for α-T_3 particles can be numerically calculated using the MMP method, it does not yield the scattering matrix. For a given state, the degree of confinement is directly related to the peak width of the DOS distribution. However, counting peak numbers and estimating their width are not reliable, especially when the distributions overlap. We thus seek alternative measures to characterize confinement.

One experimentally accessible measure is the magnetic moment [212, 213, 214] defined as

$$\mu_B = -\frac{e}{2}\int (\mathbf{r}\times\mathbf{j})d^2\mathbf{r}, \tag{2.15}$$

where $\mathbf{r}$ is the position, $\mathbf{j}$ is the current, and the integration is with respect to the

interior of the cavity. In Ref. [213], it was shown that in topological insulators with impurities, only the states with a large magnetic moment can survive. The robustness of the states against impurities can be measured by the magnetic moment, justifying its use to characterize the degree of confinement.

For the α-T_3 wave, the associated current is given by $\mathbf{j} = v_g \Psi^\dagger \mathbf{S}^\alpha \Psi$ with the following components in the 2D Cartesian coordinates:

$$\begin{aligned} j_x &= 2v_g \mathrm{Re}\left[\psi_B^*(\psi_A \cos\phi + \psi_C \sin\phi)\right], \\ j_y &= -2v_g \mathrm{Im}\left[\psi_B^*(\psi_A \cos\phi - \psi_C \sin\phi)\right]. \end{aligned} \tag{2.16}$$

We compare the confinement properties for three types of cavities with distinct classical dynamics: circular (integrable), elliptical (mixed), and stadium (fully chaotic) cavities. The size parameters for these cavities are the same as those in Fig. 2.8. We fix the incident wavevector $k_0 R = 0.1$, change the potential height V, and calculate the magnetic moment for each quasi-bound state. Since the current is proportional to the square of the wavefunction, for each state we normalize the magnetic moment by $|\Psi|^2$ to get the average:

$$\overline{\mu}_B = \frac{1}{N} \sum_{i=1}^{N} \frac{|\int \mathbf{r} \times \mathbf{j} d^2\mathbf{r}|}{\int |\Psi|^2 d^2\mathbf{r}}, \tag{2.17}$$

where the sum is over all quasibound states in a wavevector range. Figure 2.10 shows, for the three types of cavities, the average magnetic moment versus α. Note that, because of the unbroken T-symmetry for $\alpha = 0, 1$, the net current is zero and so the corresponding average magnetic moment is zero.

For $\alpha < 0.1$, confinement is strengthened as the value of α is increased. For $\alpha > 0.2$, the opposite trend occurs: confinement is weakened as α becomes larger. For all three cavities, the maximum value of $\overline{\mu}_B$ is achieved for $\alpha \approx 0.1$, indicating the strongest possible confinement there among the possible α-T_3 materials.

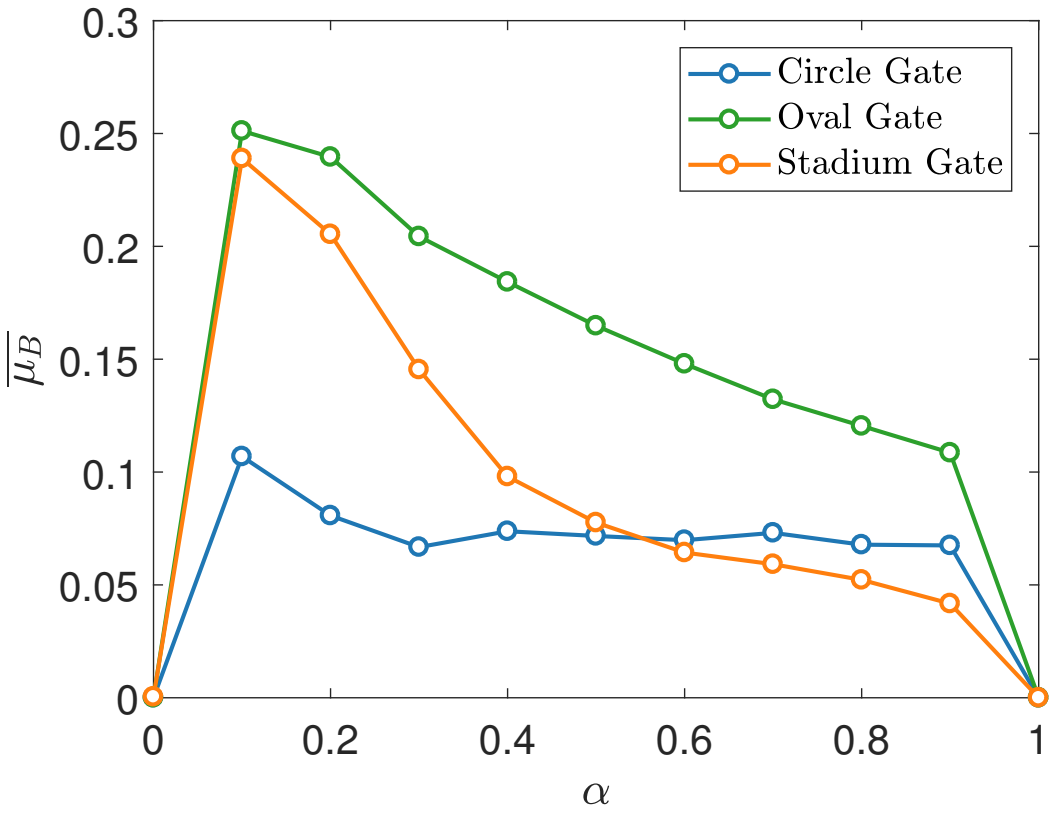

Figure 2.10: Average magnetic moments $\overline{\mu}_B$ versus α for three different cavities with characteristically distinct classical dynamics. The incident wavevector is set to be $k_0 R = 0.1$ for the circular cavity. The areas of the other two cavities are the same as that of the circular cavity, and the corresponding wavevectors are set as in Fig. 2.8. The elliptical cavity has eccentricity $\gamma = 1.1$ and the parameters of the stadium cavity are $r_0/l_0 = 1$ and $L = 2\pi r_0 + 2l_0$. The normalized magnetic moment is averaged for $VR \in [20, 40]$, in which there are hundreds of quasibound states for all three cavities. The average magnetic moment is zero for $\alpha = 0$ or $\alpha = 1$ because of T-symmetry preservation, and is maximized for a small but non-zero value of α. This indicates that, among the possible α-T_3 materials, one with the strongest possible confinement occurs in the regime of small (but non-zero) α values.

Certain features of Fig. 2.10 can be heuristically understood, as follows. In general, wavefunctions concentrating along the boundary will contribute to a large magnetic moment and hence, strong confinement. For small values of α, the third spinor component ψ_C is proportional to αh with h being a kind of Bessel function. However, the boundary condition stipulates $h \propto \alpha^{-2}$. As a result, the third component along

the boundary is proportional to α^{-1}, leading to a larger magnetic moment for small but non-zero values of α. Note that, in the original definition Eq. (2.15), the magnetic moment depends on both the magnitude of the current $\mathbf{j}$ and the position $\mathbf{r}$. While the T-symmetry is broken for all three cavities for $\alpha \leq 0$, for the elliptical and stadium cavities, there is one additional broken symmetry: the circular symmetry. As a result, the average magnetic moments are larger than that for the circular cavity.

2.5 Conclusion and Discussion

Given a specific type of quantum materials, the issue of confinement of quasiparticles is of both fundamental importance and applied value. We have addressed the confinement issue in α-T_3 materials that represent a broad spectrum of state-of-the-art 2D Dirac materials. In terms of the lattice interaction parameter α whose value lies in the unit interval, graphene with pseudospin-1/2 quasiparticles is at the lower end of the spectrum ($\alpha = 0$) while at the other end sits the pseudospin-1 material ($\alpha = 1$). Exactly at the two ends of the spectrum, the T-symmetry is preserved but for any material in between, the symmetry is broken. While experimental realizations of α-T_3 materials have been achieved only at the two ends of the spectrum, advances in nanotechnologies and materials science may make it possible to create the 2D Dirac-Weyl materials with arbitrary values of α in the near future. A pertinent theoretical question is thus what kind of α-T_3 materials would have the best confinement property, especially in the quantum-dot regime. In general, associated with T-symmetry breaking is broken level degeneracy, facilitating confinement. It is thus intuitively expected that any α-T_3 material for $\alpha \neq 0$ would have better confinement properties than graphene and the pseudospin-1 lattice. We have indeed observed this feature. In fact, we have found that optimal confinement is achieved for materials near the graphene end, i.e., those with small and non-zero α values.

The general approach we have undertaken in this study is quantum scattering from a cavity generated by a purely static electric potential. Focusing on the quantum-dot regime, we have studied a number of cavities with characteristically distinct classical dynamics: circular, annular, elliptical, and stadium cavities. For the circular cavity with classical integrable dynamics, the scattering matrix can be analytically obtained, making feasible a confinement analysis based on the scattering cross sections and the Wigner-Smith time delay. Typically, a number of resonant peaks arise in the plots of these quantities versus some energy-related parameter and a relatively sharp resonant peak corresponds to strong confinement, as has been verified by an analysis of backscattering and level separation. For the eccentric annular cavity that can generate the full spectrum of classical dynamics ranging from integrable to chaotic, the scattering matrix can still be analytically calculated through some proper coordinate transform. For the elliptical and stadium cavities, where the classical dynamics are mixed for the former and fully chaotic for the latter, the scattering matrix approach is infeasible. We have adopted an efficient numerical method, the MMP method, to deal with these cavities. In particular, the method enables the three-component spinor wavefunction for an arbitrary value of α to be numerically computed, based on which experimentally accessible measures of confinement such as the magnetic moment can be calculated. Through analyzing the magnetic moment, we find that, for small α values, for the confined modes, the third component of the spinor wavefunction is typically larger than the other two components, generating wavefunctions concentrating along the cavity boundary with a large magnetic moment.

For all these cases, we have found that in the region of small α values ($\alpha < 0.1$), confinement is enhanced as the value of α is increased. For $\alpha > 0.2$, confinement is weakened with a continuous increase in the value of α. Thus, for the α-T_3 lattice system, the best confinement is achieved for a small but non-zero value of α. This

phenomenon is general and holds regardless of the nature of the classical dynamics.

2.6 Potential Scattering of α-T_3 Particles: Basics

Free space solutions. In the free space, the Hamiltonian for α-T_3 wave is

$$H = v_g \boldsymbol{S}^\alpha \cdot \boldsymbol{p}, \tag{2.18}$$

where $\mathbf{S}^\alpha = (S_x^\alpha, S_y^\alpha)$ are 3×3 generalized Pauli matrices. The two matrices, together with the third one S_z^α, obey the Levi-Civita symbol in three dimensions. For massless quasiparticles in α-T_3 materials, the matrix S_z^α does not arise in the Hamiltonian. Letting $\phi = \tan^{-1}\alpha$, we have

$$S_x^\alpha = \begin{pmatrix} 0 & \cos\phi & 0 \\ \cos\phi & 0 & \sin\phi \\ 0 & \sin\phi & 0 \end{pmatrix}, \tag{2.19}$$

$$S_y^\alpha = -i \begin{pmatrix} 0 & \cos\phi & 0 \\ -\cos\phi & 0 & \sin\phi \\ 0 & -\sin\phi & 0 \end{pmatrix}. \tag{2.20}$$

For $\phi = 0$, S_x^α and S_y^α reduce to the 2×2 Pauli matrices for pseudospin-1/2 particles. For $\phi = \pi/4$, the Hamiltonian reduces to that for pseudospin-1 particles. For $\alpha \leq 0$, we have $\sin\phi \approx \alpha$.

In the polar coordinates, the eigenequation becomes

$$\begin{aligned} \begin{pmatrix} 0 & \cos\phi\hat{L}_- & 0 \\ \cos\phi\hat{L}_+ & 0 & \sin\phi\hat{L}_- \\ 0 & \sin\phi\hat{L}_+ & 0 \end{pmatrix} \begin{pmatrix} \cos\phi h_{l-1} e^{-i\theta} \\ i\kappa h_l \\ -\sin\phi h_{l+1} e^{i\theta} \end{pmatrix} e^{il\theta} \\ = E \begin{pmatrix} \cos\phi h_{l-1} e^{-i\theta} \\ i\kappa h_l \\ -\sin\phi h_{l+1} e^{i\theta} \end{pmatrix} e^{il\theta}, \end{aligned} \tag{2.21}$$

where

$$\hat{L}_{\pm} = -ie^{\pm i\theta}\left(\partial_r \pm i\frac{\partial_\theta}{r}\right). \tag{2.22}$$

and $\kappa = \text{sign}(k)$ is the band index.

Comparing with the solutions for the pseudospin-1/2 system, we have that h_l are the Bessel type of functions. Let $h_l^{(0)} = J_l$ be the Bessel function, and $h_l^{(1)} = H_l^{(1)}$ and $h_l^{(2)} = H_l^{(2)}$ be the first and second kind of Hankel functions. Furthermore, we let

$${}^k\psi_l^{(0,1,2)}(\mathbf{r}) \equiv \frac{1}{\sqrt{2\pi}}\begin{pmatrix} \cos\phi h_{l-1}^{(0,1,2)}(kr)e^{-i\theta} \\ i\kappa h_l^{(0,1,2)}(kr) \\ -\sin\phi h_{l+1}^{(0,1,2)}(kr)e^{i\theta} \end{pmatrix} e^{il\theta}, \tag{2.23}$$

where $\mathbf{r} = (r, \theta)$.

Boundary conditions. For a finite potential, we write the three-component spinor wavefunction as

$$\Psi(r,\theta) = \begin{pmatrix} \psi_A \\ \psi_B \\ \psi_C \end{pmatrix} = \begin{pmatrix} \mathcal{R}_A(r)e^{-i\theta} \\ \mathcal{R}_B(r) \\ \mathcal{R}_C(r)e^{i\theta} \end{pmatrix} e^{il\theta}, \tag{2.24}$$

where $\mathcal{R}_A(r)$, $\mathcal{R}_B(r)$, and $\mathcal{R}_C(r)$ are the corresponding radial components. The eigenequation $H\Psi = E\Psi$ thus becomes

$$\begin{aligned} -iv_g \begin{pmatrix} 0 & \cos\phi\left(\frac{d}{dr} + \frac{l}{r}\right) & 0 \\ \cos\phi\left(\frac{d}{dr} - \frac{l-1}{r}\right) & 0 & \sin\phi\left(\frac{d}{dr} + \frac{l+1}{r}\right) \\ 0 & \sin\phi\left(\frac{d}{dr} - \frac{l}{r}\right) & 0 \end{pmatrix} \begin{pmatrix} \mathcal{R}_A(r) \\ \mathcal{R}_B(r) \\ \mathcal{R}_C(r) \end{pmatrix} \\ = [E - V(r)] \begin{pmatrix} \mathcal{R}_A(r) \\ \mathcal{R}_B(r) \\ \mathcal{R}_C(r) \end{pmatrix}. \end{aligned} \tag{2.25}$$

Suppose the change in the potential at the cavity boundary $r = R$ is finite. Integrating Eq. (2.25) in the infinitesimal interval $r \in [R - \eta, R + \eta]$, where $\eta \sim 0$, we obtain

$$\begin{aligned}
\mathcal{R}_B(R-\eta) &= \mathcal{R}_B(R+\eta), \\
\cos\phi\mathcal{R}_A(R-\eta)+\sin\phi\mathcal{R}_C(R-\eta) &= \sin\phi\mathcal{R}_A(R+\eta)+\cos\phi\mathcal{R}_C(R+\eta)
\end{aligned} \tag{2.26}$$

Together with the angular part of the spinor wavefunction, we get the boundary conditions as

$$\begin{aligned}
\psi_B^< &= \psi_B^>, \\
\cos\phi\psi_A^< e^{i\theta}+\sin\phi\psi_C^< e^{-i\theta} &= \cos\phi\psi_A^> e^{i\theta}+\sin\phi\psi_C^> e^{-i\theta}.
\end{aligned} \tag{2.27}$$

where $<$ and $>$ denote the wavefunctions inside and outside of the boundary, respectively. Note that, across the boundary, the second spinor component must be continuous, but only a linear combination of the first and third components is required to be continuous.

The current density for α-T_3 particles is

$$\boldsymbol{j} = v_g \Psi^\dagger \boldsymbol{S}^\alpha \Psi, \tag{2.28}$$

whose components are given by

$$\begin{aligned}
j_x &= 2v_g \text{ Re } \left[\psi_B^*(\psi_A\cos\phi+\psi_C\sin\phi)\right], \\
j_y &= -2v_g \text{ Im } \left[\psi_B^*(\psi_A\cos\phi-\psi_C\sin\phi)\right].
\end{aligned} \tag{2.29}$$

Elastic scattering theory. For scattering of α-T_3 particles from a potential, the general wavefunction can be written as

$$\begin{aligned}
\Psi &= \Psi_{in}+\Psi_{out} \\
&= \sum_l a_l \left[{}^k\psi_l^{(2)} + \sum_{l'} S_{ll'}\,{}^k\psi^{(1)}\right] \\
&= \sum_l a_l \left[2\,{}^k\psi_l^{(0)} + \sum_{l'}(S_{ll'}-\delta_{ll'})\,{}^k\psi^{(1)}\right],
\end{aligned} \tag{2.30}$$

where the incident and reflected waves are expressed as the second and first kinds of Hankel functions, respectively, $S_{ll'}$ is the scattering matrix between incident wave with angular momentum l and scattering wave with angular momentum l'. Let $T_{ll'} \equiv S_{ll'} - \delta_{ll'}$ and $\chi_{in} \equiv \sum_l a_l {}^k\psi_l^{(0)}$. For certain choice of a_l, χ_{in} corresponds to the plane wave. The far-field behavior of the wavefunction is

$$\lim_{k_0 r \gg 1} \Psi = \chi_{\text{in}} + \frac{f(\theta, \theta')}{\sqrt{-ir}} \begin{pmatrix} \cos\phi e^{-i\theta} \\ \kappa \\ \sin\phi e^{i\theta} \end{pmatrix} e^{ikr}. \tag{2.31}$$

Comparing with the assumption in Eq. (2.30), we get

$$\frac{f(\theta, \theta')}{\sqrt{-ir}} \begin{pmatrix} \cos\phi e^{-i\theta} \\ \kappa \\ \sin\phi e^{i\theta} \end{pmatrix} e^{ikr} = \lim_{kr \gg 1} \sum_l a_l \sum_{l'} T_{ll'} \begin{pmatrix} \cos\phi H^{(1)}_{l'-1}(kr) e^{-i\theta} \\ i\kappa H^{(1)}_{l'}(kr) \\ -\sin\phi H^{(1)}_{l'+1}(kr) e^{i\theta} \end{pmatrix} e^{il'\theta}. \tag{2.32}$$

Using the asymptotic property of the Hankel function [215]

$$H_l^{(1)}(z) \approx \sqrt{\frac{2}{\pi z}} \exp\left[i \left(z - \frac{1}{2} l\pi - \frac{1}{4}\pi \right) \right] \tag{2.33}$$

and comparing both sides of Eq. (2.32), we get

$$\begin{aligned} f_l(\theta, \theta') &= \sqrt{\frac{2}{\pi k}} \sum_l \sum_{l'} a_l T_{ll'} (-i)^{l'} e^{il'\theta} \\ &= \sqrt{\frac{2}{\pi k}} \sum_l f_l(\theta') e^{il\theta} \end{aligned} \tag{2.34}$$

where

$$f_l(\theta') = \sum_m a_m T_{ml} (-i)^l. \tag{2.35}$$

Once the scattering matrix has been obtained, we can analyze the far-field behaviors. For example, the differential cross section is

$$\frac{d\sigma}{d\theta} = |f(\theta, \theta')|^2 = \frac{2}{\pi k} \left| \sum_l f_l(\theta') e^{il\theta} \right| \tag{2.36}$$

and the total scattering cross section is given by

$$\sigma(\theta') = \oint d\theta |f(\theta, \theta')|^2 = \frac{4}{k} \sum_{ll'} a_l (TT^\dagger)_{ll'} a^*_{l'}. \tag{2.37}$$

2.7 Scattering from a Circular Cavity

Solution of scattering wavefunctions. the Incident wave is

$$\chi_{\text{in}}(\mathbf{r}) = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos\phi \exp(i\theta') \\ \kappa_0 \\ \sin\phi \exp(-i\theta') \end{pmatrix} e^{ik_0 r\cos\theta}, \tag{2.38}$$

where θ' is the incident angle and $k_0 = |E|/v_g$ is the incident wavevector. Using the Jacobi-Anger identity [215]

$$e^{iz\cos\theta} \equiv \sum_{l=-\infty}^{\infty} i^l J_l(z) e^{il\theta}, \tag{2.39}$$

we can expand the plane wave in the polar coordinates as

$$\begin{aligned} \chi_{\text{in}}(\mathbf{r}) &= \frac{1}{\sqrt{2}} \sum_l i^{l-1} \begin{pmatrix} \cos\phi J_{l-1}(k_0 r) e^{i(l-1)\theta} \\ i\kappa_0 J_l(k_0 r) e^{il\theta} \\ -\sin\phi J_{l+1}(k_0 r) e^{i(l+1)\theta} \end{pmatrix} \\ &= \sqrt{\pi} \sum_{l=-\infty}^{l=\infty} i^{l-1}\, {}^{k_0}\psi_l^{(0)}. \end{aligned} \tag{2.40}$$

From Fig. 2.1, there are two regions. In region I (outside of the cavity), the wavevector is k and the band index is $\kappa_0 = \text{Sign}E$. Inside the cavity, the wavevector is $q = |E - V|/v_g$ and the band index is $\kappa_1 = \text{Sign}(E - V)$. The wavefunction in region I is

$$\Psi^{(\text{I})}(\mathbf{r}) = \chi_{\text{in}}(\mathbf{r}) + \sqrt{\pi} \sum_{l=-\infty}^{\infty} i^{l-1} A_l\, {}^{k_0}\psi_l^{(1)}, \tag{2.41}$$

where A_l's are the expansion coefficients. Similarly, in region II, we have

$$\Psi^{(\text{II})}(\mathbf{r}) = \sqrt{\pi} \sum_{l=-\infty}^{\infty} i^{l-1} B_l\, {}^{q}\psi_l^{(0)}. \tag{2.42}$$

Imposing boundary conditions and matching the wavefunctions for each angular momentum channel, we get

$$\begin{aligned} B_l J_l(qR) &= \kappa_0\kappa_1[J_l(k_0R) + A_l H_l^{(1)}(k_0R)], \\ B_l X_l^{(0)}(qR) &= X_l^{(0)}(k_0R) + A_l X_l^{(1)}(k_0R), \end{aligned} \tag{2.43}$$

where

$$X_l^{(0,1,2)} = \cos^2\phi h_{l-1}^{(0,1,2)} - \sin^2\phi h_{l+1}^{(0,1,2)}. \tag{2.44}$$

We thus have

$$\begin{aligned} A_l &= -\frac{J_l(qR)X_l^{(0)}(k_0R) - \kappa_0\kappa_1 X_l^{(0)}(qR)J_l(k_0R)}{J_l(qR)X_l^{(1)}(k_0R) - \kappa_0\kappa_1 X_l^{(0)}(qR)H_l^{(1)}(k_0R)}, \\ B_l &= \frac{H_l^{(1)}(k_0R)X_l^{(0)}(k_0R) - X_l^{(1)}(k_0R)J_l(k_0R)}{H_l^{(1)}(k_0R)X_l^{(0)}(qR) - \kappa_0\kappa_1 X_l^{(1)}(k_0R)J_l(qR)}. \end{aligned} \tag{2.45}$$

Recalling the definition of the scattering matrix S^{cd}

$$\Psi^{(\mathrm{I})} = \sum_{l=-\infty}^{\infty} a_l \left({}^{k_0}\psi_l^{(2)} + S_l^{cd\,{}^{k_0}}\psi_l^{(1)} \right) \tag{2.46}$$

and comparing this with Eq. (2.41), we obtain

$$S_l^{cd} = 1 + 2A_l = -\frac{J_l(qR)X_l^{(2)}(k_0R) - \kappa_0\kappa_1 X_l^{(0)}(qR)H_l^{(2)}(k_0R)}{J_l(qR)X_l^{(1)}(k_0R) - \kappa_0\kappa_1 X_l^{(0)}(qR)H_l^{(1)}(k_0R)}. \tag{2.47}$$

From Eqs. (2.36) and (2.37), we obtain the differential cross section as

$$\frac{d\sigma}{d\theta} = \frac{1}{\pi k_0}\left|\sum_l (S_l^{cd} - 1)e^{il\theta}\right|^2 \tag{2.48}$$

and the following total scattering cross section

$$\sigma = \frac{2}{k_0}\sum_l \left|S_l^{cd} - 1\right|^2, \tag{2.49}$$

where the factor $(2/k_0)$ is the result of plane wave normalization.

Far field properties. In the quantum-dot regime $k_0R \ll 1$, for pseudospin-1/2 scattering the main contribution to the scattering wavefunction comes from the $\pm 1/2$ angular momentum channels. For pseudospin-1 scattering [96], the main contribution comes from the lowest angular momentum channel $l = 0$. For α-T_3 scattering, we have

$$A_1 = -\frac{J_1(qR)X_1^{(0)}(k_0R) - \kappa_0\kappa_1 X_1^{(0)}(qR)J_1(k_0R)}{J_1(qR)X_1^{(1)}(k_0R) - \kappa_0\kappa_1 X_1^{(0)}(qR)H_l^{(1)}(k_0R)}. \tag{2.50}$$

Recall the asymptotic properties of the Bessel and Hankel functions [215]:

$$\lim_{z\to 0} J_0(z) \approx 1, \tag{2.51}$$

$$\lim_{z\to 0} J_l(z) \approx \frac{1}{\Gamma(l+1)}\left(\frac{1}{2}z\right)^l, \tag{2.52}$$

$$\lim_{z\to 0} H_0^{(1)}(z) \approx -H_0^{(2)}(z) \approx \frac{2i}{\pi}\ln(z), \tag{2.53}$$

$$\lim_{z\to 0} H_l^{(1)}(z) \approx -H^{(2)}(z) \approx -\frac{i}{\pi}\Gamma(l)\left(\frac{1}{2}z\right)^{-l}, \tag{2.54}$$

for $l > 0$. For negative values of l, we have $J_{-l} = (-1)^l J_l$ and $H_{-l}^{(1,2)} = (-1)^l H_l^{(1,2)}$. We treat α as a perturbation parameter. For $\alpha \leq 0$, the dominant perturbation term is $J_1(qR)H_1^{(1)}(k_0R)$ in the denominator of Eq. (2.50), which has the form $\alpha^2 H_2^{(1)}(k_0R)$. We get

$$A_1 = \frac{A_1^{(\alpha=0)}}{1 + \eta\alpha^2/(k_0R)}, \tag{2.55}$$

where η is a parameter. Note that A_1 decays fast in the quantum-dot regime. The backscattering ratio is given by

$$\delta = \frac{d\sigma/d\theta|_{\theta=\pi}}{d\sigma/d\theta|_{\theta=0}} = \frac{|A_0 - A_1|^2}{|A_0 + A_1|^2} \propto \alpha^4(k_0R)^{-2}. \tag{2.56}$$

Low field property and level degeneracy. We analyze the maximum value of B_l. For $l \neq 0$, the maximum occurs when the denominator of Eq. (2.45) vanishes. In

the pseudospin-1/2 case $\alpha = 0$, Eq. (2.43) becomes

$$\begin{aligned} B_l J_l(qR) &= \kappa_0\kappa_1[J_l(k_0R) + A_l H_l^{(1)}(k_0R)], \\ B_l J_{l-1}(qR) &= J_{l-1}(k_0R) + A_l H_{l-1}^{(1)}(k_0R). \end{aligned} \tag{2.57}$$

Using the transform $l \to -l+1$ and noting $J_{-l} = (-1)^l J_l$ and $H_{-l}^{(1)} = (-1)^l H_l^{(1)}$, we get $B_{-l+1} = \kappa_0\kappa_1 B_l$ and $A_{-l+1} = A_l$. For pseudospin-1 ($\alpha = 1$) scattering, Eq. (2.43) becomes

$$\begin{aligned} B_l J_l(qR) &= \kappa_0\kappa_1[J_l(k_0R) + A_l H_l^{(1)}(k_0R)], \\ B_l(J_{l-1}(qR) - J_{l+1}(qR)) &= J_{l-1}(k_0R) - J_{l+1}(k_0R) + A_l(H_{l-1}(k_0R) - H_{l+1}(k_0R)). \end{aligned} \tag{2.58}$$

With $l \to -l$ and using $J_{-l} = (-1)^l J_l$ and $H_{-l}^{(1)} = (-1)^l H_l^{(1)}$, we get $A_{-l} = A_l$ and $B_{-l} = B_l$. The wave function is degenerate with respect to the transform $l \to -l$.

For a general α-T_3 system, Eq. (2.43) contains the parameter α. We have

$$\begin{aligned} B_l J_l(qR) &= \kappa_0\kappa_1[J_l(k_0R) + A_l H_l^{(1)}(k_0R)], \\ B_l X_l^{(0)}(qR) &= X_l^{(0)}(k_0R) + A_l X_l^{(1)}(k_0R). \end{aligned} \tag{2.59}$$

For $\alpha \leq 0$, we have $\alpha \approx \sin\phi$ and, hence,

$$X_l^{(1)} \approx H_{l-1}^{(1)}\left[1 + \eta l(l-1)\left(\frac{\alpha}{kR}\right)^2\right], \tag{2.60}$$

where η is a parameter. In the numerator of B_l, $X_l^{(1)}(k_0R)$ cancels off $J_l(k_0R)$ for small k_0R and $l \neq 0$. The denominator has a similar behavior as in the far field case:

$$H_l^{(1)}(k_0R)J_l(qR) = \kappa_0\kappa_1 H_{l-1}^{(1)}\left[1 + \eta l(l-1)\left(\frac{\alpha}{k_0R}\right)^2\right] J_l^{(1)}(qR). \tag{2.61}$$

To satisfy this equation for small values of α, we consider $q \to q + \Delta_l$. Comparing the two sides of the equations, we get

$$\Delta_l \propto \frac{\alpha^2 l}{k_0R}, \tag{2.62}$$

for $l > 0$. For α close to one, we have $\cos\phi \approx \cos(\pi/4) + (\alpha - 1)\sin(\pi/4)$. Following a similar analysis, we have that $H_{l+1}^{(1)}(k_0 R)$ in the denominator is dominant and $H_l^{(1)}(k_0 R)$ is a perturbation. We get

$$\Delta_l \propto (1-\alpha)\frac{k_0 R}{l} \text{ For } |l| > 0. \tag{2.63}$$

2.8 Scattering of α-T_3 Wave from a Ring Cavity

Consider the ring cavity in Fig. 2.5(a1), where two boundaries divide the whole space into three regions. The wavevectors in the three regions are $k_0 = |E|/v_g$, $k_1 = |E - V_1|/v_g$, and $k_2 = |E - V_2|/v_g$ with the respective band indices $\kappa_0 = \text{sign}E$, $\kappa_1 = \text{sign}(E - V_1)$, and $\kappa_2 = \text{sign}(E - V_2)$. The wavefunctions in the three regions are given by

$$\Psi^{(\text{I})}(\mathbf{r}) = \sum_{l=-\infty}^{\infty} \left({}^{k_0}\psi_l^{(2)} + S_l^{ring\ k_0}\psi_l^{(1)} \right), \tag{2.64}$$

$$\Psi^{(\text{II})}(\mathbf{r}) = \sum_{l=-\infty}^{\infty} C_l \left({}^{k_1}\psi_l^{(2)} + S_l^{cd\ k_1}\psi_l^{(1)} \right), \tag{2.65}$$

$$\Psi^{(\text{III})}(\mathbf{r}) = \sum_{l=-\infty}^{\infty} D_l\, {}^{k_2}\psi_l^{(0)}. \tag{2.66}$$

Let $E_l = C_l S_l^{cd}$. Matching the boundary conditions for each angular momentum channel, we get

$$\begin{pmatrix} \kappa_1 H_l^{(2)}(k_1R_2) & -\kappa_2 J_l(k_2R_2) & \kappa_1 H_l^{(1)}(k_1R_2) & 0 \\ X_l^{(2)}(k_1R_2) & X_l^{(0)}(k_2R_2) & X_l^{(1)}(k_1R_2) & 0 \\ \kappa_1 H_l^{(2)}(k_1R_1) & 0 & \kappa_1 H_l^{(1)}(k_1R_1) & -\kappa_0 H_l^{(1)}(k_0R_1) \\ X_l^{(2)}(k_1R_1) & 0 & X_l^{(1)}(k_1R_1) & -X_l^{(1)}(k_0R_1) \end{pmatrix} \begin{pmatrix} C_l \\ D_l \\ E_l \\ S_l^{ring} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \kappa_0 H_l^{(2)}(k_0R_1) \\ X_l^{(2)}(k_0R_1) \end{pmatrix}. \tag{2.67}$$

The solutions of the coefficients are

$$\begin{aligned} C_l &= \kappa_0\kappa_1 \frac{H_l^{(2)}(k_0R_1) + H_l^{(1)}(k_0R_1)S_l^{cd}}{H_l^{(2)}(k_1R_1) + H_l^{(1)}(k_1R_1)S_l^{cd}}, \\ D_l &= \kappa_0\kappa_1 C_l \frac{H_l^{(2)}(k_1R_2) + H_l^{(1)}(k_1R_2)S_l^{cd}}{J_l(k_2R_2)}. \end{aligned} \tag{2.68}$$

We thus get the scattering matrix for the ring cavity as

$$S_l^{ring} = -\frac{\kappa_0 x_l H_l^{(2)}(k_0R_1) - \kappa_1 y_l X_l^{(2)}(k_0R_1)}{\kappa_0 x_l H_l^{(1)}(k_0R_1) - \kappa_1 y_l X_l^{(1)}(k_0R_1)}. \tag{2.69}$$

where

$$\begin{aligned} x_l &= X_l^{(2)}(k_1R_1) + X_l^{(1)}(k_1R_1)S_l^{cd}, \\ y_l &= H_l^{(2)} + H_l^{(1)}(k_1R_1)S_l^{cd}, \end{aligned}$$

with S_l^{cd} given by Eq. (2.47).

2.9 Scattering of α-T_3 Wave from an Eccentric Circular Cavity

Similar to scattering from a ring cavity, in an eccentric circular cavity there are three distinct regions with wavevectors k_0, k_1 and k_2. Because of the eccentricity,

there is mixing of wavefunctions from different angular-momentum channels. We can write

$$\Psi^{(\mathrm{I})}(\mathbf{r}) = \sum_{l=-\infty}^{\infty} a_l^0 \left[{}^{k_0}\psi_l^{(2)} + \sum_{l'=-\infty}^{\infty} S_{ll'}\, {}^{k_0}\psi_{l'}^{(1)} \right], \tag{2.70}$$

$$\Psi^{(\mathrm{II})}(\mathbf{r}) = \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} {}^{l}a_m^1 \left[{}^{k_1}\psi_l^{(2)} + \sum_{m'=-\infty}^{\infty} S_{mm'}^{od}\, {}^{k_1}\psi_{l'}^{(1)} \right], \tag{2.71}$$

where S^{od} is the scattering matrix for the inner circle, a transformed version of the inner circle in the corresponding ring cavity:

$$S^{od} = U^{-1} S^{cd} U, \tag{2.72}$$

where the transformation matrix is $U = [U_{l\mu}] = [J_{\mu-l}(k_1\xi)]$ and its inverse is $U^{-1} = [U_{ml}^{-1}] = [J_{m-l}(k_1\xi)]$. We match the boundary conditions

$$\begin{aligned} a_l^0 X_l^{(2)}(k_0 R_1)\delta_{lm} &= {}^{l}a_m^1 X_m^{(2)}(k_1 R_1) + \sum_j {}^{l}a_j^1 S_{jm}^{od} X_m^{(1)}(k_1 R_1), \\ i\kappa_0 \left[a_l^0 H_l(k_0 R_1)\delta_{lm} \right] &= i\kappa_1 \left[{}^{l}a_m^1 H_m^{(2)}(k_1 R_1) + \sum_j {}^{l}a_j^1 S_{jm}^{od} H_m^{(1)}(k_1 R_1) \right], \end{aligned} \tag{2.73}$$

and define

$$\begin{aligned} \mathbb{A}^0 = [a_l^0 \delta_{lm}], \mathbb{A} &= [{}^{l}a_m^1], \\ \mathbb{C}^{(1,2)} = [X_l^{(1,2)}(k_0 R_1)\delta_{lm}], \mathbb{D} &= [H_l^{(1,2)}(k_0 R_1)\delta_{lm}], \\ \mathbf{c}^{(1,2)} = [X_l^{(1,2)}(k_1 R_1)\delta_{lm}], \mathbf{d} &= [H_l^{(1,2)}(k_1 R_1)\delta_{lm}]. \end{aligned} \tag{2.74}$$

Equation (2.73) can be rewritten as a matrix equation

$$\begin{aligned} \mathbb{A}^0\mathbb{C}^{(2)} + \mathbb{A}^0 S \mathbb{C}^1 &= \mathbb{A}\mathbf{c}^{(2)} + \mathbb{A}S^{od}\mathbf{c}^{(1)}, \\ \kappa_0 \left[\mathbb{A}^0\mathbb{D}^{(2)} + \mathbb{A}^0 S \mathbb{D} \right] &= \kappa_1 \left[\mathbb{A}\mathbf{d}^2 + \mathbb{A}S^{od}\mathbf{d}^{(1)} \right]. \end{aligned} \tag{2.75}$$

We obtain

$$S = -\frac{\mathbb{D}^{(2)} - \kappa_0\kappa_1 \mathbb{D}^{(2)}\mathbb{E}}{\mathbb{D}^{(1)} - \kappa_0\kappa_1 \mathbb{D}^{(1)}\mathbb{E}}, \tag{2.76}$$

where $\mathbb{E} = \mathbb{F}^{-1}\mathbb{D}$, $\mathbb{F}$ and $\mathbb{G}$ are defined as

$$\begin{aligned} \mathbb{F} &= \mathbf{c}^{(2)} + S^{od}\mathbf{c}^{(1)}, \\ \mathbb{G} &= \mathbf{d}^{(2)} + S^{od}\mathbf{d}^{(1)}, \end{aligned}$$

and

$$\mathbb{A} = \frac{\mathbb{A}^0\mathbb{C}^{(2)} + \mathbb{A}^0 S\mathbb{C}^{(1)}}{\mathbf{c}^{(2)} + S^{cd}\mathbf{c}^{(1)}}, \tag{2.77}$$

with $\mathbb{A}^0$ being the coefficient for the incident wave.

2.10 MMP Method for Solving α-T_3 Wave Scattering from an Arbitrary Domain

We describe a generalized MMP method that can be used to solve the scattering of α-T_3 wave in an efficient way. The method was originated in optics [202, 203, 204, 205, 206] and recently adopted for pseudospin-1 wave scattering [201]. Consider the simple case of a single scattering cavity, where the regions outside of and inside it are denoted as I and II, respectively, with wavevectors k_1 and k_2 as well as band indices κ_1 and κ_2. Suppose a plane spinor wave is incident on the cavity, which can be written as

$$\chi^{\text{in}}(\mathbf{r}) = \begin{pmatrix} \chi_A^{\text{in}} \\ \chi_B^{\text{in}} \\ \chi_C^{\text{in}} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos\phi \exp(-i\theta') \\ \kappa_0 \\ \sin\phi \exp(i\theta') \end{pmatrix} e^{ik_0 r\cos\theta}. \tag{2.78}$$

To calculate the scattering wavefunction, we place a number of "poles" inside and outside of the cavity, as shown in Fig. 2.11. Poles inside and outside of the cavity are labeled as m_{II} and m_{I}, respectively, with their total numbers M_{II} and M_{I}. Each pole is treated as a fictitious source. To avoid the problem of singularity at a source, the scattering wave function outside of the cavity is given as the superposition of the waves from all the poles inside the cavity, while the wavefunction inside the cavity

is determined by the poles outside, as illustrated in Figs. 2.11(a-c). For a fictitious source, the base function is

$$\Psi_l(\mathbf{r}) = \frac{1}{\sqrt{2\pi}} \begin{pmatrix} \cos\phi h_{l-1}^{(0,1,2)}(kr)e^{-i\theta} \\ i\kappa h_l^{(0,1,2)}(kr) \\ -\sin\phi h_{l+1}^{(0,1,2)}(kr)e^{i\theta} \end{pmatrix} e^{il\theta}, \tag{2.79}$$

where $h_l^{(0,1,2)}(z)$ denotes some type of Bessel function. For each source, we choose the outgoing wave as the basis, i.e., the Hankel function of the first kind: $H_l^{(1)}$. The wavefunction outside of the cavity can be written in terms of the poles inside as

$$\Psi^{(\mathrm{I})}(\mathbf{r}) = \begin{pmatrix} \psi_A^{(\mathrm{I})} \\ \psi_B^{(\mathrm{I})} \\ \psi_C^{(\mathrm{I})} \end{pmatrix} = \sum_{m_{\mathrm{II}}} \sum_l F_l^{m_{\mathrm{II}}} \begin{pmatrix} \cos\phi H_{l-1}^{(1)}(k_0 d_{m_{\mathrm{II}}})e^{-i\theta_{m_{\mathrm{II}}}} \\ i\kappa_0 H_l^{(1)}(k_0 d_{m_{\mathrm{II}}}) \\ -\sin\phi H_{l+1}^{(1)}(k_0 d_{m_{\mathrm{II}}})e^{i\theta_{m_{\mathrm{II}}}} \end{pmatrix} e^{il\theta_{m_{\mathrm{II}}}}, \tag{2.80}$$

where

$$\begin{aligned} d_{m_{\mathrm{II}}} &\equiv |\mathbf{d}_{m_{\mathrm{II}}}| = |\mathbf{r} - \mathbf{r}_{m_{\mathrm{II}}}|, \\ \theta_{m_{\mathrm{II}}} &\equiv \mathrm{Angle}(\mathbf{r} - \mathbf{r}_{m_{\mathrm{II}}}), \end{aligned}$$

and $F_l^{m_{\mathrm{II}}}$ is the expansion coefficient for the eigenvector from the m_{II}th pole with angular momentum index l located at $\mathbf{r}_{m_{\mathrm{II}}}$.

Similarly, the wave function inside the cavity is determined by the poles outside of the cavity:

$$\Psi^{(\mathrm{II})}(\mathbf{r}) = \begin{pmatrix} \psi_A^{(\mathrm{II})} \\ \psi_B^{(\mathrm{II})} \\ \psi_C^{(\mathrm{II})} \end{pmatrix} = \sum_{m_{\mathrm{I}}} \sum_l F_l^{m_{\mathrm{I}}} \begin{pmatrix} \cos\phi H_{l-1}^{(1)}(q d_{m_{\mathrm{I}}})e^{-i\theta_{m_{\mathrm{I}}}} \\ i\kappa_1 H_l^{(1)}(q d_{m_{\mathrm{I}}}) \\ -\sin\phi H_{l+1}^{(1)}(q d_{m_{\mathrm{I}}})e^{i\theta_{m_{\mathrm{I}}}} \end{pmatrix} e^{il\theta_{m_{\mathrm{I}}}}, \tag{2.81}$$

where

$$\begin{aligned} d_{m_{\mathrm{I}}} &\equiv |\mathbf{d}_{m_{\mathrm{I}}}| = |\mathbf{r} - \mathbf{r}_{m_{\mathrm{I}}}|, \text{ and} \\ \theta_{m_{\mathrm{I}}} &\equiv \mathrm{Angle}(\mathbf{r} - \mathbf{r}_{m_{\mathrm{I}}}). \end{aligned}$$

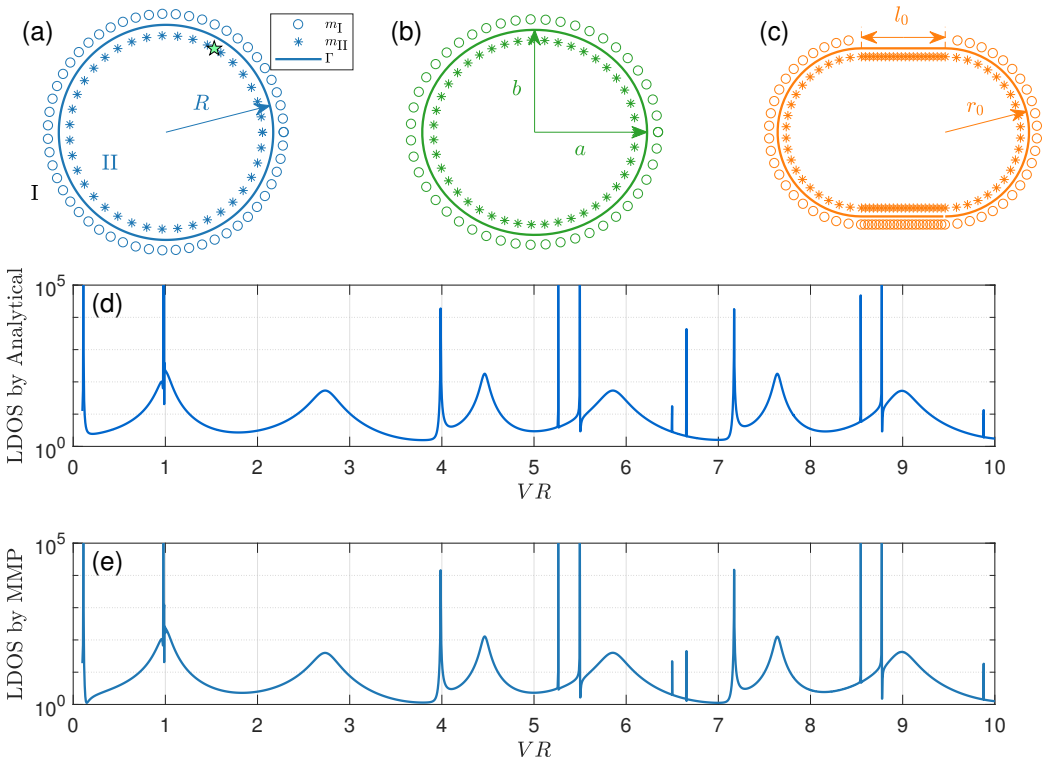

Figure 2.11: Illustration of MMP method for α-T_3 wave scattering and validation. (a-c) Distributions of interior and exterior poles for the circular, elliptical, and stadium cavities, respectively. For each cavity, the solid curve represents the cavity boundary Γ, and the regions outside of and inside the cavity are labeled as I and II, respectively. (d) For the circular cavity, analytically obtained local density of state (LDOS) versus the potential parameter VR at the test point ($r = 0.9$ and $\theta = \pi/3$) specified by the green pentagram in (a). (e) The LDOS at the same point obtained from MMP. The agreement between the results in (d) and (e) is excellent.

To match the boundary conditions, the three components of the spinor wavefunction must satisfy

$$\begin{aligned} \left(\chi_B^{\mathrm{in}} + \psi_B^{(\mathrm{I})}\right)\Big|_{\mathbf{r}\in\Gamma} &= \psi_B^{(\mathrm{II})}\Big|_{\mathbf{r}\in\Gamma}, \\ \left[\cos\phi\left(\chi_A^{\mathrm{in}} + \psi_A^{(\mathrm{I})}\right) + \sin\phi\left(\chi_C^{\mathrm{in}} + \psi_C^{(\mathrm{I})}\right)\right]\Big|_{\mathbf{r}\in\Gamma} &= \\ \left(\cos\phi\psi_A^{(\mathrm{II})} + \sin\phi\psi_C^{(\mathrm{II})}\right)\Big|_{\mathbf{r}\in\Gamma}. & \end{aligned}$$

Expanding this equation in terms of the poles and the corresponding eigenvectors,

we obtain

$$\begin{aligned}
\sum_{m_{\rm II}}\sum_{l} {}^{j}P^{(\rm I)}_{lm_{\rm II}}F^{m_{\rm II}}_{l} - \sum_{m_{\rm I}} {}^{j}P^{(\rm II)}_{lm_{\rm I}}F^{m_{\rm I}}_{l} &= -{}^{j}\chi^{\rm in}_{B}, \\
\sum_{m_{\rm II}}\sum_{l} {}^{j}Q^{(\rm I)}_{lm_{\rm II}}F^{m_{\rm II}}_{l} - \sum_{m_{\rm I}} {}^{j}Q^{(\rm II)}_{lm_{\rm I}}F^{m_{\rm I}}_{l} &= \\
&\quad -\left(\cos\phi\,{}^{j}\chi^{\rm in}_{A} + \sin\phi\,{}^{j}\chi^{\rm in}_{C}\right),
\end{aligned} \tag{2.82}$$

where

$$\begin{aligned}
{}^{j}P^{(\rm I)}_{lm_{\rm II}} &= i\kappa_0 H^{(1)}_{l}(k_0|\mathbf{r}_j - \mathbf{r}_{m_{\rm II}}|)e^{il\theta_{m_{\rm II}}}, \\
{}^{j}P^{(\rm II)}_{lm_{\rm I}} &= i\kappa_1 H^{(1)}_{l}(k_1|\mathbf{r}_j - \mathbf{r}_{m_{\rm I}}|)e^{il\theta_{m_{\rm I}}}, \\
{}^{j}Q^{(\rm I)}_{lm_{\rm II}} &= \cos^2\phi H^{(1)}_{l-1}(k_0|\mathbf{r}_j - \mathbf{r}_{m_{\rm II}}|)e^{i(l-1)\theta_{m_{\rm II}}} \\
&\quad - \sin^2\phi H^{(1)}_{l+1}(k_0|\mathbf{r}_j - \mathbf{r}_{m_{\rm II}}|)e^{i(l+1)\theta_{m_{\rm II}}}, \\
{}^{j}Q^{(\rm II)}_{lm_{\rm I}} &= \cos^2\phi H^{(1)}_{l-1}(k_1|\mathbf{r}_j - \mathbf{r}_{m_{\rm I}}|)e^{i(l-1)\theta_{m_{\rm I}}} \\
&\quad - \sin^2\phi H^{(1)}_{l+1}(k_1|\mathbf{r}_j - \mathbf{r}_{m_{\rm I}}|)e^{i(l+1)\theta_{m_{\rm I}}},
\end{aligned} \tag{2.83}$$

and

$${}^{j}\chi^{\rm in}_{A,B,C} = \chi^{\rm in}_{A,B,C}(\mathbf{r}_j). \tag{2.84}$$

We discretize the boundary into J points. For each point, we truncate the angular momentum as $l \in [-L, L]$. Let $N_L = 2L+1$ be the number of eigenvectors associated with different values of the angular momentum and let $N = N_L \times (M_{\rm I} + M_{\rm II}) = N_{\rm I} + N_{\rm II}$. The boundary conditions lead to the following matrix equation

$$\mathbb{M}_{2J\times N}\cdot\mathbf{F}_{N\times 1} = -\mathbf{G}_{2J\times 1}, \tag{2.85}$$

where $\mathbb{M}$ can be expanded in terms of the matrices $\mathbb{P}$ and $\mathbb{Q}$, which are typically not

square matrices:

$$\mathbb{M}_{2J\times N} = \left(\begin{array}{c|c} \mathbb{P}^{(\mathrm{I})} & -\mathbb{P}^{(\mathrm{II})} \\ \hline \mathbb{Q}^{(\mathrm{I})} & -\mathbb{Q}^{(\mathrm{II})} \end{array}\right)_{2J\times N} . \tag{2.86}$$

The two vectors in Eq. (2.85) are given by

$$\mathbf{F}_{N\times 1} = \begin{pmatrix} F_{-L}^{1_{\rm II}} \\ \vdots \\ F_{l}^{1_{\rm II}} \\ F_{l}^{2_{\rm II}} \\ \vdots \\ F_{l}^{M_{\rm II}} \\ \vdots \\ F_{L}^{M_{\rm II}} \\ \hline F_{-L}^{1_{\rm I}} \\ \vdots \\ F_{l}^{1_{\rm I}} \\ F_{l}^{2_{\rm I}} \\ \vdots \\ F_{l}^{M_{\rm I}} \\ \vdots \\ F_{L}^{M_{\rm I}} \end{pmatrix}_{N\times 1} , \mathbf{G}_{2J\times 1} = \begin{pmatrix} {}^{j}\chi_{B}^{\rm in} \\ \vdots \\ {}^{j}\chi_{B}^{\rm in} \\ \vdots \\ {}^{J}\chi_{B}^{\rm in} \\ \hline \cos\phi\, {}^{1}\chi_{A}^{\rm in} + \sin\phi\, {}^{1}\chi_{A}^{\rm in} \\ \vdots \\ \cos\phi\, {}^{j}\chi_{A}^{\rm in} + \sin\phi\, {}^{j}\chi_{A}^{\rm in} \\ \vdots \\ \cos\phi\, {}^{J}\chi_{A}^{\rm in} + \sin\phi\, {}^{J}\chi_{A}^{\rm in} \end{pmatrix}_{2J\times 1} , \tag{2.87}$$

and the matrices $\mathbb{P}$ and $\mathbb{Q}$ are given by

$$\mathbb{P}^{(\lambda)} = \begin{pmatrix} {}^{1}P^{(\lambda)}_{-L1_{\overline{\lambda}}} \cdots {}^{1}P^{(\lambda)}_{-LM_{\overline{\lambda}}} \cdots {}^{1}P^{(\lambda)}_{l1_{\overline{\lambda}}} \cdots {}^{1}P^{(\lambda)}_{lM_{\overline{\lambda}}} \cdots {}^{1}P^{(\lambda)}_{LM_{\overline{\lambda}}} \\ {}^{2}P^{(\lambda)}_{-L1_{\overline{\lambda}}} \cdots {}^{2}P^{(\lambda)}_{-LM_{\overline{\lambda}}} \cdots {}^{2}P^{(\lambda)}_{l1_{\overline{\lambda}}} \cdots {}^{2}P^{(\lambda)}_{lM_{\overline{\lambda}}} \cdots {}^{2}P^{(\lambda)}_{LM_{\overline{\lambda}}} \\ \vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \\ {}^{j}P^{(\lambda)}_{-L1_{\overline{\lambda}}} \cdots {}^{j}P^{(\lambda)}_{-LM_{\overline{\lambda}}} \cdots {}^{j}P^{(\lambda)}_{l1_{\overline{\lambda}}} \cdots {}^{j}P^{(\lambda)}_{lM_{\overline{\lambda}}} \cdots {}^{j}P^{(\lambda)}_{LM_{\overline{\lambda}}} \\ \vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \\ {}^{J}P^{(\lambda)}_{-L1_{\overline{\lambda}}} \cdots {}^{J}P^{(\lambda)}_{-LM_{\overline{\lambda}}} \cdots {}^{J}P^{(\lambda)}_{l1_{\overline{\lambda}}} \cdots {}^{J}P^{(\lambda)}_{lM_{\overline{\lambda}}} \cdots {}^{J}P^{(\lambda)}_{LM_{\overline{\lambda}}} \end{pmatrix}_{J\times N_{\overline{\lambda}}}, \tag{2.88}$$

and

$$\mathbb{Q}^{(\lambda)} = \begin{pmatrix} {}^{1}Q^{(\lambda)}_{-L1_{\overline{\lambda}}} \cdots {}^{1}Q^{(\lambda)}_{-LM_{\overline{\lambda}}} \cdots {}^{1}Q^{(\lambda)}_{l1_{\overline{\lambda}}} \cdots {}^{1}Q^{(\lambda)}_{lM_{\overline{\lambda}}} \cdots {}^{1}Q^{(\lambda)}_{LM_{\overline{\lambda}}} \\ {}^{2}Q^{(\lambda)}_{-L1_{\overline{\lambda}}} \cdots {}^{2}Q^{(\lambda)}_{-LM_{\overline{\lambda}}} \cdots {}^{2}Q^{(\lambda)}_{l1_{\overline{\lambda}}} \cdots {}^{2}Q^{(\lambda)}_{lM_{\overline{\lambda}}} \cdots {}^{2}Q^{(\lambda)}_{LM_{\overline{\lambda}}} \\ \vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \\ {}^{j}Q^{(\lambda)}_{-L1_{\overline{\lambda}}} \cdots {}^{j}Q^{(\lambda)}_{-LM_{\overline{\lambda}}} \cdots {}^{j}Q^{(\lambda)}_{l1_{\overline{\lambda}}} \cdots {}^{j}Q^{(\lambda)}_{lM_{\overline{\lambda}}} \cdots {}^{j}Q^{(\lambda)}_{LM_{\overline{\lambda}}} \\ \vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \quad \cdots \quad \vdots \\ {}^{J}Q^{(\lambda)}_{-L1_{\overline{\lambda}}} \cdots {}^{J}Q^{(\lambda)}_{-LM_{\overline{\lambda}}} \cdots {}^{J}Q^{(\lambda)}_{l1_{\overline{\lambda}}} \cdots {}^{J}Q^{(\lambda)}_{lM_{\overline{\lambda}}} \cdots {}^{J}Q^{(\lambda)}_{LM_{\overline{\lambda}}} \end{pmatrix}_{J\times N_{\overline{\lambda}}}, \tag{2.89}$$

where $\lambda \in \{\mathrm{I}, \mathrm{II}\}$ and $\overline{\lambda} = \{\mathrm{I}, \mathrm{II}\} - \lambda$.

Because of the finite truncation, the total basis is not complete, so Eq. (2.86) can be satisfied only approximately. To achieve high accuracy requires $J \gg N$. Equation (2.86) can be solved by the pseudo-inverse method [$\mathbf{F} = -\mathrm{pinv}(\mathbb{M} * \mathbf{G})$ in Matlab].

From the expansion coefficients $\mathbf{F}$, we can get the scattering wavefunction in the whole space. To validate the MMP method, we set $\alpha = 1/3$ in the circular cavity and calculate the local density of states (LDOS) as a function of the parameter VR both analytically and using the MMP method. The results are shown in Figs. 2.11(d) and 2.11(e). There is a good agreement between the results.

The error of the MMP method can be estimated as

$$\mathrm{SSE} = \frac{\|\mathbb{M} * \mathbf{F} + \mathbf{G}\|}{\|\mathbf{G}\|}. \tag{2.90}$$

For the circular and the elliptical cavities, the relative errors are typically smaller than 10^{-5}. For the stadium cavity, the error is bounded by 0.05. In the quantum-dot regime where the wavevector inside the cavity is large, a large number of poles are needed and they should be placed as close to the boundary as possible.

2.11 Appendix

Following are the relevant publication on the topics presented in this chapter

C.-D. Han, H.-Y. Xu, and Y.-C. Lai*, "Electrical confinement in a spectrum of two-dimensional Dirac materials with classically integrable, mixed, and chaotic dynamics," Physical Review Research **2**, 013116, (2020).

Chapter 3

ADAPTABLE HAMILTONIAN NEURAL NETWORKS

The rapid growth of research in exploiting machine learning to predict chaotic systems has revived a recent interest in Hamiltonian Neural Networks (HNNs) with physical constraints defined by the Hamilton's equations of motion, which represent a major class of physics-enhanced neural networks. We introduce a class of HNNs capable of adaptable prediction of nonlinear physical systems: by training the neural network based on time series from a small number of bifurcation-parameter values of the target Hamiltonian system, the HNN can predict the dynamical states at other parameter values, where the network has not been exposed to any information about the system at these parameter values. The architecture of the HNN differs from the previous ones in that we incorporate an input parameter channel, rendering the HNN parameter–cognizant. We demonstrate, using paradigmatic Hamiltonian systems, that training the HNN using time series from as few as four parameter values bestows the neural machine with the ability to predict the state of the target system in an entire parameter interval. Utilizing the ensemble maximum Lyapunov exponent and the alignment index as indicators, we show that our parameter-cognizant HNN can successfully predict the route of transition to chaos. Physics-enhanced machine learning is a forefront area of research, and our adaptable HNNs provide an approach to understanding machine learning with broad applications.

3.1 Introduction

A daunting challenge in machine learning is the lack of understanding of the inner working of the artificial neural networks. As machine learning has been increasingly

incorporated into many vital structures and systems that support the functioning of the modern society, it is imperative to develop a general understanding of the inner gears of the underlying neural networks. For example, feed-forward neural networks or multilayer perceptrons constitute the fundamentals of modern deep learning machines with broad applications in image, video and audio processing [216]. Such a neural machine typically consists of an input layer, a large number of hidden layers, and an output layer. From the input layer on, nodes in the same layer do not interact with each other, but they are connected with the nodes in the next layer via a set of weights and biases whose values are determined through training, where the paradigmatic method of stochastic gradient descent (SGD) [217] is often used. How the networks in different layers work together to solve a specific problem remains unknown. In another line of research, reservoir computing, a class of recurrent neural networks [218, 219, 220, 221], has gained considerable momentum since 2017 as a powerful paradigm for model-free, fully data driven prediction of nonlinear and chaotic dynamical systems [222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238]. A reservoir computing machine constitutes an input layer, a single hidden layer, and an output layer. Differing from the network structure of a multilayer perceptron, the network in the hidden layer of a reservoir computing machine has a complex topology in which the nodes are coupled with each other following some probability distribution. Another difference is that, in feed-forward neural networks, only the weights and biases connecting the hidden layer and the output layer neurons are determined by training, while in reservoir computing those parameters as well as the weights of the complex network in the hidden layer are pre-defined. A well trained reservoir machine can generate accurate prediction of the state evolution of a chaotic system for a duration that is typically several times longer than that which can be achieved using the traditional methodologies in nonlinear time series analysis.

This is remarkable, considering the hallmark of chaos: sensitive dependence on initial conditions, which rules out long-term prediction. Yet, there is little understanding of how the internal network dynamics of reservoir computing machines behave or "manage" to replicate accurately (for some amount of time) the chaotic evolution of the true system.

At the present, to develop a general explainable framework to encompass various types of machine learning is not feasible. In this regard, a promising direction of pursuit is the so-called physics-enhanced machine learning, in which the neural networks are designed to solve specific physics problems with the goal to enhance the learning efficiency through exploiting the underlying physical principles or constraints. The idea was articulated almost three decades ago [160], when the principle of Hamiltonian mechanics was incorporated into the design of neural networks, leading to *Hamiltonian Neural Networks* (HNNs) that have recently gained renewed attention [153, 154, 155, 156, 157]. Comparing with traditional neural networks, in an HNN, the energy is conserved. It has been demonstrated that, an HNN can be trained to possess the power to predict the dynamical evolution of the target Hamiltonian system in both integrable and chaotic regimes, provided that the network is trained with data taken from the same set of parameter values at which the prediction is to be made [153, 154, 155, 156, 157]. Recently the principle of HNN has been generalized [239] to systems described by the Lagrangian equation of motion [240] and a general type of ordinary differential equations [241] or coordinate transforms [242, 243] with applications in robotics [244, 245].

In this paper, we address adaptability, a fundamental issue in machine learning, of Hamiltonian neural networks. More precisely, we consider the situation where a target Hamiltonian system can experience slow drift or sudden changes in some parameters. Slow environmental variations can lead to adiabatic parameter drifting,

while external disturbances can lead to sudden parameter changes. We ask if it is possible to design HNNs, which are trained with data from a small number of parameter values of the target system, to have the predictive power for parameter values that are not in the training set. Inspired by the recent work on predicting critical transitions and collapse in dissipative dynamical systems based on reservoir computing [148, 149, 150, 151, 152], we articulate a class of HNNs whose input layer contains a set of channels that are specifically used for inputting the values of the distinct parameters of interest to the neural network. The number of the parameter channels is equal to the number of freely varying parameters in the target Hamiltonian system. The simplest case is where the target system has a single bifurcation or control parameter so only one input parameter channel to the neural network is necessary. We demonstrate that, by incorporating such a parameter channel into a feed-forward type of HNNs and conducting training using time series data from a small number of bifurcation parameter values (e.g., four), we effectively make the HNN adaptable to parameter variations. That is, the so-trained HNN has inherited the rules governing the dynamical evolution of the target Hamiltonian system. When a parameter value of interest, which is not in the training parameter set, is fed into the HNN through the parameter channel, the machine is capable of generating dynamical behaviors that statistically match those of the target system at this particular parameter value. The HNN has thus become adaptable because it has never been exposed to any information or data from the target system at this parameter value, yet the neural machine can reproduce the dynamical behavior. Using the Hénon-Heiles model as a prototypical target Hamiltonian system, we demonstrate that our adaptable HNN can successfully predict the dynamical behaviors, integrable or chaotic, for any parameter values that are reasonably close to those in the training parameter set. Remarkably, by feeding a systematically varying set of bifurcation parameter values into the

parameter channel, the HNN can successfully predict the transition to chaos in the target Hamiltonian system, which we characterize using two measures: the ensemble maximum Lyapunov exponent and the alignment index. It is worth emphasizing that, in the existing literature on HNNs [153, 154, 155, 156, 157], training and prediction are done at the same set of parameter values of the target Hamiltonian system, but our work goes beyond by making the HNN significantly more powerful with enhanced and expanded predictability.

We remark that, in physics, machine learning has been exploited to solve difficult problems in particle physics [246, 247], quantum many-body systems [248], inverse design in optical systems [249], and quantum information [250, 158]. However, the working mechanisms of the underlying neural networks remain largely unknown [251]. The physics enhanced HNNs studied here are different from these applications, as we focus on exploiting physical principles to enable neural networks with unprecedented predictive power with respect to parameter variations.

In Sec. 3.2, we describe the architecture of the articulated parameter-cognizant HNNs and the method of training. In Sec. 3.3, we present results of predicting the dynamical behavior of the Hénon-Heiles system in a wide parameter region, including the transition to chaos based on calculating the ensemble maximum Lyapunov exponent and the minimum alignment index. In general, the prediction accuracy depends on how "close" the desired parameter value is to the training regime. In Sec. 3.4, we address a number of pertinent issues such as the choosing of the training parameter values, multiple parameter channels, and HNNs for a Hamiltonian system defined by the one-dimensional Morse potential. A summarizing discussion and speculations are offered in Sec. 3.5.

3.2 Parameter-cognizant Hamiltonian Neural Networks

The central idea for physics-enhanced machine learning is to "force" the dynamical evolution of the neural network to follow certain physical rules or constraints, examples of which are Hamilton's equations of motion [153, 154, 155, 156, 157], Lagrangian equations [240], or the principle of least action [252, 253]. In particular, the structure of HNNs is such that the underlying neural dynamical system is effectively a Hamiltonian system for which the energy is conserved during the evolution. Different from previous work [153, 154, 155, 156, 157], the bifurcation parameter of the target Hamiltonian system serves as an input "variable" to the neural network through an additional input channel so that the HNN learns to associate the input time series with the specific value of the bifurcation parameter. Using time series from a small number of distinct bifurcation parameter values to train the HNN, it can gain the ability to "sense" the changes in the dynamics (or dynamical "climate") of the target system with the bifurcation parameter.

The structure of our articulated parameter-cognizant HNN is shown in Fig. 3.1, where the input contains three parts: the position and momentum variables of the target system, and the bifurcation parameter. To be concrete, we use two hidden layers, where each layer contains 200 artificial neurons (nodes). The third layer is the output, which contains a single node whose dynamical state corresponds to the Hamiltonian of the target system. Let $\mathbf{y}$ denote the set of dynamical variables of each layer. The transform from the dynamical variables in the ith layer to those in the $(i+1)$th layer follows the following rule:

$$\mathbf{y}^{i+1} = \boldsymbol{\sigma}^i(\mathcal{W}^i \cdot \mathbf{y}^i + \mathbf{b}^i), \tag{3.1}$$

where $\boldsymbol{\sigma}^i$ is a given nonlinear activation function, $\mathcal{W}^i$ is the weight matrix and $\mathbf{b}^i$ is bias vector associated with the neurons in the ith layer, which are to be determined

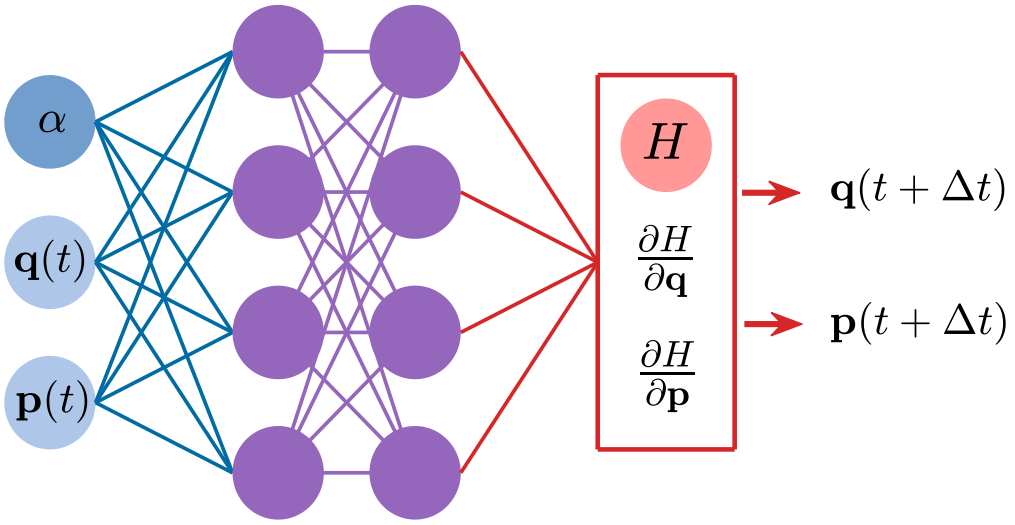

Figure 3.1: Structure of parameter-cognizant HNN. The input channels are denoted by the blue circles, which are connected to the first hidden layer (purple circles). The blue circle denoted by "α" is the parameter input channel that feeds the value of the bifurcation parameter of the target Hamiltonian system, together with the time series $\mathbf{q}(t)$ and $\mathbf{p}(t)$ through the corresponding input channels, into the first hidden layer. There are two hidden layers. The output variables are the partial derivatives of the Hamiltonian of the target system with respect to the canonical coordinates and momenta, together with the Hamiltonian, which determine the dynamical state at the next time step.

through training. We set the output as the spatial derivatives of the input variables to force the dynamics of the neural network to follow the Hamilton's equations of motion. The derivatives are calculated through the back prorogation algorithm. Once the output is known, the loss function defined as

$$\mathcal{L} = \left\| \frac{\partial H}{\partial \mathbf{q}} + \frac{d\mathbf{p}_{\text{real}}}{dt} \right\| + \left\| \frac{\partial H}{\partial \mathbf{p}} - \frac{d\mathbf{q}_{\text{real}}}{dt} \right\|, \tag{3.2}$$

can be calculated. Through the training process, we optimize the weights and biases in Eq. (3.1) by minimizing the loss function. This is done by using the standard SGD method [217]. The whole process from network construction and training to carrying out the prediction is accomplished by using the open source package Tensorflow and Keras [254, 255].

Table 3.1 summarizes the structure and parameters of our HNN. It has about

Table 3.1: Specifications of HNN

Description	Values
Number of hiden layers	2
Neurons per layer	200
Optimizer	Adam
Epochs	500
Activation functioin	tanh

40,000 unknown parameters to be optimally determined through training. The computation can be quite efficient even without using parallel or GPU acceleration. The anticipation is that, after training with time series data from a small number of distinct bifurcation parameter values, the HNN can predict the dynamical behavior of the target Hamiltonian system in a wide parameter interval, where the parameter variations are implemented through the input parameter channel to the HNN.

3.3 Adaptable Hamiltonian Neural Networks for Predicting Transition to Chaos

To test the adaptability of our parameter-cognizant HNN for predicting state evolution and dynamical transitions in Hamiltonian systems, we use the paradigmatic Hénon-Heiles model [256]. It is a two-degrees-of-freedom system for investigating distinct types of Hamiltonian dynamics including integrable, mixed, and chaotic behaviors and the transitions among them. It was originated from the gravitational three-body system [256], with applications in contexts such as molecular dynamics [257, 258, 259].

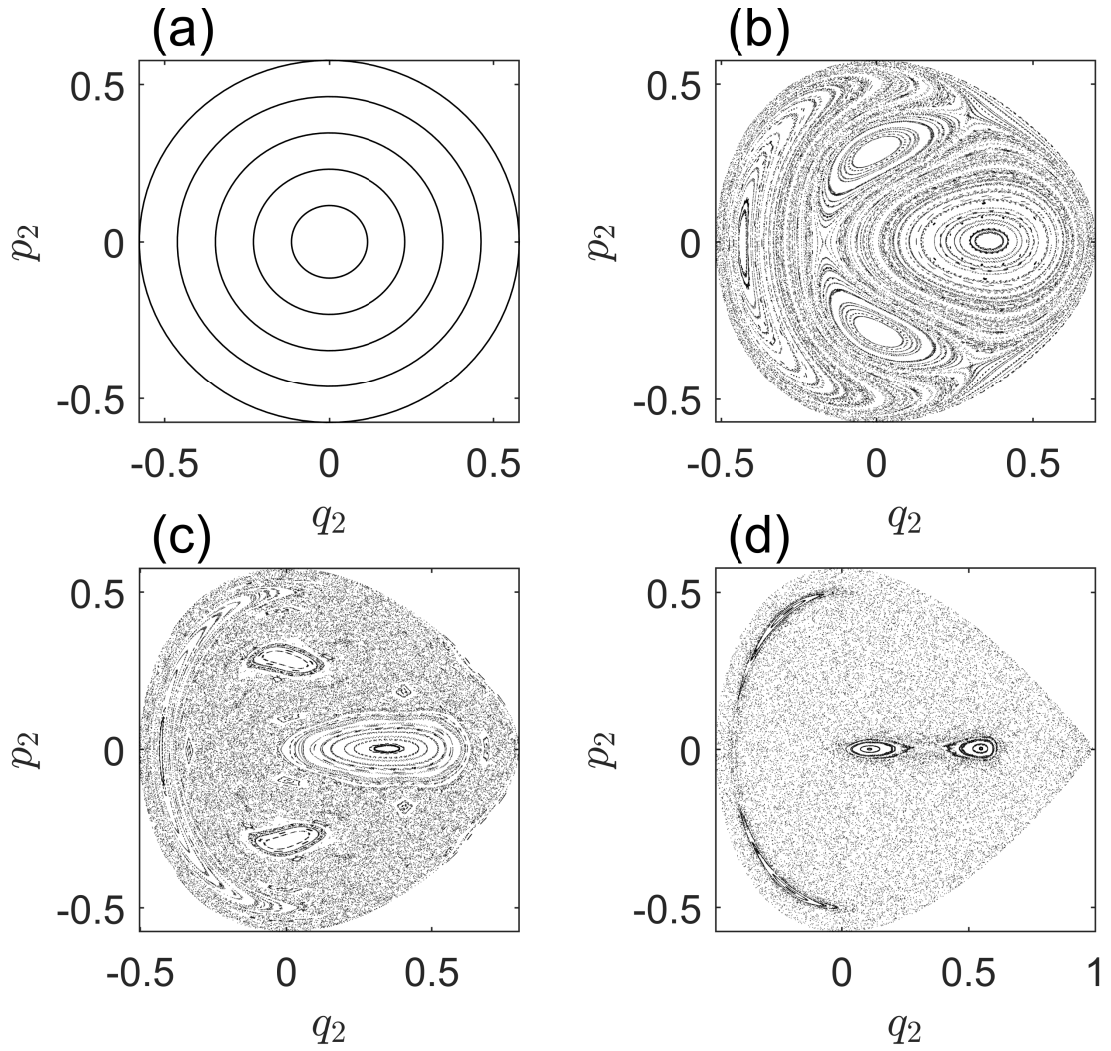

Figure 3.2: Distinct types of dynamical behaviors in the Hénon-Heiles system. The energy value is $E = 1/6$. (a-d) Integrable, mixed, small and largely chaotic dynamics for $\alpha = 0$, 0.7, 0.9, and 1.0, respectively, on the Poincaré surface of section defined by $q_1 = 0$.

3.3.1 System Description

The Hénon-Heiles Hamiltonian is

$$H = \frac{1}{2}\left(p_1^2 + p_2^2\right) + \frac{1}{2}\left(q_1^2 + q_2^2\right) + \alpha\left(q_1^2 q_2 - \frac{1}{3}q_2^3\right), \tag{3.3}$$

where q_1 and q_2 denote the coordinates, p_1 and p_2 are the corresponding momenta, $\alpha > 0$ is a bifurcation parameter that sets the magnitude of the nonlinear potential function describing, e.g., the dissociation energy in molecules [257, 258, 259]. The

dynamics of system Eq. (3.3) depend not only on α, but also on the energy E of the system that is conserved during the dynamical evolution. The maximum value of the potential function is $E_{\max} = 1/(6\alpha^2)$. For $\alpha = 0$, $E_{\max}$ diverges, so all trajectories are bounded. For $\alpha > 0$, if the particle energy exceeds $E_{\max}$, the Hamiltonian system becomes open with scattering trajectories that can escape to infinity. To train the adaptable HNN, bounded trajectories are required, so we set $0 \leq \alpha \leq 1$ and $E \leq 1/6$. (For particle energy above the threshold, chaotic scattering dynamics and fractal geometry can arise [260, 261, 262].) As the value of α increases from zero to one, characteristically different dynamical behaviors can arise, such as integrable, mixed, and chaotic. In particular, for $\alpha = 0$, the nonlinear term in Eq. (3.3) disappears and the system becomes a harmonic oscillator - an integrable system. In this case, the entire phase space contains periodic and quasiperiodic orbits only, as shown in Fig. 3.2(a). As α increases from zero, the system becomes nonlinear and chaotic seas amid the Kol'mogorov-Arnol'd-Moser (KAM) islands can arise in the phase space, giving rise to mixed dynamics, as shown in Fig. 3.2(b) for $\alpha = 0.7$ and $E = 1/6$. For $\alpha = 0.9$, $\alpha = 1$ and $E = 1/6$, most trajectories in the phase space are chaotic, as shown in Figs. 3.2(c) and 3.2(d).

3.3.2 Training and Testing of Adaptability

Our goal of training is to "instill" certain adaptable power into the HNN. To achieve this, we choose a number of distinct values of the bifurcation parameter α. For each α value, we randomly choose initial conditions with energy below the escape threshold $E_{\max}$ and numerically integrate the Hamilton's equations of motion to generate particle trajectories in the phase space. Because of the mixed dynamics, the training data contain both integrable and chaotic orbits. Specifically, the time interval of the trajectory is $0 \leq t \leq 1000$, which contains hundreds of oscillation

Table 3.2: List of training parameters for Hénon-Heiles system

Description	Values
Neural network ensembles	20
Energy samples	7
Orbit per energy	1
Orbit length	1000
Time step	0.1
Training parameter set	$\alpha \in \{0.2, 0.4, 0.6, 0.8\}$

cycles, and we collect training data using the sampling time step $dt = 0.1$. The energy associated with the training data is maintained to be constant to within 10^{-6}.

In general, the weights and biases of the adaptable HNN determined by the SGD method depend on the training data set. To reduce the prediction error, an ensemble of HNNs can be used [152]. Concretely, for each value of α, we generate 20 different sets of data for training, leading to an ensemble of 20 HNNs. The parameter setting for training is listed in Tab. 3.2.

After the training, all the weights and biases in Eq. (3.1) are determined. The Hamiltonian and its derivatives for each network in the HNN ensemble can be evaluated for any input, leading to the average derivative values. To characterize the prediction accuracy for different values of the bifurcation parameter, we use the root-mean square error (RMSE) that can be calculated from the difference between the HNN predicted and the true orbits. For $\alpha = 0$, the motion is integrable so the predicted orbit is always close to some real orbit, leading to exceedingly small errors. In this case, we take advantage of one feature of HNNs that it directly yields the Hamiltonian function, from which the potential function can be calculated. It is thus convenient to use the relative error between the predicted potential function and the

true one to characterize the HNN performance, which is defined as

$$\langle \Delta V \rangle \equiv \frac{\overline{|V_{\rm pred} - V_{\rm real}|}}{\overline{V_{\rm real}}}, \tag{3.4}$$

where the average is taken in the region of $V_{\rm real} < 1/6$. The predicted potential profile is given by $V_{\rm pred} = H_{\rm pred} - C$, where $C = \min(H_{\rm pred})$ so that the minimum value of $V_{\rm pred}$ is zero. Note that the average in Eq. (3.4) is calculated from an integral in a 2D domain in the physical space, for which the boundary of the domain needs to be specified. A natural choice of the criterion to set the boundary would be $V(x, y) < E_{\rm max} = 1/6$, but occasionally the predicted orbit will diverge. Numerically, there are different ways to overcome this difficulty. For example, if the boundary is set according to the criterion: $\max(V_{\rm pred}, V_{\rm real}) < 1/6$, then almost all orbits are bounded, rendering calculable the error $\langle \Delta V \rangle$.

We demonstrate that HNN can be used to reconstruct the Hamiltonian of the target system. Consider phase-space points for $H(\alpha, q_1, q_2, p_1, p_2) < 1/6$ and expand the Hamiltonian about the origin using the Taylor series:

$$H_{\rm pred}(\alpha) = \sum_{i_1, i_2, i_3, i_4} \beta_{i_1 i_2 i_3 i_4}(\alpha) q_1^{i_1} q_2^{i_2} p_1^{i_3} p_2^{i_4}, \tag{3.5}$$

where β's are the expansion coefficients, and the sum is taken according to $0 \leq \text{sum}(i_1, i_2, i_3, i_4) \leq 3$, which contains in total 35 terms. Comparing with true Hamiltonian Eq. (3.3), only six terms are non-zero.

We train the HNN at four values of the parameter: $\alpha \in \{0.2, 0.4, 0.6, 0.8\}$. For each α value, we choose seven random initial conditions with their energies below the threshold. Figure 3.3(a) shows the relative error in predicting the potential function for $\alpha \in (0, 1)$. The interval in α can be divided into two parts: the shaded region $\alpha \in [0.2, 0.8]$ that contains the four values of α used in training, and the blank regions on both side of the shaded region. In the shaded region, the relative error

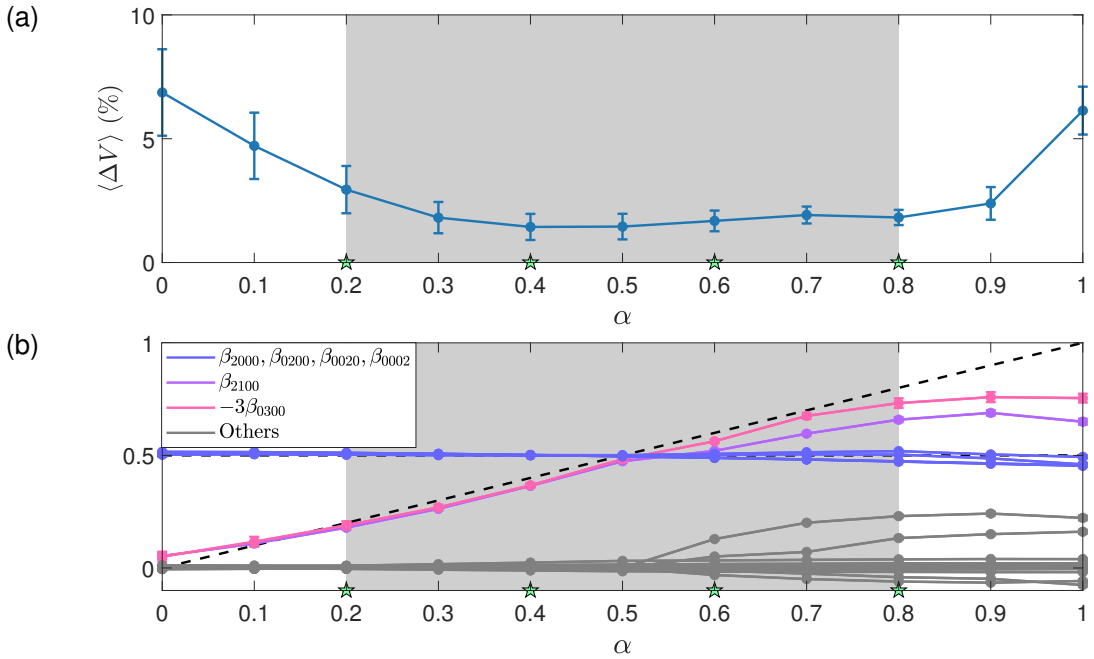

Figure 3.3: (a) Relative error in predicting the potential function of the Hénon-Heiles system. Shown is the error versus the bifurcation parameter α in the unit interval. The HNN is trained with time series data from four values of α: $\alpha \in \{0.2, 0.4, 0.6, 0.8\}$ (the four green pentagons). In the shaded region that contains these four values of α, the relative error is less than 2%, demonstrating the adaptability of the HNN in predicting the target Hamiltonian system for parameter values not in the training set. The adaptability extends even outside the shaded region but with larger error (still within 8% though). (b) Coefficients of the Taylor expansion for H_{pred} versus the bifurcation parameter α, where β_{2000}, β_{0200}, β_{0020} and β_{0002} correspond to the first four square terms in Eq. (3.3) whose true value is $1/2$, and β_{2100} and β_{0300} correspond to the two cubic terms that are proportional to α. Other terms in the expansion do not appear in the original Hamiltonian, among which the first two largest ones are β_{3000} and β_{1200} that correspond to other cubic potential terms.

is less than 2%, but the error increases away from the shaded region. Figure 3.3(b) shows the expansion coefficients for the predicted Hamiltonian. Comparing with the terms in the real Hamiltonian, our HNN predicts accurately the linear terms. For the nonlinear terms, the HNN reproduces the behavior with the variation in the bifurcation parameter α, where the errors are small in the shaded region in Fig. 3.3(b) but relatively large outside the region.

To examine the adaptability of our parameter-cognizant HNN in more detail, we

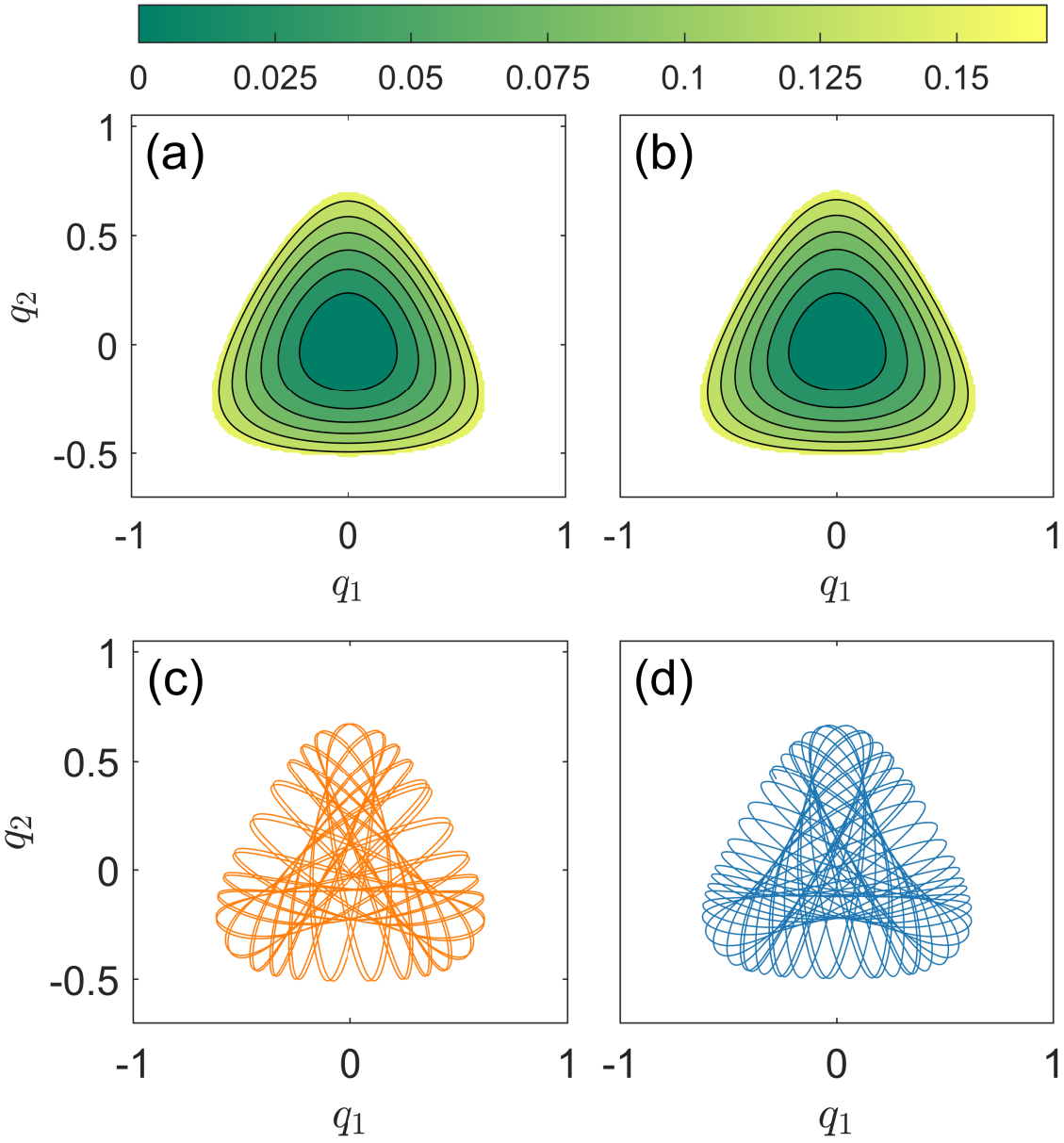

Figure 3.4: Testing the adaptability of HNN for parameter values in between two training points. (a,b) True and predicted contour maps of the potential function for $\alpha = 0.7$, respectively, where the latter is obtained by extracting the Hamiltonian at different positions with constant momentum input with normalization. (c,d) True and predicted orbits from the initial condition $[q_1, q_2, p_1, p_2] = [0, 0, 1/\sqrt{6}, 1/\sqrt{6}]$, respectively, which are quasiperiodic.

take, for example, $\alpha = 0.7$ in between the two training points $\alpha = 0.6$ and $\alpha = 0.8$, for which the vast majority of the orbits with energy $E = 1/6$ are quasiperiodic, as

can be seen from Fig. 3.2(b). Figures 3.4(a) and 3.4(b) show the true and predicted potential functions for $E < 1/6$, respectively, which are essentially indistinguishable. Figures 3.4(c) and 3.4(d) show some representative true and predicted orbits starting from the same initial condition, which agree with each other qualitatively but differ in detail. Particularly worth emphasizing is the fact that for the predicted quasiperiodic orbit, the energy can be maintained at a constant value. In fact, we have tested the method of reservoir computing [222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238] for predicting the orbit and find that, while it typically yields a more accurate orbit in short time (e.g., a few cycles), in the long run the energy is not conserved and the prediction error becomes large. Overall, since the testing bifurcation parameter value $\alpha = 0.7$ is sandwiched between two training points, our parameter-cognizant HNN exhibits a strong adaptability.

For $\alpha = 1$, with energy $E = 1/6$, most of the orbits are chaotic, where the portion of the KAM islands in the phase space becomes relatively insignificant, as shown in Fig. 3.2(d). In the case, the contour map of the true potential function has a triangular shape, as shown in Fig. 3.5(a). The predicted potential contour map is shown in Fig. 3.5(b), which agrees reasonably well with the true one. Figures 3.5(c) and 3.5(d) show a true and the predicted chaotic orbits from the same initial condition. While their details are different, the HNN predicts correctly that the orbit is chaotic. In fact, as will be shown in Sec. 3.3.3, for the two quantities characterizing the statistical behavior of the orbit, e.g, the maximum Lyapunov exponent and the alignment index, the predicted orbit yields the same results as those from the true orbit. In general, the closer the testing parameter value is to one of the training points, the higher the prediction accuracy.

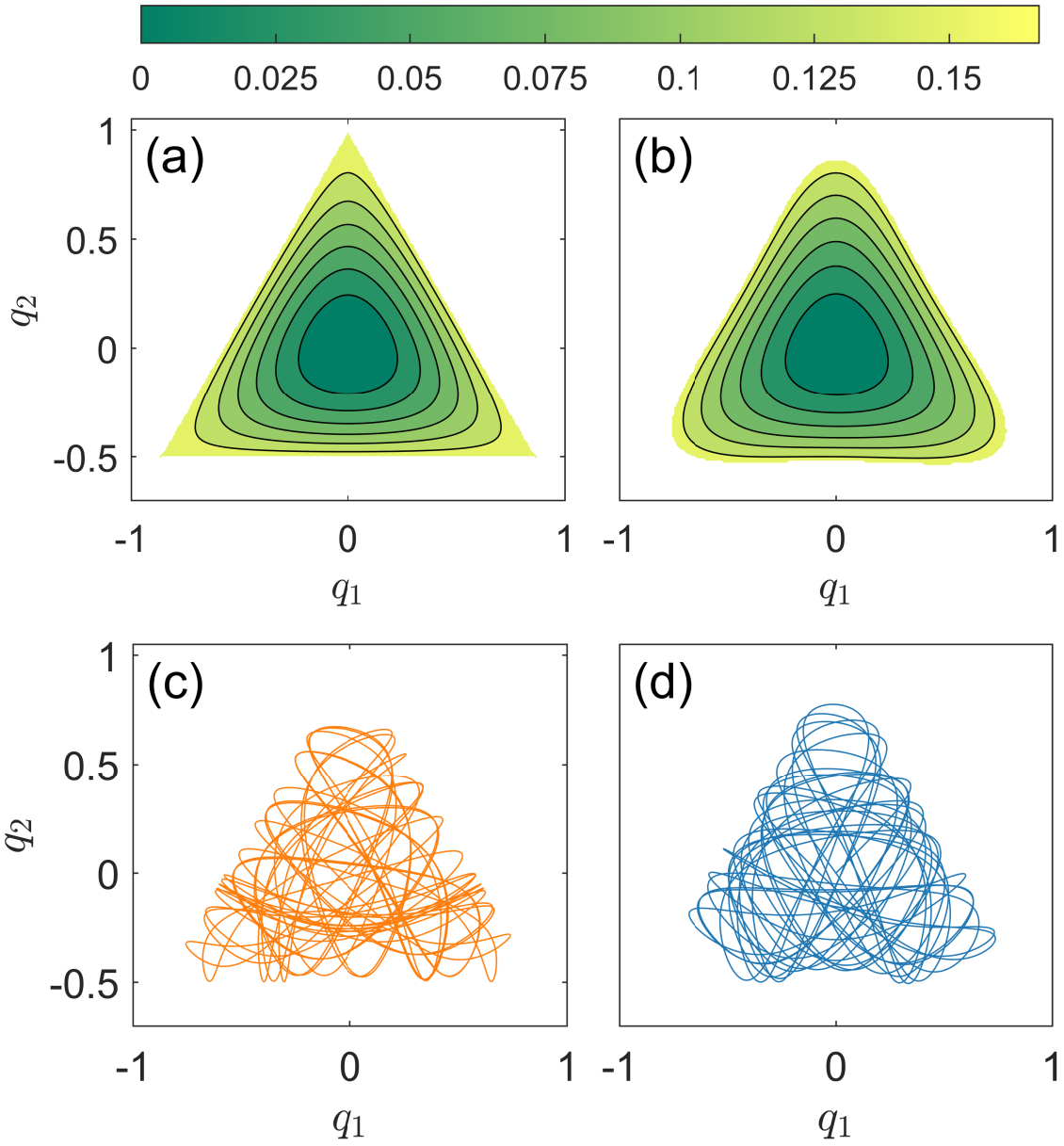

Figure 3.5: Testing the adaptability of HNN for parameter values outside the training interval. (a,b) True and predicted contour maps of the potential function for $\alpha = 1$. (c,d) True and predicted orbits from the same initial condition $[q_1, q_2, p_1, p_2] = [0, 0, 1/\sqrt{6}, 1/\sqrt{6}]$, which differ in detail but are both chaotic.

3.3.3 Adaptable Prediction of a Hamiltonian System

In a typical Hamiltonian system, the route of transition to ergodicity as a nonlinearity parameter increases is as follows [263]. In the weak nonlinear regime, the

system is integrable, where the motions are quasiperiodic and occur on tori generated by different initial conditions, as illustrated in Fig. 3.2(a) for the Hénon-Heiles system. As the nonlinearity parameter α increases, chaotic seas of various sizes emerge, leading to a mixed phase space, as exemplified in Fig. 3.2(b). In the regime of strong nonlinearity, e.g., $\alpha = 1$, most of the phase space constitutes chaotic seas with only a small fraction still occupied by KAM islands, as shown in Fig. 3.2(c) and (d). Here we provide strong evidence for the adaptability of our parameter-cognizant HNN by demonstrating that it can accurately predict the transition scenario, with training conducted based on time series from only a handful values of the nonlinearity parameter.

Distinct from dissipative systems in which random initial conditions in the basin of attraction of an attractor (periodic or chaotic) lead to trajectories that all end up in the same attractor, in Hamiltonian systems different initial conditions typically lead to different dynamically invariant sets. Because of this feature of Hamiltonian systems, to investigate the transition scenario, computations from initial conditions in the whole phase space leading to a statistical assessment and characterization of the resulting orbits are necessary. We focus on two statistical quantities: the largest Lyapunov exponent and the minimum alignment index, where the former characterizes the exponential separation rate of infinitesimally close trajectories and the latter measures the relative "closeness" of two arbitrary vectors along the trajectory (e.g., whether they become parallel, anti-parallel, or neither) [264]. For a chaotic trajectory, an infinitesimal vector stretches or contracts exponentially along the unstable or the stable direction, respectively. As a result, a random vector will approach the unstable direction along the trajectory and two random vectors will align with each other quickly. In particular, given two initial vectors $\mathbf{u}_1^0$ and $\mathbf{u}_2^0$, after i time steps,

they become $\mathbf{u}_1^i$ and $\mathbf{u}_2^i$, respectively. The minimum alignment index is defined as

$$\gamma^i \equiv \min(\|\mathbf{u}_1^i + \mathbf{u}_2^i\|, \|\mathbf{u}_1^i - \mathbf{u}_2^i\|). \tag{3.6}$$

When chaos sets in, the value of γ^i will quickly approach zero with time.

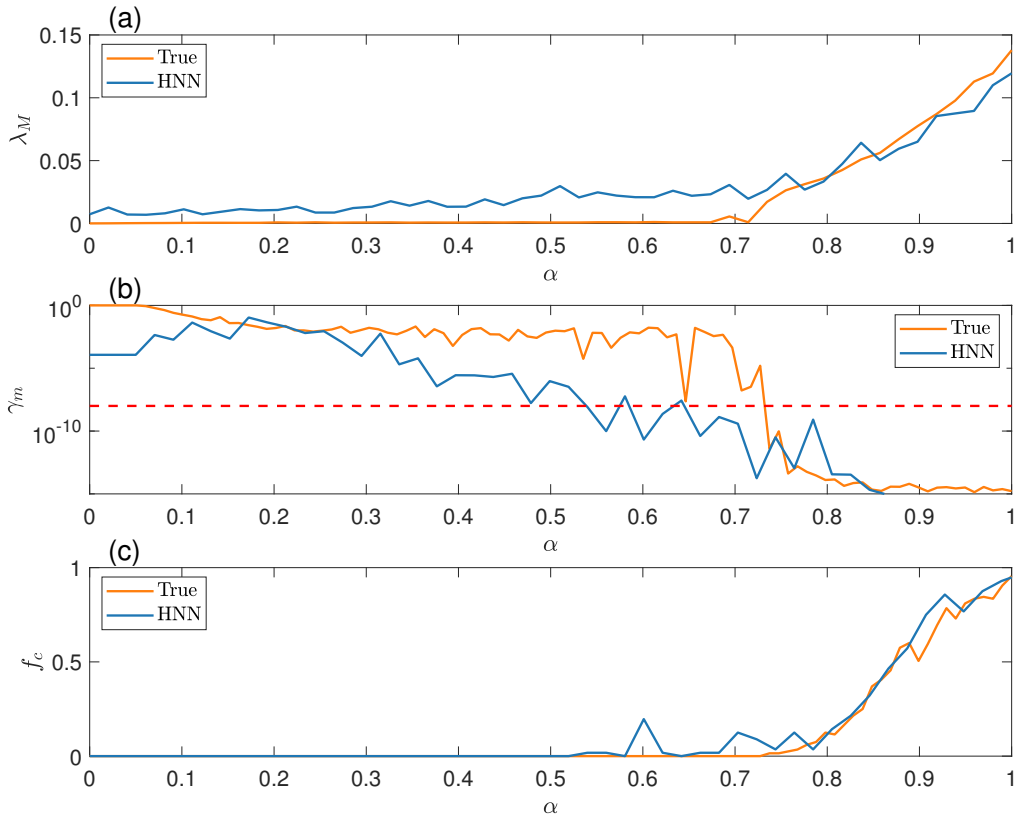

Figure 3.6: Test of adaptability of parameter-cognizant HNN in predicting transition to chaos in the Hénon-Heiles system. (a-c) The ensemble maximum Lyapunov exponent λ_M, the ensemble minimum alignment index γ_m together with the threshold 10^{-8}, and the fraction f_c of chaos in the phase space versus the nonlinearity parameter α, respectively. Transition to chaos occurs about $\alpha \leq 0.7$. The orange and blue colors correspond to the true and HNN predicted results, respectively. There is a reasonable agreement between the predicted and true behaviors, attesting to the adaptable predictive power of the HNN.

For a properly trained HNN with its weights and biases determined, the output contains the predicted Hamiltonian whose partial derivatives with respect to the coordinate and momentum vectors can be calculated directly based on the architecture of the neural network. These partial derivatives constitute the velocity field of the

underlying dynamical system, whose Jacobian matrix can then be determined, from which the machine predicted Lyapunov exponents and the alignment index can then be calculated (see Appendix 3.6). The true values of the Lyapunov exponents and the minimum alignment index can be calculated directly from the original Hamiltonian (3.3) of the target system.

In our calculation, we take 100 equally spaced values of the bifurcation parameter in the unit interval: $\alpha \in [0, 1]$. For each α value, we choose 200 random initial conditions and calculate, for each initial condition, the values of the largest Lyapunov exponent and the minimum alignment index. A trajectory is deemed chaotic [265] if the largest exponent is positive and the minimum alignment index is less than 10^{-8}. We denote the maximum Lyapunov exponent and the minimum alignment index from the ensemble of 200 trajectories as λ_M and γ_m, respectively, which are functions of α. Another quantity of interest is the fraction of chaotic trajectories, denoted as f_c, which also depends on α. The triplet of characterizing quantities, λ_M, γ_m, and f_c, can be calculated from the HNN and from the original Hamiltonian as a function of α. A comparison can then be made to assess the adaptable power of prediction of our parameter-cognizant HNN.

Figures 3.6(a-c) show the machine predicted and true values of λ_M, γ_m, and f_c versus α, respectively, for particle energy $E = 1/6$. It can be seen that chaos arises for $\alpha \leq 0.7$, at which λ_M becomes positive, γ_m decreases to 10^{-8}, and f_c begins to increase from zero. In Fig. 3.6(a), the true value of λ_M for $0 \leq \alpha < 0.7$ is essentially zero, but the HNN predicted λ_M is slightly positive. The remarkable feature is that both types of λ_M value begins to increase appreciably for $\alpha > 0.7$. In fact, there is a reasonable agreement between the true and predicted behavior of λ_M. Similar features are present in the behaviors of γ_m and f_c versus α, as shown in Figs. 3.6(b) and 3.6(c), respectively. These results are strong evidence that our parameter-cognizant HNN

is capable of adaptable prediction of distinct dynamical behaviors in Hamiltonian systems.

3.4 Issues Pertinent to Adaptability of Hamiltonian Neural Networks

We address the adaptability of HNNs by asking the following three questions. First, can the adaptability of HNNs be enhanced by increasing the number of training values of the bifurcation parameter? Second, can adaptability be achieved with multiple parameter channels? Third, does adaptability hold for different target Hamiltonian systems?

3.4.1 Effect of Number of Training Parameter Values

So far, we have used four distinct values of the bifurcation parameter to train our parameter-cognizant HNN. We now investigate if the adaptability can be enhanced by increasing the number of training parameter values. Here by "enhancement" we mean a reduction in the overall errors of predicting the Hamiltonian in a parameter interval that contains values not in the training set. To test this, we conduct the following numerical experiment. We choose $N \geq 3$ training parameter values and, for each parameter value, we train the HNN M times using an ensemble of time series collected from M energy values below the threshold (10 time series from 10 random initial conditions with energy below the threshold). To make the comparison meaningful, we choose the values of M and N such that NM is approximately constant. In particular, for Simulation #1, we set $N = 3$: $\alpha = 0.25$, 0.5 and 0.75, and $M = 9$. For Simulation #2, we choose $N = 4$: $\alpha = 0.2$, 0.4, 0.6, and 0.8, and $M = 7$. For Simulation #3, we have $N = 5$: $\alpha = 0.1$, 0.3, 0.5, 0.7, and 0.9, and $M = 5$. For each simulation, we calculate the ensemble error $\langle \Delta V \rangle$ in predicting the potential function as defined in Eq. (3.4) for $0 \leq \alpha \leq 1$. The results are shown in Fig. 3.7. It

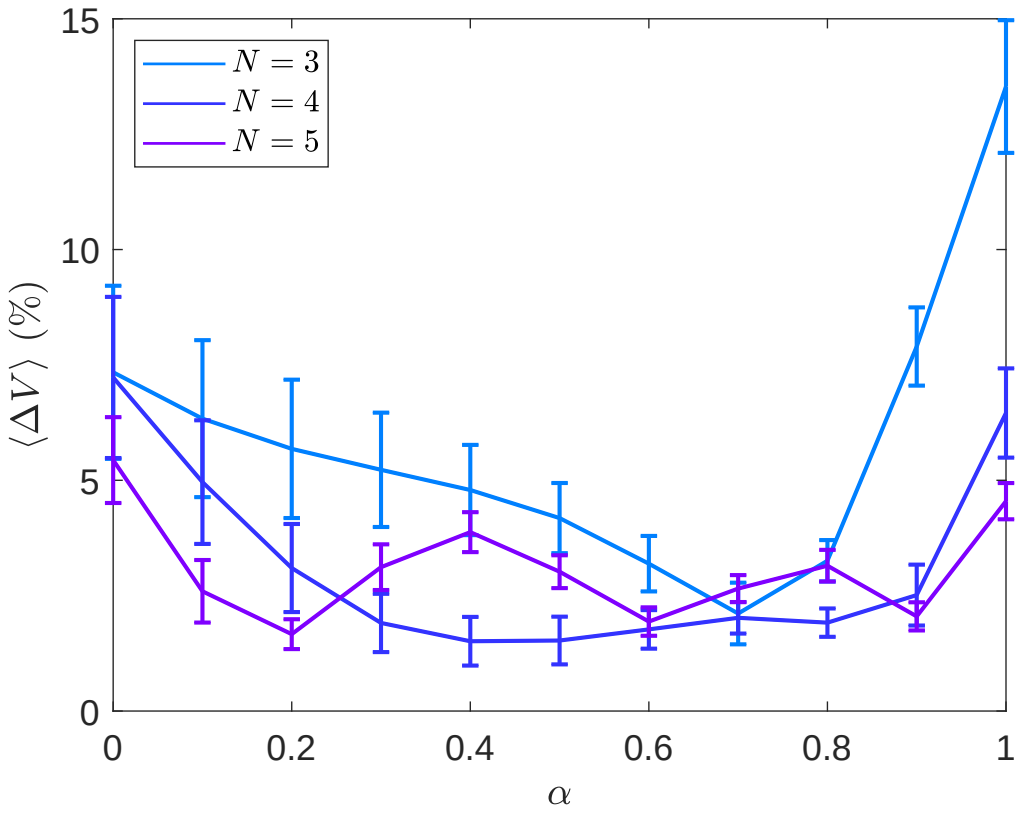

Figure 3.7: Effect of increasing the number N of training parameter values on the adaptable prediction error. Shown are the ensemble errors $\langle \Delta V \rangle$ in predicting the potential function for the three simulation settings explained in the text. Increasing N beyond four does not lead to a significant reduction in the errors, indicating that the HNN has already acquired the needed adaptability with training at four different values of the bifurcation parameter.

can be seen that the errors for $N = 3$ are generally larger than those for $N > 3$, but the errors for the two cases ($N = 4$ and 5) are approximately the same, indicating that increasing N above four will not lead to a significant reduction in the errors of adaptable prediction. That is, by training with multiple time series from four values of the bifurcation parameter, the HNN has already acquired the necessary adaptability for predicting the system behavior at other nearby parameter values.

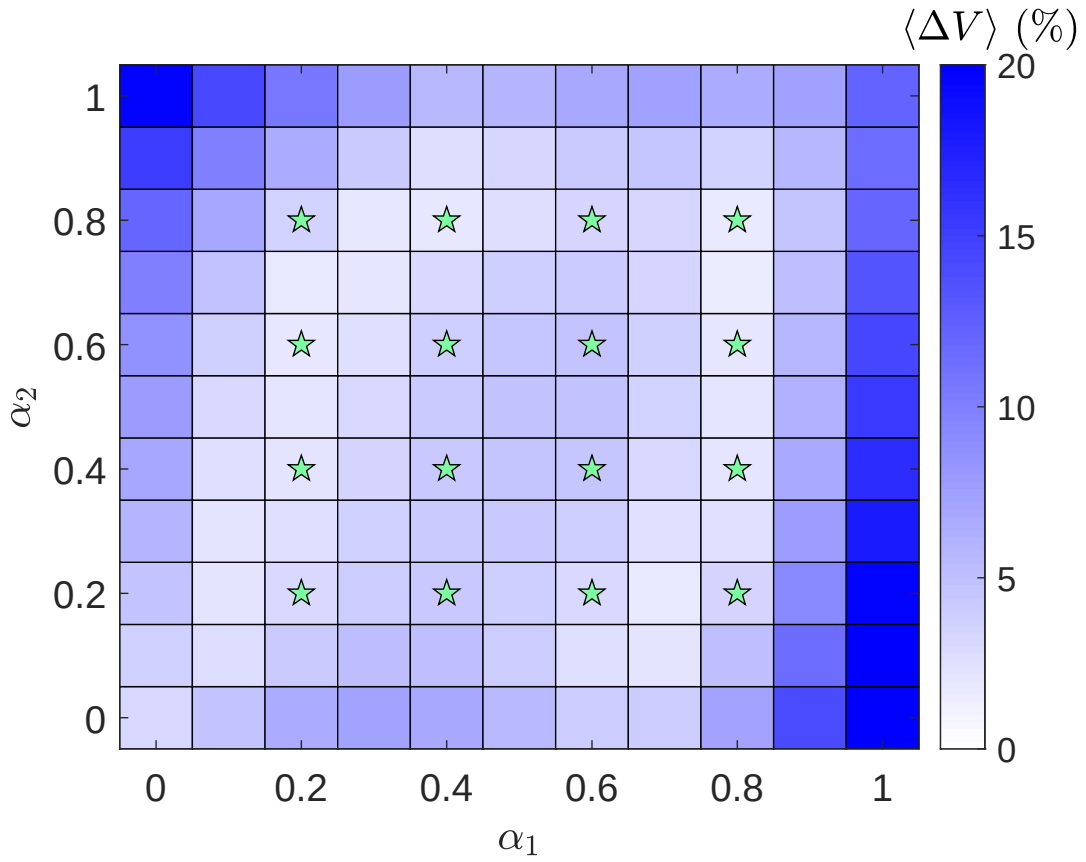

Figure 3.8: Prediction performance of an HNN with two input parameter channels. The target Hamiltonian system is given by the asymmetric Hénon-Heiles system as defined by Eq. (3.7). Shown is the color-coded ensemble prediction error $\langle \Delta V \rangle$ in the (α_1, α_2) plane. Training is conducted at the 16 points indicated by the green pentagons. The prediction error is small ($< 5\%$) in the central region $(\alpha_1, \alpha_2) \in [0.2, 0.8]$.

3.4.2 HNNs with Two Parameter Channels

We construct parameter-cognizant HNNs with more than one parameter channel. For this purpose, we modify the Hénon-Heiles Hamiltonian Eq. (3.3) to

$$H = \frac{1}{2}\left(p_1^2 + p_2^2\right) + \frac{1}{2}\left(q_1^2 + q_2^2\right) + \alpha_1 q_1^2 q_2 - \frac{\alpha_2}{3} q_2^3, \tag{3.7}$$

where α_1 and α_2 are two independent bifurcation parameters, requiring two independent parameter input channels to the HNN. The energy threshold for bounded motions can be evaluated numerically. We conduct training for a number of com-

binations of α_1 and α_2 values: $\alpha_1, \alpha_2 \in \{0.2, 0.4, 0.6, 0.8\}$. The training data are generated as follows: for each parameter pair, we choose five energy values below the threshold and, for each energy value, a single time series is collected. After the training is done, we predict the potential function for $\alpha_1, \alpha_2 \in [0, 1]$ with the interval 0.1 in each direction of parameter variation. Figure 3.8 shows the color-coded ensemble prediction error $\langle \Delta V \rangle$ in the (α_1, α_2) plane. For some combinations of α_1 and α_2 with a relatively large difference in their values, the threshold energy is less than 1/6. For such cases, the integration domain in Eq. (3.4) is modified accordingly based on the threshold value. It can be seen that, in the parameter region $(\alpha_1, \alpha_2) \in (0.2, 0.8)$, the prediction error is about 5%, while the errors outside the region tend to increase. At the two off-diagonal corners, the errors are the largest, due to the strong asymmetry in the potential profile. Figure 3.8 demonstrates that HNNs with two parameter channels can be trained to be adaptable for prediction.

3.4.3 HNNs for a Diatomic Molecule System

We consider a different target Hamiltonian system, a system defined by the one-dimensional Morse potential that describes the interaction between diatomic molecule [266]. This system was previously used in the study of chaotic scattering [267, 268]. The Hamiltonian is given by

$$H = \frac{p_x^2}{2} + V(x) \equiv \frac{p_x^2}{2} + [1 - \exp(-a(x - x_0))]^2 - 1, \tag{3.8}$$

where the potential function $V(x)$ has a minimum value at $x = x_0$ with $V(x_0) = -1$ and $V(x \to \infty) \to 0$. Taking the minimum potential value as the reference point for energy E, all orbits are bounded for $E < 1$. We set $x_0 = 1$ and choose a as the bifurcation parameter. The training data are generated from four different values of a: $a = 0.5$, 1.0, 2.0, and 4.0 where, for each training parameter value, an ensemble of

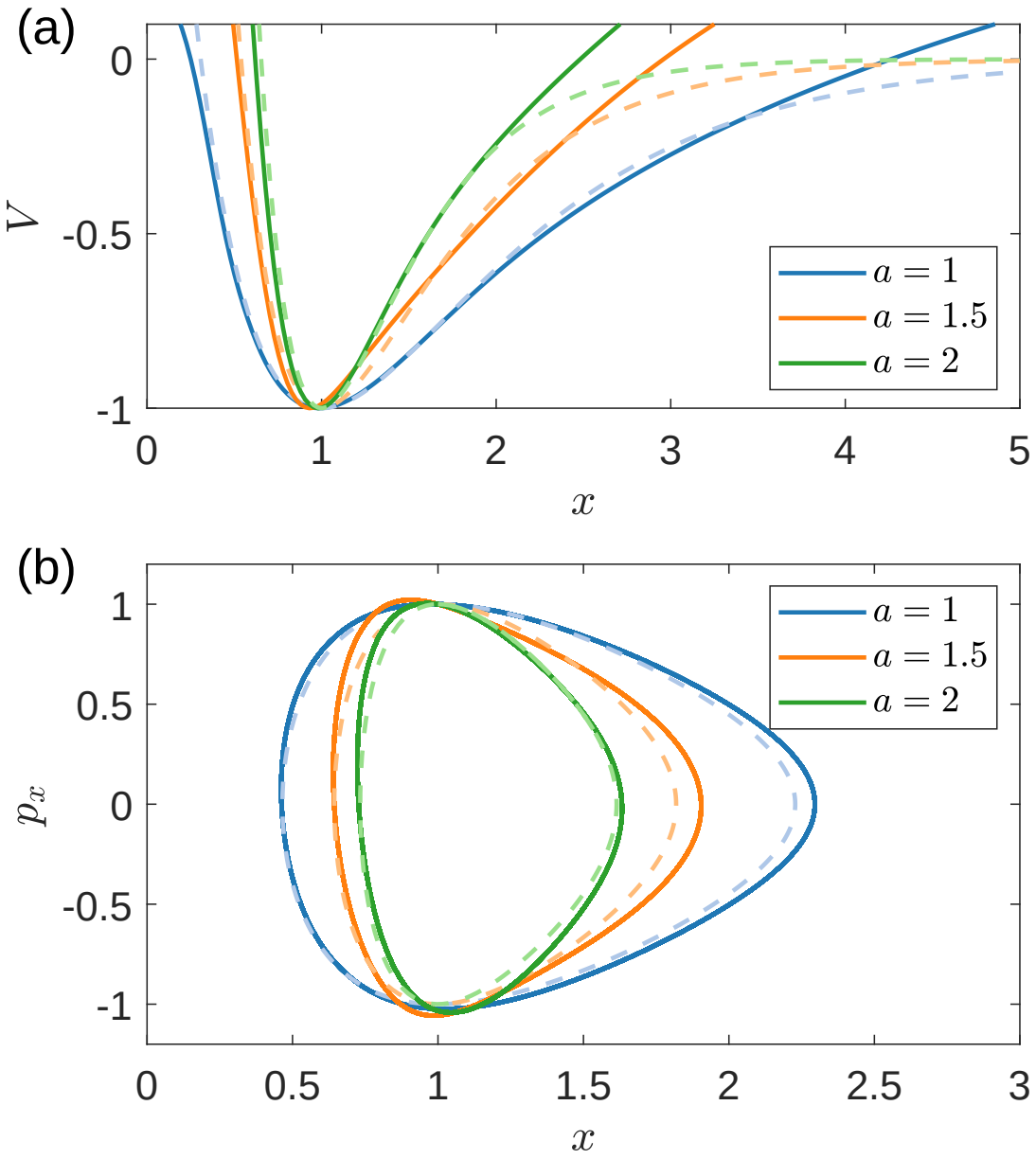

Figure 3.9: Parameter-cognizant HNN trained for the one-dimensional Morse system. (a) Shown are the predicted potential profiles for $a = 1.0$, 1.5, and 2.0 (solid curves), together with the corresponding true profiles (dashed curves). The predicted potential function for $a = 1.5$ is not the interpolation of those for $a = 1.0$ and $a = 2.0$, attesting to the adaptable predictive power of the HNN. (b) True and predicted orbits in phase space from $[x, p_x] = [1, 1]$.

five values of energy is used, resulting an ensemble of 20 independent time series. The time span for each time series is $0 \leq t \leq 100$ with the sampling time step $\Delta t = 0.1$.

Figure 3.9 shows the predicted potential profile for $a = 1.5$, together with those for the two training points $a = 1$ and $a = 2$. The result is accurate for x around the minimum potential point, but large errors arise when the position is far away from the minimum point. A plausible reason is that, for large values of x, the potential varies slowly, resulting in small changes in the momentum. As a result, the corresponding portions of the time series exhibit less variation, leading to large prediction errors by the HNN. The trained HNN has apparently gained certain adaptability, as the prediction result for $a = 1.5$ is not the interpolation of those for $a = 1.0$ and $a = 2.0$.

3.5 Discussion

Developing adaptable machine learning in general has broad applications to critical problems of current interest. For example, a problem of paramount importance is to predict how a system may behave in the future when some key parameters of the system may have drifted, based on information that is available at the present. As an example, suppose an ecosystem is currently in a normal state. Due to the environmental deterioration, some of its parameters such as the carrying capacity and/or the species decay rates will have drifted in the future. Is it possible to predict if the system will collapse when certain amount of parameter drift has occurred, when the system equations are not known and the only available information is time series data that can be measured *prior to but including the present*? Adaptable machine learning offers a possible solution. For example, it has been demonstrated recently [152] that incorporating a parameter-cognizant mechanism into reservoir computing machines enables prediction of possible critical transition and system collapse in the future for any given amount of parameter drift. However, the state-of-the-art reservoir computing schemes under intensive current research [222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238] do not taken into account physi-

cal constraints such as energy conservation, so they are suitable but for dissipative dynamical systems.

Combining the laws of physics and traditional machine learning has the potential to significantly enhance the performance and predictive power of neural networks. It has been demonstrated recently that enforcing the Hamilton's equations of motion in the traditional feed-forward neural networks can lead to improvement in the prediction accuracy for Hamiltonian systems in both integrable and chaotic regimes [153, 154, 155, 156, 157]. In these studies, training and prediction are conducted for the same set of parameter values of the target Hamiltonian system, so the underlying Hamiltonian neural networks are not adaptable in the sense that they are not capable of predicting the dynamical behavior of the system at a different parameter setting.

Do adaptable HNNs that we have developed have any practical significance? From the point of view of making predictions of the future states of Hamiltonian systems subject to parameter drifting, the answer is perhaps no. The main reason is that HNNs require all coordinate and momentum time series. For example, one may be interested in predicting whether a complicated many-body astrophysical system may lose its stability and become mostly chaotic in the future, where the only available information is the position and momentum measurements prior to or at the present when the system is still in a mostly integrable regime. As the laws of physics for this system are known, the data required for training is not a lesser burden than knowing the Hamiltonian itself. Nonetheless, our work generates insights into the working of HNNs, as follows.

Our parameter-cognizant, adaptable HNNs have a parameter input channel to the standard multilayer network with the loss function stipulated by the Hamilton's equations of motion, and are capable of successful prediction of transition to chaos in Hamiltonian systems. In particular, through training with coordinate and momentum

time series from four different values of the bifurcation (nonlinearity) parameter, the machine gains adaptability as evidenced by its successful prediction of the dynamical behavior of the target system in an entire parameter interval containing the training parameter values. That is, the benefits of training are that the HNN has learned not only the dynamical "climate" of the target Hamiltonian system but also how the "climate" changes with the bifurcation parameter. Machine learning can thus be viewed as a process by which the neural network self-adjusts its dynamical evolution rules to incorporate those of the target system.

When systematically varying values of the bifurcation parameter are fed into the HNN, it can predict the transition to chaos from a mostly integrable regime, as determined by the ensemble maximum Lyapunov exponent and minimum alignment index as well as the fraction of chaos as a function of the bifurcation parameter. For a single parameter channel, the adaptable predictive power is achieved insofar as the training parameter set contains at least three or four distinct values. For an HNN with duplex parameter channels, the size of the required training parameter set should be at least four by four. Adaptable prediction has also been accomplished for a different Hamiltonian system defined by the Morse potential function. We expect the principle of designing parameter-cognizant HNNs and the training method devised in this paper to hold for general Hamiltonian systems.

One issue is the dependence of the energy surface on the bifurcation parameter. As the parameter changes continuously, the energy surface will evolve accordingly. If we intend to predict the system dynamics for some specific value of the bifurcation parameter for a fixed energy value, the training data sets should contain time series collected from a larger energy value to cover the pertinent phase space region at the desired energy value.

It should also be noted that, using HNNs to predict the transition from integrable

dynamics to chaos in the Hénon-Heiles system was first reported [156], which relied on using energy E as the control parameter for a fixed value of the nonlinearity parameter (e.g., $\alpha = 1$). Here we have studied the transition using α as the bifurcation parameter for a fixed energy value (e.g., $E = 1/6$). The two routes are equivalent because the Hénon-Heiles system Eq. (3.3) possesses a three-fold symmetry in the configuration space. Such an equivalence also arises in systems whose potential function contains nonlinear square terms, e.g., the classical ϕ^4 or FPU model [269, 270]. However, for the two-parameters Hamiltonian Eq. (3.7) studied in this paper, the three-fold symmetry is broken, destroying the equivalence between varying the nonlinearity parameter and energy. In fact, for Hamiltonian systems such as the Morse and double-pendulum systems, the equivalence does not hold either [266, 268, 271]. Our adaptable HNN does not rely on any such equivalence, and can be effective in predicting the transition to chaos in any type of Hamiltonian systems.

3.6 Algorithm for Calculating the Lyapunov Exponent and Alignment Index of Hamiltonian Neural Networks

Given a dynamical system $d\mathbf{x}/dt = \mathbf{f}(\mathbf{x})$, the Jacobi matrix is given by $\mathcal{J} = \partial\mathbf{f}/\partial\mathbf{x}$. For a Hamiltonian system, the dynamical variables are $\mathbf{x} \equiv [\mathbf{q}, \mathbf{p}]^T$ and

$$\mathbf{f}(\mathbf{q}, \mathbf{p}) = \left[\frac{\partial H}{\partial \mathbf{p}}, -\frac{\partial H}{\partial \mathbf{q}}\right]. \tag{3.9}$$

For an HNN, in principle, the Hamiltonian H is given by a sequence of operations of the neural network with the weights and biases in Eq. (3.1) determined by training. An alternative but efficient approach to calculating the Jacobian matrix $\mathcal{J}$ is the finite-difference method. In particular, for a given initial condition, we generate an orbit of N points with time interval dt and calculate $\mathcal{J}$ at each time step. Let the sequence of Jacobian matrices be denoted as $\mathcal{J}(t_0), \mathcal{J}(t_1), \cdots, \mathcal{J}(t_N)$, and let $\mathcal{Y}(t_0)$

be the identity matrix $\mathcal{I}$. If the phase space of the target Hamiltonian system is D-dimensional ($D = 4$ for the Hénon-Heiles system), there are D Lyapunov exponents. Let $\boldsymbol{\lambda}$ be the vector of the D Lyapunov exponents: $\boldsymbol{\lambda} \equiv (\lambda_1, \lambda_2, \ldots, \lambda_D)^T$ and set the initial values of the exponents to be zero: $\boldsymbol{\lambda}(t_0) = (0, 0, \ldots, 0)^T$. After N steps, we have

$$\mathcal{Y}(t_N) = \mathcal{J}(t_N) \cdot \mathcal{Y}(t_{N-1}). \tag{3.10}$$

Carrying out the QR decomposition of the matrix $\mathcal{Y}(t_N)$ with the resulting matrices denoted as $\mathcal{Q}$ and $\mathcal{R}$, we have

$$\lambda_j(t_N) = \lambda_j(t_{N-1}) + \log|R_{jj}|, \tag{3.11}$$

$$\mathcal{Y}(t_N) = \mathcal{Q}, \tag{3.12}$$

where R_{jj} is the jth diagonal element of the matrix $\mathcal{R}$. The Lyapunov exponents are given by

$$\lambda_j = \lim_{N\to\infty} \frac{\lambda_j(t_N)}{N dt},\ j = 1, \ldots, D. \tag{3.13}$$

The maximum Lyapunov exponent is $\lambda_M = \max_j(\lambda_j)$.

To calculate the alignment index, we introduce a matrix $\mathcal{M}$ and set it to be the identity matrix at the initial time: $\mathcal{M}(t_0) = \mathcal{I}$. Let $\mathbf{u}_1(t_0) = [1, 0, 0, 0]^T$ and $\mathbf{u}_2(t_0) = [0, 1, 0, 0]^T$ be two linearly independent vectors at the initial time. After N steps, we have

$$\begin{aligned} \mathcal{M}(t_N) &= (\mathcal{I} + \mathcal{J}(t_N)dt) \cdot \mathcal{M}(t_{N-1}), \\ \mathbf{u}_{1,2}(t_N) &= \mathcal{M}(t_N) \cdot \mathbf{u}_{1,2}(t_0). \end{aligned} \tag{3.14}$$

Normalizing the vectors $\mathbf{u}_{1,2}(t_N)$ by their respective magnitude to have the unit length, we obtain the minimum alignment index as

$$\gamma_m = \lim_{N\to\infty} \min(\|\mathbf{u}_1(t_N) + \mathbf{u}_2(t_N)\|, \|\mathbf{u}_1(t_N) - \mathbf{u}_2(t_N)\|). \tag{3.15}$$

3.7 Appendix

Following are the relevant publication on the topics presented in this chapter

C.-D. Han, B. Glaz, M. Haile, and Y.-C. Lai*, "Adaptable Hamiltonian neural networks," Physical Review Research **3**, 023156, (2021).

Chapter 4

TOMOGRAPHY OF TIME-DEPENDENT QUANTUM HAMILTONIANS WITH MACHINE LEARNING

Interacting quantum Hamiltonians are fundamental to quantum computing. Data-based tomography of time-independent quantum Hamiltonians has been achieved, but an open challenge is to ascertain the structures of time-dependent quantum Hamiltonians using time series measurements taken locally from a small subset of the spins. Physically, the dynamical evolution of a spin system under time-dependent driving or perturbation is described by the Heisenberg equation of motion. Motivated by this basic fact, we articulate a physics-enhanced machine-learning framework whose core is Heisenberg neural networks. In particular, we develop a deep learning algorithm according to some physics-motivated loss function based on the Heisenberg equation, which "forces" the neural network to follow the quantum evolution of the spin variables. We demonstrate that, from local measurements, not only can the local Hamiltonian be recovered, but the Hamiltonian reflecting the interacting structure of the whole system can also be faithfully reconstructed. We test our Heisenberg neural machine on spin systems of a variety of structures. In the extreme case in which measurements are taken from only one spin, the achieved tomography fidelity values can reach about 90%. The developed machine-learning framework is applicable to any time-dependent systems whose quantum dynamical evolution is governed by the Heisenberg equation of motion.

4.1 Introduction

Quantum computation based on spin is a fundamental component of quantum information science and technology [272]. Recently, it has been demonstrated that manipulating 50 spins can generate the computational capability beyond any kind of classical computers, leads to quantum supremacy [273, 274]. From a network point of view, the information exchange between any pair of spins can be regarded as a link between the two spins. When the interactions associated with all spin pairs are taken into account, the end result is effectively a network, giving rise to the subfield called qubit or spin system [275, 276, 277]. Experimentally, a multispin coupling system can be realized using cavity quantum electrodynamics [278], ion traps [279], or superconducting qubits [273, 274].

There are two types of quantum Hamiltonians: time-dependent or time-dependent. In the former case, the system can be decomposed into a sequence of quantum gates [280], resembling a classical circuit structure. Since the Hamiltonian is constant over time, this effectively leads to quantum adiabatic computing systems, where quantum computing algorithms can be performed on the ground states [281]. For time-dependent Hamiltonians, both the spin coupling and an external, time-varying field [282] or an output control signal [283] are present. It was argued that in time-varying spins systems, the problem of switch off can be mitigated and the computation speed can be enhanced [284]. In general, an external field can serve to increase the computational capacity even for relatively simple spins system structures [283, 285]. However, it is challenging to analyze and realize time-dependent control of spins systems. Recently, the idea of embedding a time-dependent Hamiltonians into a time independent one was studied, but the generality or universal applicability of this approach remains unknown [286, 287]. In addition, deep neural-

network based techniques to quantum state classification and reconstruction have been proposed [288, 289].

In recent years, the inverse problem of spins systems has attracted a great deal of attention. The basic question is, given only limited access to the system, i.e., only part of the system can be measured, can the global structure of the spins system be determined? Previous efforts focused on monitoring the Hamiltonian as a function of time through the eigenstate realization algorithm (ERA) [290, 291], compressive sensing [292, 293], machine learning [286, 287], or quantum quench [294]. The basic idea is to find the coefficients of the power series terms constituting the Hamiltonian in some basis. However, when applying to time-dependent quantum Hamiltonians, these approaches are limited to systems of a single spin or those with a special type of external field [293, 295, 296]. The general difficulty is that the functional form of the time signal generates an optimization problem in infinite dimensions [297], rendering inapplicable any optimization algorithm designed for finding a finite number of parameters. For the methods based on the eigenstates, difficulties arise when the system changes too fast with time [290, 291, 296]. Recently, in [298] the author verified different methods to reconstruct time-independent quantum Hamiltonians. A method based on unsupervised learning was proposed in [299]

In this paper, we solve a general class of inverse problems in spins systems by exploiting machine learning [158]. Our work was partly inspired by the recent work on the classical Hamiltonian Neural Networks (HNNs) [160, 153, 155, 156], where the basic idea is to introduce a physics-based, customized loss function to "force" the dynamical evolution of the system to follow that stipulated by the classical Hamilton's equations. However, the existing HNNs are not directly applicable to quantum Hamiltonians, thereby requiring new approaches. Our idea originates from the basic physical consideration that the dynamical evolution of spins systems is governed

by the Heisenberg equations of motion. We are thus motivated to develop a class of Heisenberg Neural Networks (HENNs) by exploiting deep learning to predict the Hamiltonian but under the constraint of the Heisenberg equations of motion. The HENNs have the advantage of guaranteeing that the underlying quantum evolution possesses the Hermitian structure. Our assumptions for the quantum Hamiltonian are listed as follows. (i) Quantum Hamiltonian varies continuously with time, a situation that can be expected to hold in experiments in general. (ii) The measured spins are coupled with the rest of the spins directly or indirectly. (iii) There are no Pauli terms in the Hamiltonian commute with each other.

Our main results are the following. The original time-dependent quantum Hamiltonians whose structure is to be determined based on incomplete local measurements and the HENN that is an artificial neural network for predicting the Hamiltonian of the original system. We treat the dynamical evolution of the original system in terms of both Schrödinger and Heisenberg pictures. We demonstrate that, with only local measurements, the local Hamiltonian can be recovered, similar to the solution of the local Hamiltonian learning problem [296]. In particular, defining the tomography fidelity as the ratio between the correctly predicted links and the total possible number of links in the underlying spins systems, we find that the fidelity can reach 90% even when the number of spins measured is much smaller than the system size. In fact, the predicted Hamiltonian contains the global information about the coupling profile of the original quantum Hamiltonians. We note that the problem of network reconstruction or tomography has been well studied in classical nonlinear dynamical systems [300, 301, 302, 303, 304], and there was also a study of structure identification for time-independent quantum Hamiltonians [305]. Our work goes beyond the relevant literature in that we have successfully articulated and validated a general machine learning framework of quantum tomography for time-dependent quantum

Hamiltonians.

Before presenting the details of our work, we offer four remarks.

Remark 1. It is worth noting that the problem of initialization in quantum Hamiltonians tomography is challenging and has not been solved. The current state of the field is that, even when measurement is done locally, initial state preparation for the whole system is still required. This is the case for algorithms such as ERA [290, 291]. Even for the problem of local Hamiltonian recovery, the requirement is that initially the system be in an eigenstate [296]. To our knowledge, the only case where initial state preparation is not required is when the spins system structure is already known [306]. The new feature of our work is a machine-learning approach to tomography of *time-dependent* quantum Hamiltonians without any prior knowledge about the network structure. As in the existing studies treating time-independent quantum Hamiltonians [290, 291, 292, 293, 286, 287], a large number of given initial states is required. However, state preparation methods are nowadays available [307, 308, 309].

Remark 2. One of the relating topics is called quantum process tomography (QPT) [310, 311]. The idea is detecting the system information based on observations. The procedure are describing as follows: Preparing an initial quantum state, perform observation and then repeat for different initial states until it's enough for matrix inversion. This method needs 16^n number of observations, which is exponential growth as we increasing the number of spins n, but the measurement can be performed locally to reconstruct the Hamiltonian. The idea is widely used for time independent Hamiltonian. Some experiments for 2 or 3 spins systems are shown in [312, 313, 314].

Remark 3. Recently, another paper used a Recurrent Neural Network (RNN) to learn the time dependent and time independent quantum Hamiltonian from single spin measurement [315]. Comparing with our results, in [315] the structure for the

spins system assumed to be known and in our paper we try to detect the coupling structure. Another difference is that in [315] they use recurrent neural network, comparing with HENN the training is time consuming.

In Sec. 4.2, we describe the HENN learning framework. In Sec. 4.3, we test our machine-learning method using a variety of time-dependent quantum Hamiltonians, which include networks with short- or long-range interactions and two quantum gates. In Appendix 4.5, we present analytic results with HENNs for one- and three-spin systems.

4.2 Time-dependent Quantum Hamiltonians and Heisenberg Neural Networks

Consider a system of spins coupled by an external field. The Hamiltonian is

$$H(t) = h^{(1)} + f(t)h^{(2)}, \tag{4.1}$$

where $h^{(1,2)}$ represent the time-independent Hamiltonian and $f(t)$ is a continuous function of time that is the result of the application of a time-dependent electrical or magnetic field. Suppose the system is initially in the state $|\psi_0\rangle$ at $t = 0$. In the Schrödinger picture where the state evolves with time but the operators are time-invariant, at time t the expectation value of an operator A is given by $\langle A\rangle_t$. In the Heisenberg picture where the state does not change with time but the operators do, an operator evolves according to the Heisenberg equation

$$\frac{dA^H}{dt} = i[H^H(t), A^H(t)], \tag{4.2}$$

where the superscript H specifies that the corresponding matrix is in the Heisenberg picture, $H^H(t) = U_{t,0}^\dagger H(t) U_{t,0}$, and $U_{t,0}$ and $H(t)$ do not commute with each other due to the time dependence. Once $H^H(t)$ is known, the corresponding Hamiltonian in the Schrödinger picture $H(t)$ can be determined. The goal is to solve the Heisenberg equation based the observations of A.

Since Eq. (4.2) is a set of linear equations in $H^H(t)$, for any time t the equations are solvable if the number of non-equivalent equations is no less than the number of unknown elements. That is, the non-commutative operators at all times, $A^H(t)$, are required to be known. This is a key difference from time-independent systems, where $H^H(t) = H$ so one operator at any time, $A^H(t)$, can be used as the non-commutative operator. In this case, once the observations (e.g., time series) are sufficient, the Hamiltonian can be fully determined [290, 291].

The number of independent elements in the Hamiltonian matrix provides another angle to appreciate the complexity of the problem. In particular, for a system with n spins, at a specific time t, the Hamiltonian in Eq. (4.1) can be represented by a Hermitian matrix in terms of the $N = 2^n$ linearly independent states. There are altogether $N^2 = 4^n$ bases for an $N \times N$ Hamiltonian matrix that is Hermitian. To fully solve Eq. (4.2) will thus need all the 4^n measurements at a given time. For example, for a two-spin system, there are four linearly independent states, so in principle 16 observations are needed. These observations can be generated by the direct product of the Pauli matrices $S_{\alpha,\beta} = \sigma_\alpha^1 \otimes \sigma_\beta^2$, where α and β are integers ranging from 0 to 3, which correspond to the identity and the three Pauli matrices $\sigma_x, \sigma_y, \sigma_z$. For these 16 matrices, one is an identity that commutes with all other matrices. Consequently, we need at least $2^N - 1 = 4^n - 1$ measurements to fully determine the Hamiltonian.

When the quantum states of all spins can be measured, it is straightforward to obtain the Hamiltonian matrix through Eq. (4.2). A difficult situation is that only a small fraction of the spins in the network, e.g., one or two, can be measured. Experimentally, measurements or observations have been reported for one-, two-, and three-spin systems [316, 317, 318]. The pertinent question is, what can we learn about the whole network system when only local measurements in some subspace of the full

space are available? To address this question, we decompose the Hamiltonian as

$$H = \underbrace{H_o + H_i}_{H_{o'}} + H_h, \tag{4.3}$$

where H_o is the subspace Hamiltonian for the observed spins, H_i represents the interaction between the observed and the nonobservable spins, and H_h is Hamiltonian for the nonobservable spins. Let $H_{o'} \equiv H_o + H_i$, which is the sub-Hamiltonian that contains information directly related to the observed spins.

Take a three-spin system as an example, as shown in Fig. 4.1(a). The three spins are labeled with $1, 2, 3$ and we assume that only the first spin can be measured. The subspace of H_o contains 3 bases corresponding to the Pauli matrices for the first spin $H_o = \sigma^1$, and H_i contains two-body interactions between the first spin and the second or the third spin and the three-body interaction:

$$H_i = \sigma^1(\sigma^2 + \sigma^3 + \sigma^2\sigma^3).$$

The subspace Hamiltonian H_h contains the Pauli matrices for the second and third spins as well as the two-body interaction between them: $H_h = \sigma^2 + \sigma^3 + \sigma^2\sigma^3$. Overall, this is a three-node spin system, where the nodal interactions represent different links. For each node, three independent quantities (the three Pauli matrices) are needed to characterize the spin polarization, which generate different combinations of coupling. A unique feature of spin systems, which is not present in classical complex networks, is that one link can couple more than two nodes.

The decomposition scheme in (4.3) is valid only in the Schrödinger picture. In the Heisenberg picture, different subspace are mixed together in the time evolution, so all the subspace must be simultaneously determined. For limited observations, the solutions of the Heisenberg equation can be nonunique. To overcome this difficulty, we exploit machine learning to predict the Hamiltonian. Inspired by the work of HNN

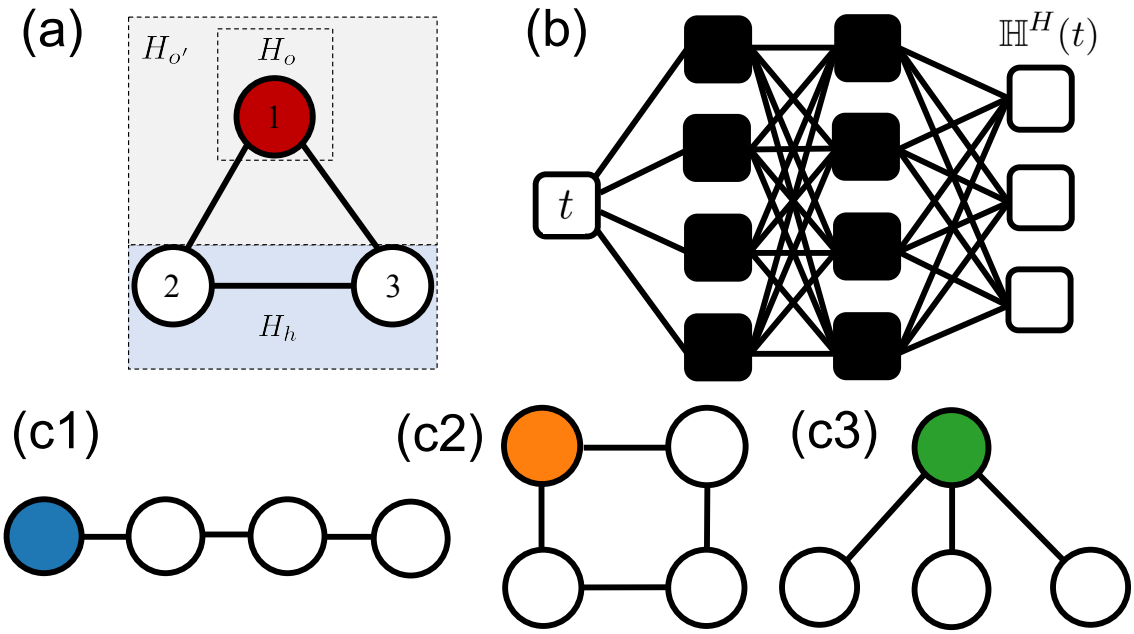

Figure 4.1: Three- and four- spins systems and the machine learning architecture. (a) Schematic illustration of local observation of the Hamiltonian in a three-spin system. Say only one node or spin (the red one) can be observed, which corresponds to the Hamiltonian H_o. The Hamiltonian for the nonobservable nodes and their interactions are labeled as H_h. (b) Machine learning (neural network) architecture, where the nodes in the hidden layers are represented by black squares, the input and output are denoted by open squares, and the various weighted links (solid line segments) connect the input to the output. The input is one-dimensional: it is simply the time variable t. The output constitutes the elements of the matrix $\mathbb{H}^H(t)$ whose size is determined by the size of the spin system. The weight associated with each link is calculated by the auto-gradient method to minimize the custom loss Eq. (4.4). (c1-c3) Possible architectures of a 4-spin system, where the filled circle represents the observable spin in the network. The networks in (c1-c3) have a chain, a cyclic, and a tree structure, respectively. For $n = 3$ and $n = 5$, structures similar to those in (c1-c3) exist.

whose loss function is based on the Hamilton's equations of motion [160, 153, 155, 156] for time-independent spin systems, we articulate a general class of HENNs that conform with the Heisenberg equations of motion with broad applicability to both time-dependent and time-independent spin systems.

Figure 4.1(b) shows our neural network architecture with two hidden layers. The input is one-dimensional: it is simply the time t. Each layer is a convolution of the preceding layer: $\mathbf{a}_l = \tau(\mathcal{W}_l \cdot \mathbf{a}_{l-1} + \mathbf{b}_l)$, where $\mathbf{a}_l$ is the state vector of the lth layer,

$\mathcal{W}$ is the weighted matrix connecting layers l and $l-1$, $\mathbf{b}_l$ is the bias vector of layer l, τ is a nonlinear activation function, e.g., $\tau = \tanh$. The matrix $\mathcal{W}$ and the bias vector $\mathbf{b}_l$ are to be determined through training based on spin measurements. The output is the Hamiltonian matrix in the Heisenberg picture. In our study, we use two hidden layers, each with 200 nodes. The neural network is built by Tensorflow and the Keras package [255]. We use the Stochastic Gradient Descent (SGD) and adaptive momentum (Adam) methods to determine the optimal weighted matrix $\mathcal{W}_l$ and the bias vector $\mathbf{b}_l$ by minimizing an appropriate loss function [319]. In particular, we define our loss function as the mean square error in the time derivatives of the observation as

$$\mathcal{L} = \sum_{\text{Observations}} \left| \langle \dot{A}(t) \rangle_{\text{real}} - \langle \dot{A}(t) \rangle_{\text{pred}} \right|^2, \tag{4.4}$$

where $\dot{A}_{\text{pred}}(t) = i[H^H(t), A(t)]$, and the matrix $H^H(t)$ is the output of the HENN. Once the time derivatives for some given observations are known, we input them to the loss function as the target to train the HENN and subsequently to predict the Hamiltonian. Due to incomplete measurement and finite optimization steps, the predicted Hamiltonian varies over different rounds of training. It is thus necessary to take the statistical average of the prediction and to calculate the variance.

The training data are time series of measurements. After the time series is known we can get the corresponding derivative for the loss function. Therefore, the total number of training data is the product of three thing, number of linear independent states, number of observations, number of time discrete points. In Table 4.1 we show the number of training data for each part. For example, for a chain structure with $n = 4$ and observing one spin, the number of training data is $300 \times 3 \times 100 = 90,000$. Since HENN are trained for a specific structure. Once the Hamiltonian change the neural network need to be retrained.

Table 4.1: List of training parameters for HENN

Description	Values
Number of initial states for $n = 3$, 4 and 5	$100, 300, 1100$
Number of observations for one,two spins	$3, 15$
Time discrete points	100

It can be shown that the sub-Hamiltonian $H_{o'}$ in Eq. (4.3) containing information directly related to the observed spins can be recovered [296]. For the subsystem not directly related to the observed spin, its Hamiltonian H_h cannot be fully recovered. However, we can show that, in the subspace of H_h, if the machine predicted coupling value between two nodes is smaller than some threshold, then it effectively indicates null coupling. This means that our HENN is capable of determining the coupling configuration for the quantum Hamiltonian based on if the predicted Hamiltonian matrix elements are zero or finite, providing a solution to the tomography problem for the whole system. In particular, the tomography contains two types of information: whether the spins are coupled and if so, how they are coupled. The first one is related to the spatial structure of the network, as exemplified in Fig. 4.1(a), where spin 1 is coupled to spin 2 and 3. The second type of information gives the the type of coupling among all possible coupling configurations determined by the spin polarization vector at each node.

The prediction phase of our HENN thus consists of the following steps.

First, for a given quantum spin system, we take measurement A from some part of the system and calculate the corresponding matrix elements $\mathbb{A}^H(t)$ based on the linear equation

$$\langle\psi_0|A^H(t)|\psi_0\rangle = \langle A\rangle_t.$$

To obtain the matrix elements, the number of linearly independent initial states must

be larger than the number of independent elements of the matrix. Specifically, for a spin system with n spins, at least 4^n linearly independent initial states are needed.

Second, we build up a neural network as in Fig. 4.1(b) with input time t and output as the matrix elements $\mathbb{H}^H(t)$. We train the network using the loss function defined in Eq. (4.4). After the HENN is properly trained, we evaluate the Hamiltonian for a given time series, and convert it into the Schrödinger picture. The coupling among the nodes can be obtained from the decomposition

$$\begin{aligned}\mathbb{H}(t) &= c_0(t)\mathbb{I} + \sum_{i,j} c_{ij}(t)\sigma_j^i, \\ &+ \sum_{i,j,k,m} c_{ijkm}(t)\sigma_j^i\sigma_m^k + \cdots \equiv \sum_i c_i(t)S_i, \end{aligned} \tag{4.5}$$

where S_i is the basis of the N-dimensional Hamiltonian matrix and $c_i(t)$ is the corresponding coupling coefficients at time t. We choose S_i to be the direct product of the Pauli matrices plus the identity matrix. The coefficients $c_i(t)$ determine the coupling configuration of the system.

Third, after obtaining the time series of the coupling coefficients, we take the time average for each basis $\bar{c}_i = \int_0^t |c_i(t)|dt$ and normalize them by their maximum value. We set some threshold: any value above which indicates an existent coupling between the corresponding spins.

To better illustrate our HENN based machine learning procedure, in Appendix 4.5, we present two explicit examples for HENN predicted Hamiltonian: a one-spin system and a three-spin system.

4.3 Results

We test the predictive power of the proposed HENNs for a number of spin systems. As noted, in a quantum spin system, the concept of links can be quite different from those in classical networks. In particular, one link is referred to as a specific way of

coupling in the underlying spin system. For a system with n spins, the total number of linearly independent states is 2^n. The total number of independent elements in the Hamiltonian matrix is 4^n, which is the total number of possible ways of coupling in the system. The links are generated by the direct products of the Pauli and identity matrices. The types of links include self-coupling, two-body interactions and long-range interactions. A quantity to characterize the machine-learning performance is the tomography fidelity, defined as the ratio between the number of correctly predicted links and the total possible number of links. Disregarding the identity matrix, we define the tomography fidelity as

$$F_t = \frac{4^n - 1 - (\#\text{ of missing links})}{4^n - 1}, \tag{4.6}$$

where the tomography is meaningful for $F_t > 50\%$. A more useful characterizing quantity is the success in identifying the structure of H_h, as this is proof that the method can not only yield the structure of the subsystem from which measurements are taken ($H_{o'}$), but also information about the complementary subsystem from which no observations are made (H_h), so that information about the whole system can be obtained. This alternative fidelity measure is defined as

$$F_{t'} = \frac{4^{n'} - 1 - (\#\text{ of missing links})}{4^{n'} - 1}. \tag{4.7}$$

where $n' = n - n_{\text{obs}}$ and n_{obs} is the number of spins from which observations are taken.

Another quantity is measuring the fidelity for local Hamiltonian $F_{o'}$ defined as

$$F_{o'} = 1 - \frac{\|H_{o'(\text{pred})} - H_{o'(\text{real})}\|}{\|H_{o'(\text{real})}\|}, \tag{4.8}$$

where $H_{o'(\text{real})}$ and $H_{o'(\text{pred})}$ are the real and predicted local Hamiltonian, respectively.

4.3.1 Tomography of Spin Systems Based on Two-body Interactions

The sub-Hamiltonians $h^{(1)}$ and $h^{(2)}$ in (4.1) of a spin system with two-body interactions are given by

$$\begin{aligned} h^{(1,2)} = & \sum_{i=1}^{n} \sum_{j=1}^{3} c_{ij}^{(1,2)} \sigma_j^i + \\ & \sum_{i=1}^{n} \sum_{j=i+1}^{n} \sum_{m=1}^{3} \sum_{l=1}^{3} w_{ij} c_{ijml}^{(1,2)} \sigma_m^i \sigma_l^j, \end{aligned} \tag{4.9}$$

where c's are random numbers between 0 and 1. The superscript of σ indicates the number of spins, which varies from 1 to n, the subscripts $1, 2, 3$ denote the x, y and z components of the spin, respectively, w_{ij} is the ijth element of the adjacency matrix as in a conventional, undirected network, where $w_{ij} = 1$ indicates there is coupling between spin i and spin j, otherwise $w_{ij} = 0$. The first term of h contains self-couplings, and the second term contains two-body couplings. Due to the exponential growth of the computational overload with the number of spins in the network, we limit out study to networks with $n \leq 5$ spins. The Hamiltonian (4.9) arises in a variety of physical situations such as the Heisenberg model or spin glass systems [320, 321]. The time dependence in the general Hamiltonian (4.1) is introduced into the network with the following "driving" function of time:

$$f(t) = \sin(\omega t + 2\pi\phi), \tag{4.10}$$

where ω and ϕ are random numbers whose values are taken between zero and one.

We test HENNs with the three structures shown in Figs. 4.1(a) and 4.1(c1-c3). The main difference among them lies in the degree of the observed node. For example, for the networks in Figs. 4.1(c1-c3), the degree of the observed spin is 9, 18, and 27, respectively. To generate the data, we choose 100, 300 and 1,100 random initial conditions for $n = 3$, 4 and 5. For each initial condition, we numerically integrate

the Heisenberg equation (4.2) for $0 < t < 5$, and extract from this time interval 100 equally spaced points as the measurement data. The calculated time series for a given initial state correspond to observations of σ_x, σ_y and σ_z for the specific local spin in the network from which measurements are taken. We take the time derivative defined in Eq. (4.4) as the loss function for training the HENN. Following the steps described in Sec. 4.2, we obtain the predicted interaction structure of the network. Comparing with the actual structure gives the tomography fidelity. Since the fidelity may vary for a different Hamiltonian, for each specific type of networks, we repeat this process 100 times.

Figure 4.2(a) shows the results of reconstructing the cyclic network of three spins in Fig. 4.1(a), where the degree of each node is 18 (excluding self-interactions) and there are 64 distinct links in the network. What is displayed is the average predicted coupling value $\bar{c}$ versus the link index, and the blue and red dots denote the existent and null links, respectively. The dashed horizontal line defined at 10% of the maximum coupling value can separate majority of the existent from majority of the null links. Figures 4.2(b) and 4.2(c) show the results from similar cyclic networks but with four and five spins, respectively, with the same legends as those in Fig. 4.2(a). These results indicate that, even when measurements are taken from only one spin, the coupling structure of the time-dependent Hamiltonian can be predicted by our HENN with a reasonably high accuracy.

A heuristic reason that the HENN is able to predict the structure of the spin system correctly from only local measurements is as follows. Recall that the input to the HENN is time, a continuous variable. The differential property of the neural network guarantees that the predicted Hamiltonian must be continuous and the time change for the predicted Hamiltonian must follow the Heisenberg equation as stipulated by the physically meaningful loss function. With these constrains, data from different

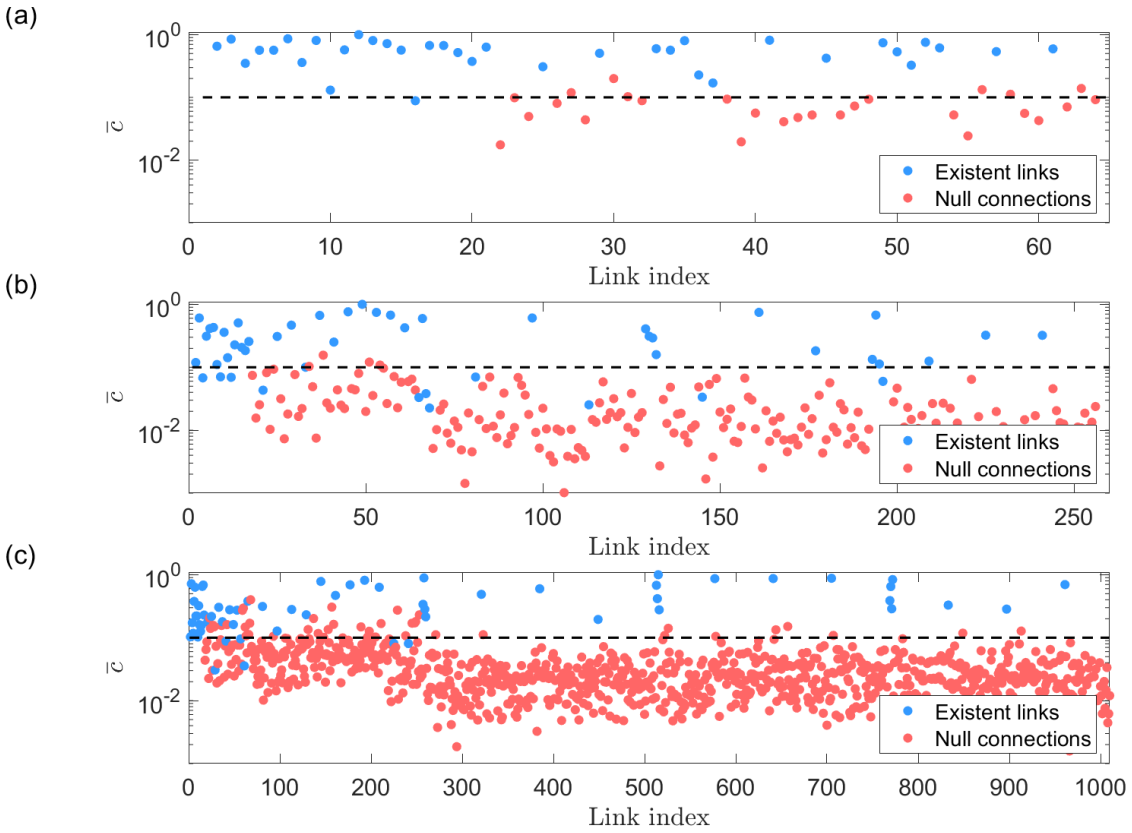

Figure 4.2: Tomography performance of HENN for cyclic networks of $n = 3$, 4, and 5 spins. The network structure is given in Fig. 4.1(a). (a) Reconstructed coupling value $\bar{c}$ versus the total number of possible links. The result is for one of the spin systems with the average tomography fidelity value from 100 random realizations. The abscissa represents the number of possible links for $n = 3$, where the basis ranges from $\sigma_0^1 \otimes \sigma_0^2 \otimes \sigma_1^3$ to $\sigma_3^1 \otimes \sigma_3^2 \otimes \sigma_3^3$, where the total number of possible links is $4^n - 1 = 63$ (with the identity matrix taken away). The blue dots represent the true, existent links, while nonexistent or null links are denoted as the red dots. The horizontal dashed line is taken at the 10% of the predicted maximum coupling value. (b,c) Results from $n = 4$ and 5, respectively, with the same legends. In all cases, the horizontal dashed line can serve as a threshold for separating majority of the existent links from majority of the null links, attesting to the ability of the machine-learning scheme to infer the whole network structure from local measurements only.

time will instill the correct physical relationships among the dynamical variables into the neural network. As a result, the difficulty of non-uniqueness of the solutions when solving the linear equations is overcome.

To further characterize the performance of HENNs for different network structures, we calculate the average fidelity measures for nine distinct networks that include the

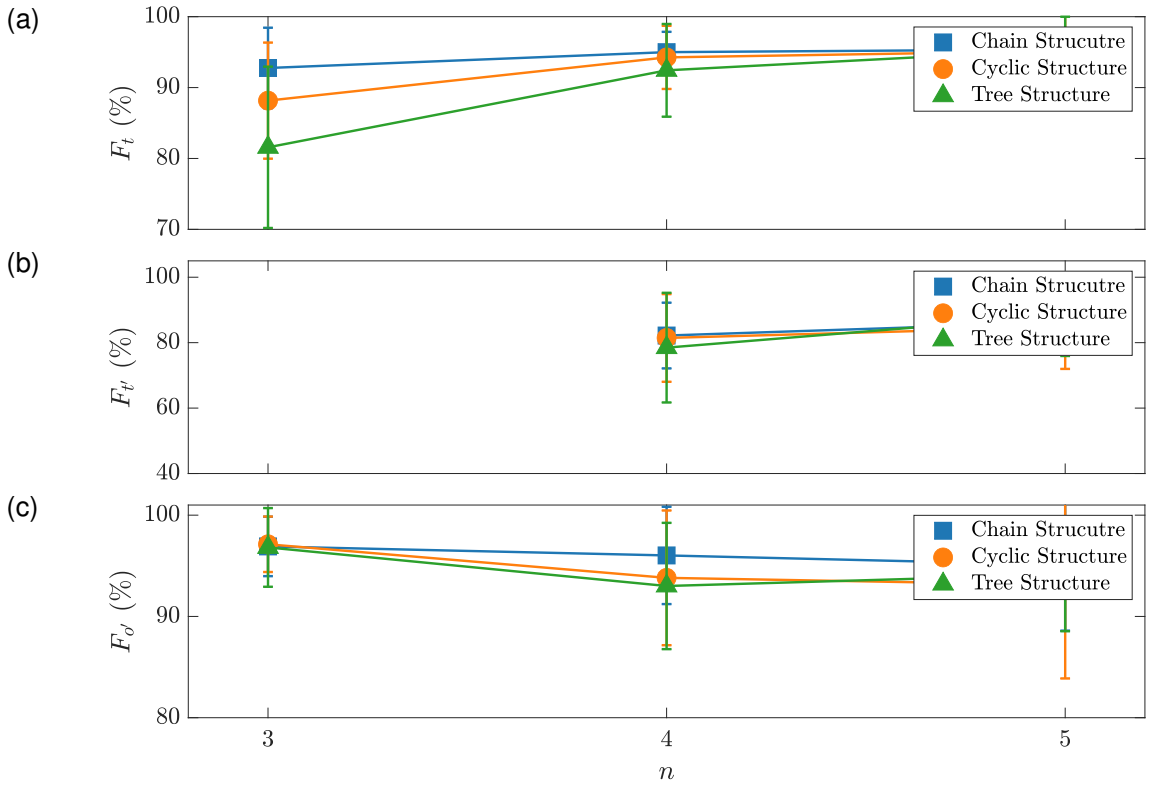

Figure 4.3: Tomography fidelity measure. (a) Average tomography fidelity F_t and the standard deviation for nine distinct networks of three types of structures in Figs. 4.1(c1-c3), respectively, each with $n = 3$, 4, and 5 spins. In all cases, local measurements from only one spin are taken. (b) Average alternative tomography fidelity $F_{t'}$ for $n = 4$ and 5, corresponding to $n' = 3$ and 4, respectively. (The results from $n = 3$ contain large statistical errors because of the the relatively small size of the system and thus are not shown so that the results for $n = 4$ and 5 can be seen clearly). In all cases, the fidelity values are above 80%, indicating the predictive power of HENN. (c) The accuracy for predicted Hamiltonians that is directly related to the observed spins $H_{o'}$. For $n = 3$, 4 and 5 spins, each contains $48, 192$ and 768 time series to compare.

chain, cyclic, and tree structures in Figs. 4.1(c1-c3), respectively, each with $n = 3$, 4, and 5 spins, as shown in Figs. 4.3(a), 4.3(b) and 4.3 (c) for the measures F_t, $F_{t'}$ and $F_{o'}$, respectively. For the network structures in Figs. 4.1(c1-c3), the degrees of the measurement spin are 9, 18, and 27, respectively. For a fixed number of spins in the network, the fidelity value decreases with the degree of the measurement spin. The reason are listed as follows: Firstly, a higher degree for the observed spins means the measurements time series is more complicated. Our calculation shows that the tomography for the Hamiltonian directly connected to the observed node $F_{o'}$ is relatively the same, but the predicted power for the hidden structure decreases. Another note is that for $n = 3$, cyclic structure and tree structure has the same degree. For $n = 3$ tree structure, we find HENN will assume there is coupling between two non observed spins but in reality there is not. Thus the tomography fidelity is small for tree structure than cyclic structure.

Note that the measure F_t is defined for the whole network, which takes into account not only the links between the measurement spin and the nonobservable spins, but also the links among the nonobservable spins, where the latter are characterized by the alternative tomography fidelity measure $F_{t'}$. Since this measure is purely for the nonobservable spins from which no measurements are taken, we expect its value to be lower than that of F_t, as shown in Fig. 4.3(b). In spite of the reduction in comparison with F_t, the values of $F_{t'}$ for $n = 4$ and $n = 5$ are still relatively high: approximately 80% and larger, attesting to the power of our HENN scheme to extract information from the nonobservable spins.

The accuracy for predicted Hamiltonians that is directly related to the observed spin is shown in Figure. 4.3 (c). Take four spins system as an example. There are 3 for the self-coupling, 27 two-body interactions, 81 three-body interactions and 81 four-body interactions. Thus we can get 192 time series to compare. In Figure. 4.3

(c) for different structures and different numbers of spins, $F_{o'}$ is larger than 90%. For spins system with tree structures, the fidelity is slightly lower than the other two structures, and as we increasing the system size, the variance becomes larger. However, there is no clear difference between size and structure effects.

Another issue is the effect for different driving functions. We choose three different driving functions $f(t) = t/5$, $f(t) = \sin(\pi t/5)$ and $f(t) = \exp[-(t-2.5)^2]$ at $t \in [0, 5]$ as an example. Each of the driving represent linear, sinusoidal and Gaussian pulse. Using the example for a cyclic structure with one spin as measurement shown in Fig. 4.1 (c2). The fidelity for tomography F_t, $F_{t'}$ and Hamiltonian $F_{o'}$ shown in Table. 4.2.

Firstly, these results indicate that for simple deriving functions $f(t) = t/5$ the fidelity is higher than others. Secondly, in Table. 4.2 the sinusoidal and Gaussian pulse contains one period. Comparing with Fig. 4.3 the fidelity is slightly higher. Since at Fig. 4.3 we randomly generated ω through 0 to 1 so the maximum contains two oscillation periods. It proves that for a more complex driving functions that has more oscillations, the errors can increase quickly.

Based on our setting, the measurement time series must start from $t = 0$, since the Hamiltonian we predict is in the Heisenberg picture. To get the coupling profile of the spins system, we need to change the Hamiltonian into the Schrödinger picture. If the time series at the beginning is missed, then even we can use the HENN to get the Hamiltonian in Heisenberg picture, change it to Schrödinger picture will be challenging. The spacing for the time series may not be the same. As long as it can maintain enough accuracy for the derivative.

Utilizing two-body coupling spin to evaluate the performance of the HENNs has certain limitations. In particular, for a given network structure, when the number of spins increases, the error appears to decrease, due mostly to the exponential growth

Table 4.2: Tomography fidelity under different driving functions

	$f(t) = t/5$	$f(t) = \sin\left(\frac{\pi}{5}t\right)$	$f(t) = \exp[-(t-2.5)^2]$
F_t	$97.4\% \pm 1.7\%$	$96.7\% \pm 1.7\%$	$96.3\% \pm 2.4\%$
$F_{t'}$	$90.7\% \pm 6.4\%$	$88.3\% \pm 6.5\%$	$87.9\% \pm 7.3\%$
$F_{o'}$	$98.0\% \pm 1.0\%$	$96.6\% \pm 1.3\%$	$95.5\% \pm 1.9\%$

in the total possible number of links in the network, which is an artifact. In some cases, the prediction results can be trivial as the system size increases. For example, if all the two-body couplings are null, then for $n = 3$, the tomography fidelity value will be about 40% because approximately 60% of the links are of the two-body type. Similarly, for $n = 4$ and 5, approximately 20% and 5% of the links are of the two-body type, leading to artificial fidelity values of about 80% and 95%, respectively. Comparing with the results in Fig. 4.3, for $n = 3$ and 4, the trivial prediction gives lower fidelity values, but the difference diminishes for $n = 5$. Consequently, based solely on two-body interactions, that the tomography fidelity increases with the system size is not synonymous to a better performance of the algorithm for larger systems. For accurate tomography of quantum spin systems, long-range interactions must be included.

4.3.2 Tomography of Quantum Spin Systems with Long-range Interactions

We consider the more general Hamiltonian that contains all short- and long-range interactions. Physical applications include the development of quantum gates such as the Toffoli or the Fredkin gate that requires three-body interaction [322], spin glass with infinite-range interactions [321], and quantum computing that requires high coherence [318, 323]. We decompose the Hamiltonian into two components, $h^{(1,2)}$, as

in Eq.(4.1), which are given by

$$h^{(1,2)} = \sum_{i_1,i_2,\cdots,i_n=0}^{3} r c^{(1,2)}_{i_1 i_2 \cdots i_n} \sigma^1_{i_1} \sigma^2_{i_2} \cdots \sigma^n_{i_n}, \tag{4.11}$$

where $c^{(1,2)}_{i_1 i_2 \cdots i_n}$ are random numbers between 0 and 1, and r takes on the values of one or zero with equal probabilities. The network comes into existence only for $r = 1$. The function $f(t)$ rendering the system time-dependent is chosen according to Eq. (4.10).

We consider systems with $n = 3$, 4 or 5 spins with 100, 300 and 1100 random initial conditions, respectively. Observing one spin leads to time series of σ_x, σ_y and σ_z from this spin. If two spins can be measured, we choose the observation variable to be $\sigma^1_\alpha \otimes \sigma^2_\beta$, where α and β are integers from zero to three, corresponding to the identity and the three Pauli matrices, respectively. Excluding the identity operation, we have 15 measured time series of 100 equally spacing points in the time interval $0 < t < 5$.

Following the procedure described in Sec. 4.2, we train the HENN to predict the coupling configurations of the spin systems with long-range interactions. Unlike the case where only two-body interactions are taken into account, here the links are chosen randomly: we consider all possible links and any specific link exists or does not exist with equal probabilities. Figure 4.4 shows the prediction performance for a network of $n = 4$ spins, where panels (a) and (b) correspond to the cases of measuring one and two spins, respectively. When only one spin is measured [Fig. 4.4(a)], most of the existent and nonexistent links can be distinguished by the 10% threshold line, yet there are still quite a few links that are on the "wrong" side. When two spins are measured, the prediction accuracy is higher as there are far fewer incorrectly predicted links. This is intuitively reasonable as measuring more spins is equivalent to imposing more constraints on the predicted Hamiltonian so as to improve the prediction accuracy.

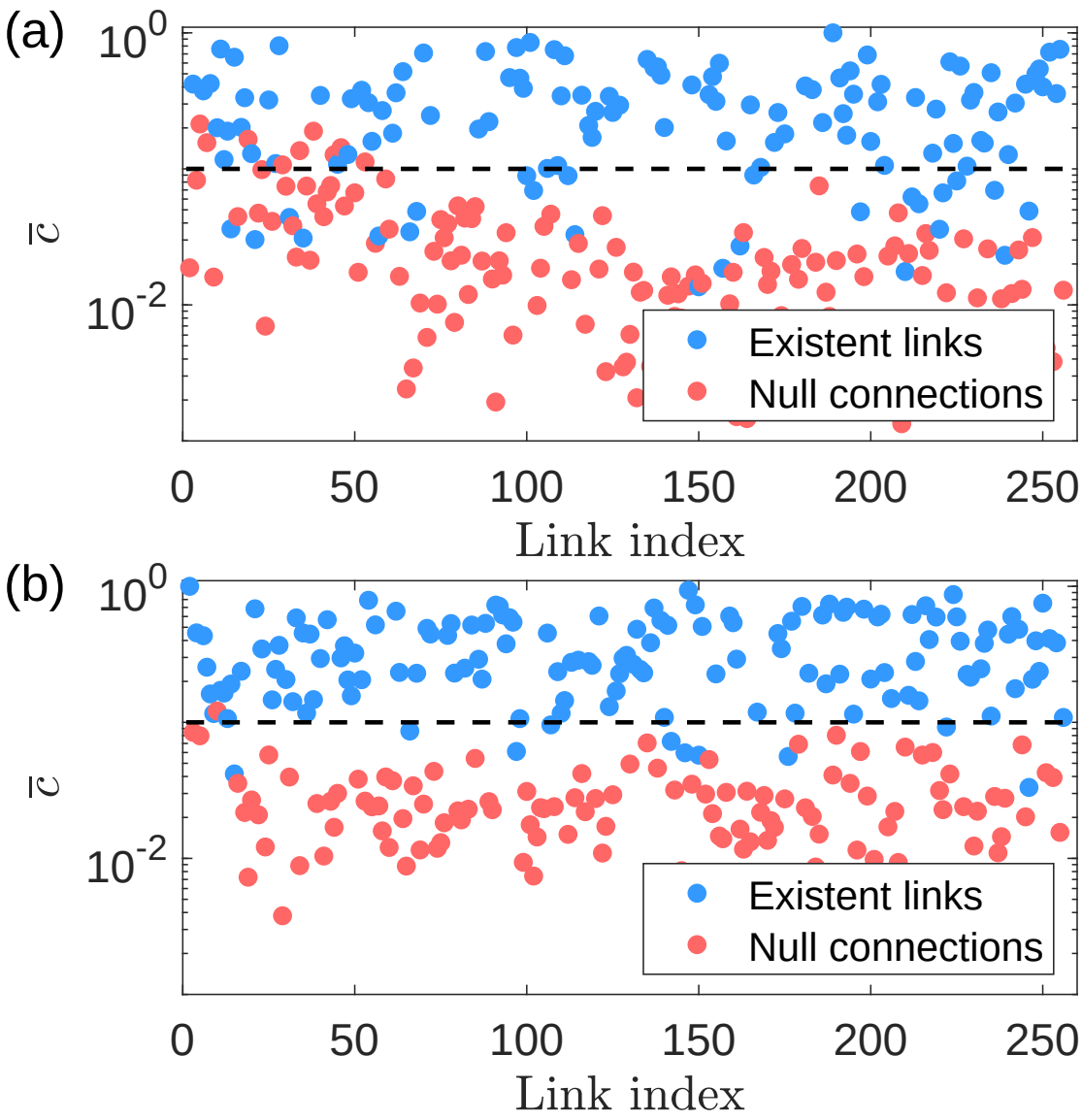

Figure 4.4: Prediction performance of HENN for a network of four spins with long-range interactions. (a) Reconstructed coupling coefficients versus based on time series measured from one spin only. The results are from one realization of the spin system with the fidelity value equal to the average fidelity value over 100 random realizations. The legends are the same as those in Fig. 4.2. (b) The corresponding results when two spins are observed with the time series as described in the text.

Figure 4.5 shows the fidelity measure of predicting random networks of $n = 3$, 4, and 5 spins, from observing one spin or two spins. As shown in Fig. 4.5(a), the fidelity value decreases as the number of spins increases. This is expected because, when observing a fixed number of spins, a larger system means more nonobservable

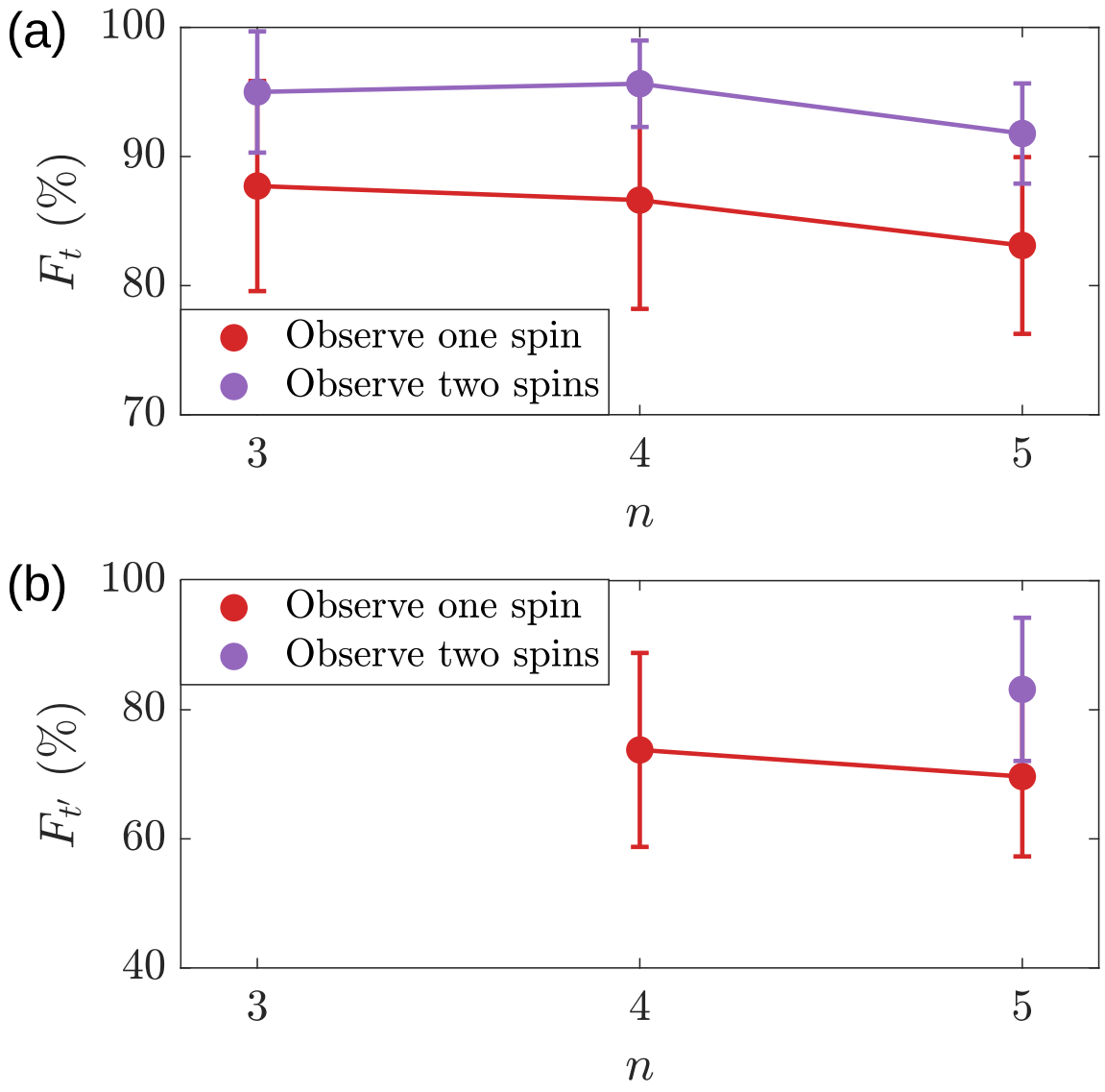

Figure 4.5: Fidelity of predicting random spin systems from observing one spin or two spins. (a) Average tomography fidelity F_t and the standard deviation. Each data point is the result of averaging over 100 random initial-condition realizations. (b) The corresponding results for $F_{t'}$. In the case of observing one spin, results for $n = 3$ contain large fluctuations. When two spins are observed, only the five-spin system generates reasonable values of $F_{t'}$. In general, observing two spins leads to higher fidelity values.

spins and leads to larger prediction uncertainties. Another expected feature is that, for a fixed system size, observing two spins leads to higher fidelity values [about 95% in Fig. 4.5(a)] as compared with the case of observing one spin [about 85% in

Fig. 4.5(a)]. Figure 4.5(b) shows that the fidelity measure with respect to the hidden structure exceeds 50%, indicating that the interactions among the nonobservable spins can be predicted with statistical confidence. In fact, as stipulated by Eq. (4.11), our HENN can predict not only the existence of the interactions but also their strength as characterized by the coefficients $c^{(1,2)}_{i_1 i_2 \cdots i_n}$.

4.3.3 Tomography of Quantum Gates

We apply our HENN framework to a class of systems that are fundamental to quantum computing: quantum logic gates. Such a gate typically consists of two or three coupled spins [272]. To be concrete, we consider the Toffoli and Fredkin gates with three spins [322] and demonstrate that HENN can perform the tomography. Experimentally, these quantum logic gates can be implemented with optical devices [324, 325] or superconducting qubits [326].

Toffoli gate is a Control-Control Not gate, i.e., when the first and second spins have the signal $|11\rangle$, the third spin will flip [322], which requires certain time, e.g., $t = 1$. By this time, the evolution operator is

$$U_{\text{Toffoli}} = \begin{pmatrix} \mathbb{I}_6 & \\ & \mathbb{X}(t=1) \end{pmatrix}, \tag{4.12}$$

where $\mathbb{I}_6$ is the 6×6 identity matrix,

$$\mathbb{X}(t=1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{4.13}$$

flips the third spin, and the off-diagonal blocks are zero. Similarly, the time evolution operator for the Fredkin gate is

$$U_{\text{Fredkin}} = \begin{pmatrix} \mathbb{I}_5 & & \\ & \mathbb{X}(t=1) & \\ & & 1 \end{pmatrix}. \tag{4.14}$$

A physical constraint for $\mathbb{X}(t)$ is that, at $t = 0$, the system does not evolve, so $\mathbb{X}(t = 0) = \mathbb{I}_2$. To build such a time evolution operator, one can use the underlying time-independent Hamiltonian as a base (Appendix 4.6), where the elements of $\mathbb{X}(t)$ are periodic functions with the fundamental frequency as the flipping rate. Searching for possible forms of $\mathbb{X}(t)$ is a basic issue in designing quantum logic gates [286, 287].

To demonstrate the applicability of HENN to quantum logic gates in a concrete manner, we choose $\mathbb{X}(t)$ as

$$\mathbb{X}(t) = \frac{1}{2}\begin{pmatrix} 1+\exp(i\pi t) & 1-\exp(3i\pi t) \\ 1-\exp(3i\pi t) & 1+\exp(i\pi t) \end{pmatrix} \tag{4.15}$$

to generate the time-dependent Hamiltonian. For training, we generate time series from $t = 0$ to $t = 1$ with the time step $dt = 0.01$ from 100 random initial conditions, and the time evolution of the dynamical variables of the third spin is taken as the measurements. For comparison, we calculate the tomography fidelity for both time-dependent and the corresponding time-independent systems.

Table 4.3: Tomography fidelity for Toffoli and Fredkin gates

	Toffoli gate	Fredkin gate
Time-dependent	$81\% \pm 3\%$	$87\% \pm 3\%$
Time-independent	$85\% \pm 4\%$	$92\% \pm 3\%$

Table 4.3 lists the values of the tomography fidelity for both Toffoli and Fredkin gates. It can be seen that the average tomography fidelity for the Fredkin gate is larger than that for the Toffoli gate. Both gates have seven links, but the Fredkin gate has more three-body coupling terms than the Toffoli gate. This means that, for the Fredkin gate, more links are directly connected to the observed spin. When the system becomes time-independent, the fidelity values are slightly higher.

4.3.4 Quantum Tomography under Noise

Noise arising from random coupling with the environment will impact the quality of quantum tomography. Previously, the issue of noise was studied in the context of Hamiltonian learning [290, 296]. Here, we study the effect of noise on our HENN based quantum tomography. A relevant point is that HENN can predict the correct form but only of the local Hamiltonian, so the Hamiltonian for the whole spin system cannot be uniquely determined. This is similar to obtaining the local Hamiltonian from local observations [296].

As a concrete example, we study a cyclic network of four spins, where measurements are taken from only one spin, which are subject to additive white Gaussian noise. The local Hamiltonian H_o constitutes the three Pauli matrices for the observed spin, giving rise to three measured time series. The Hamiltonian $H_{o'}$ consists of the three Pauli matrices for the observed spin and all the interactions that involve the observed spin. In particular, there are 27 two-body interactions, 81 three-body interactions and 81 four-body interactions. As a result, $H_{o'}$ can generate a total of 192 time series. From a different perspective, for a four spin system, the total number of dynamical variables is 256. When we measure one spin so that the other three spins are not observed, H_h contains one fourth of the dynamical variables (4^3) while $H_{o'}$ has the remaining.

Figure 4.6 shows that the fidelity value decreases, albeit slowly, as the noise amplitude increases. The robustness of HENN against weak noise roots in the goal of HENN: finding one Hamiltonian that minimizes the loss function. However, for strong noise, the derivatives in the Heisenberg equation will generate unstable solutions. Figure 4.6 also reveals the similarity between local Hamiltonian recovery and the tomography of the whole system. In particular, it demonstrates that HENN can

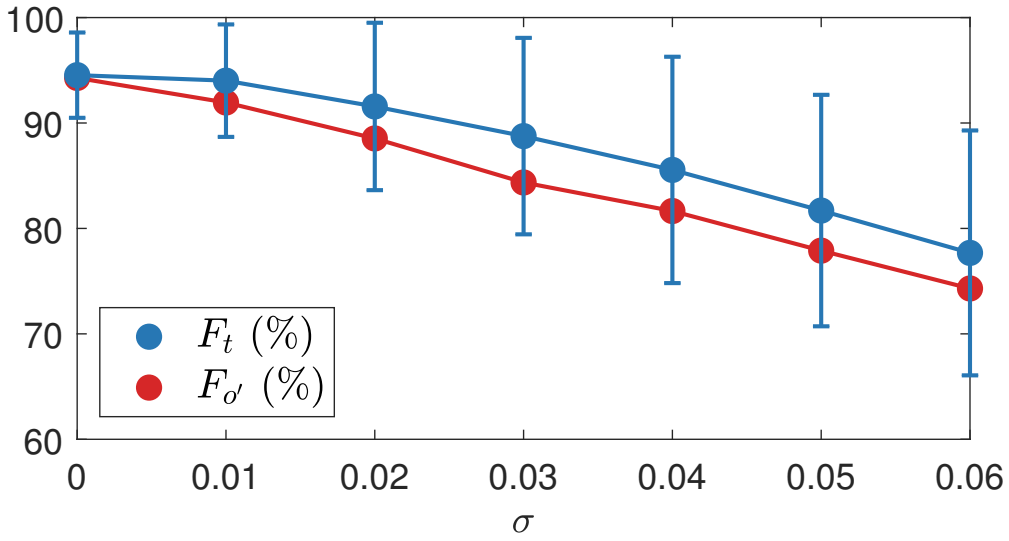

Figure 4.6: Network tomography fidelity under Gaussian noise. The system is a cyclic network of four spins constructed according to Eq. (4.9), where measurements are taken from one node. The additive white Gaussian noise has the variance $\sigma = 0.06$, which is approximately one third of the average variation of the measured time series. The two sets of data represent the two fidelity measures F_t and $F_{o'}$ versus the noise amplitude, where $F_{o'}$ characterizes the difference between the recovered local Hamiltonian and the true one.

recover not only the local Hamiltonian, but also the hidden structure of the spin system which is not a simple extension of the local Hamiltonian.

4.4 Discussion

In quantum tomography, learning time-dependent systems from partial and limited measurements remains a challenge, as it requires optimization in an infinite-dimensional space. Machine learning provides a potentially viable solution. Because of the underlying physics of the spin systems, it is necessary to incorporate the physical constraints into the learning algorithms. Historically, the idea of developing physics informed artificial neural networks was conceived almost three decades ago [160], but it has recently attracted a revived interest, particularly in the context of HNNs [153, 155, 156]. Our idea is that, for time-dependent quantum systems,

the Heisenberg representation is natural in which the operators evolve with time as governed by the Heisenberg equation. Neural networks taking into account the physical constraints manifested as the Heisenberg equation may thus provide an approach to tomography of time-dependent quantum systems. Using spin systems that have been exploited extensively in quantum computing as a paradigm, we have developed a class of Heisenberg Neural Networks (HENNs) and demonstrated that, based on the time series measurements of local spin variables, not only the local Hamiltonian but that of the whole spin system can be faithfully determined. The method is effective even when measurements are conducted on a small part of the system, e.g., measuring one spin in a five-spin system. Considering that the existing algorithms on quantum tomography of spin systems were designed for networks whose structures are completely known [327, 306, 328], our work represents a useful complement.

For quantum tomography of time-dependent interacting spin systems, we have tested a variety of network structures. In general, the tomography fidelity depends on the interacting structure of the network. For example, it is inversely proportional to the degree of the spin from which measurements are taken (Fig. 4.3) when the spin system is relatively dense, as a large degree means more interactions with the nonobservable spins. The fidelity value also depends on the number of observed spins relative to the total number of spins in the network where, naturally, measuring more spins can lead to higher fidelity values (e.g., Fig. 4.5). Indeed, a comparison of the tomography results from the Toffoli gate with those from the Fredkin gate reveals explicitly that more coupling links with the spin being measured lead to increased fidelity values, when the spin system is relatively sparse.

A number of factors can affect the tomography accuracy. For the spin systems studied, the choices of the coupling and the time dependence as characterized by the driving function $f(t)$ are arbitrary. Our study has revealed that the tomography

quality does not depend on the time signal insofar as the length of the measured time series is proper, but the HENN predicted Hamiltonian tends to deviate from the true one after approximately half of the driving period. If the time series are too short, e.g., a fraction of the driving period, or if the time series are too long, e.g., more than a few driving periods, the resulting fidelity value would decrease. Another factor that can affect the fidelity is heterogeneity in the couplings in the network. For the results in this paper, the distribution of the couplings in the spin system is assumed to be uniform, where the typical fidelity value achieved is about 90%. However, we find that large variations in the coupling strengths can make the HENN ineffective. The ratio between self- and mutual couplings can also affect the tomography, where if the former dominate the latter, the errors in the tomography can be reduced.

Possible extensions of this work are as follows. Firstly, when implementing the HENN, the initial states must be specified with a number of constraints. That is, it is necessary to know the initial quantum state of each spin initially. For time-independent systems, quantum tomography of spin systems is possible even if the initial states are not completely specified [306] or if the dimension of the system with a given coupling structure needs to be determined [329, 330]. However, to our knowledge such methods are limited in time independent structure. Such as in [306] They consider a problem to reconstruct a time independent Hamiltonian for a specific structure. Since the structure is known, the matrix representation for observable at time t can be expanded to a few unknown parameters. Once the time series for measurements is long enough the original Hamiltonian and initial states can be constructed. It would be useful to study if these approaches can be extended to time-dependent systems with partial initial conditions or incomplete information about the system.

The second issue is scalability. For spins system with n spins HENN requires

4^n of measurements, which exponential growth as we increase the number of spins. Calculating the matrix representation for the observations $A(t)$ is time-consuming. Another issue arises in the prediction part. For large system the predicted Hamiltonian will contain more terms thus make the neural network very large. Recently, neural networks with more than 100 billions parameters have been constructed [331], with applications to large systems with 100 spins [296]. In general, to scale our HENN algorithm to large time-dependent spin systems remains to be a challenge.

Thirdly is about the effect of Noise. In Noisy Quantum (NISQ) [332, 333], one of the goals is to extract the maximum quantum computational power from current devices. In [274] the fidelity for a two-spin gate is larger than 99%, but fidelity drops when quantum computing run multi cycles. In our paper the noise effect is relatively large due to the derivative is taken on the time series contains noise. One way to reduce the noise is using signal processing technics. Use low-pass filter, which only allows the low frequency signal pass. Another possible way is to change the loss function, since the noise appears in $\langle \dot{A}(t) \rangle_{\text{real}}$. We can change it a matrix form $\mathcal{L} = \sum_{\text{Observations}} \left\| \dot{\mathbb{A}}^{(H)}(t)_{\text{real}} - \dot{\mathbb{A}}^{(H)}(t)_{\text{pred}} \right\|^2$ and use matrix inversion technics to calculating $\mathbb{A}^{(H)}(t)$ and its corresponding time derivative $\dot{\mathbb{A}}^{(H)}(t)$. During the matrix inversion process, when the number of initial states is sufficiently large, which means the number of equations are much larger than the number of unknown parameters, the computed matrix $\mathbb{A}^H(t)$ should converge to the real one. Therefore, we can decrease the noise effect by taking observations on more linearly independent states for unbias Gaussian noise.

Fourthly, In classical mechanics, integrating Hamiltonian system can often be stiff, a concept that usually refers to differential equations where we have to take very small time-steps of integration so that the numerical solution remain stable [334]. Symplectic integrators are especially interesting in order to respect the conserved quantities

in Hamiltonian systems, thereby usually being more stable and structure-preserving than non-symplectic one. For example, the simplest symplectic integrator is the well-known leapfrog method, also known as the Stömer-Verlet integrator [335]. However, even the best integrators remain severely challenged by phenomena as intuitive as a mechanical rebound or a slingshot effect, which are more severe forms of stiffness. Such numerical issues are almost doomed to conflict with the inherently approximate nature of a learning algorithm. Recently there has one symplectic recurrent neural networks (SRNNs) [239]. In the three-body problem, the SRNN-trained Hamiltonian compensates for discretization errors. It can even outperform numerically solving the ODE using the true Hamiltonian at the same time-step size. We notice that we need Baker–Campbell–Hausdorff formula for time-dependent Hamiltonian to change the Hamiltonian from Heisenberg picture to Schrödinger picture [336]. It may be interesting to implement corresponding algorithms to maintain integral accuracy in quantum mechanics.

Recently, there have been efforts in understanding machine learning (the field of explainable machine learning), and physics-enhanced machine learning represents a useful perspective [251]. The HENN articulated in this paper, which includes time correlation, provides an effective way to ascertain and understand the hidden structures in the neural network. As such, our HENN may be exploited as a paradigm for explainable machine learning.

4.5 Examples of Heisenberg Neural Networks

We present a number of examples for which the HENN can be explicitly constructed.

4.5.1 One Spin System

The Hamiltonian of a single spin system under a periodic driving is

$$H(t) = \sigma_x \sin(t). \tag{4.16}$$

We use the convention $\sigma_x = \sigma_+ + \sigma_-$ and $\sigma_y = i(-\sigma_+ + \sigma_-)$, where $\sigma_{+,-}$ correspond to the creation and annihilation operators, respectively. In the basis $|1\rangle$ and $|0\rangle$, the matrix expressions of σ_x and σ_y are

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ and } \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{4.17}$$

The third spin operator is defined by $\sigma_z|1\rangle = |1\rangle$ and $\sigma_z|0\rangle = -|0\rangle$.

The HENN for the one-spin system can be constructed by following the three steps described in Sec. 4.2, as follows.

Step 1: We generate random initial conditions $|\psi_0\rangle$ given by

$$\psi_0 = \frac{1}{r_1 + r_2} \begin{pmatrix} \sqrt{r_1}\exp(i2\pi\theta_1) \\ \sqrt{r_2}\exp(i2\pi\theta_2) \end{pmatrix} = \begin{pmatrix} \phi_{11} + i\phi_{12} \\ \phi_{21} + i\phi_{22} \end{pmatrix}, \tag{4.18}$$

where $r_{1,2}$ are the initial probabilities in the respective state, $\theta_{1,2}$ are the corresponding phase variables, both $r_{1,2}$ and $\theta_{1,2}$ are uniform random numbers between zero and one. In the machine learning algorithm, all quantities are real, so it is necessary to convert the wavefunction into the summation for the real and imaginary parts.

The time evolution of ψ_0 is governed by the Schrödinger equation. The expectation value of operator A is given by $\langle\psi_t|A|\psi_t\rangle$. In the Heisenberg picture, the states do not change but the operators change with time. We expand the operator $A^H(t)$ in the basis $|1\rangle$ and $|0\rangle$, with the corresponding matrix $\mathbb{A}^H(t)$. The Hermitian property of $\mathbb{A}^H$ stipulates it must contain 4 independent elements

$$\mathbb{A}^H(t) = \begin{pmatrix} A_1^H(t) & A_2^H(t) + iA_3^H(t) \\ A_2^H(t) - iA_3^H(t) & A_4^H(t) \end{pmatrix}. \tag{4.19}$$

Both the Schrödinger and Heisenberg pictures should give the same physical results. We have

$$(\phi_{11}-i\phi_{12},\phi_{21}-i\phi_{22})\begin{pmatrix} A_1^H(t) & A_2^H(t)+iA_3^H(t) \\ A_2^H(t)-iA_3^H(t) & A_4^H(t) \end{pmatrix} \begin{pmatrix} \phi_{11}+i\phi_{12} \\ \phi_{21}+i\phi_{22} \end{pmatrix} = \langle\psi_t|A|\psi_t\rangle. \tag{4.20}$$

The unknown elements $A^H(t)$ appear in the equation in a linear fashion:

$$\begin{aligned} A_1^H(t)(\phi_{11}^2+\phi_{12}^2)+A_2^H(t)(2\phi_{11}\phi_{21}+2\phi_{12}\phi_{22})+A_3^H(t)(2\phi_{12}\phi_{21}-2\phi_{11}\phi_{22}) \\ +A_4^H(t)(\phi_{21}^2+\phi_{22}^2)=\langle\psi_t|A|\psi_t\rangle. \end{aligned} \tag{4.21}$$

At least four different initial conditions are required to solve this equation, and a further increase in the number of states changes little the result. We take A to be a Pauli matrix. It expectation value versus time for a given initial state is illustrated in in Fig. 4.7(a).

Step 2: Let the Hamiltonian of the unknown system be $H(t)$. The corresponding operator in the Heisenberg picture is $H^H(t)$. Expanding the operator in the basis, we get

$$\mathbb{H}^H(t)=\begin{pmatrix} H_1^H(t) & H_2^H(t)+iH_3^H(t) \\ H_2^H(t)-iH_3^H(t) & H_4^H(t) \end{pmatrix}. \tag{4.22}$$

The quantum evolution is governed by the Heisenberg equation:

$$\frac{dA^H(t)}{dt}=i[H^H(t),A^H(t)]. \tag{4.23}$$

For a given initial state, the Heisenberg equation can be written in the matrix form as

$$\frac{d}{dt}\langle A\rangle_t = i\psi_0^\dagger\left[\mathbb{H}^H(t)\mathbb{A}^H(t)-\mathbb{A}^H(t)\mathbb{H}^H(t)\right]\psi_0. \tag{4.24}$$

Expanding the right side, we get a summation of 24 terms. As the system size is increased, the number in the summation grows quickly. Through the matrix product, we get

$$\frac{d}{dt}\langle A\rangle_t = \mathbb{T}\boldsymbol{\phi}\boldsymbol{\phi}\mathbf{A}(t)\mathbf{H}(t), \tag{4.25}$$

where

$$\begin{aligned}
\boldsymbol{\phi} &= [\phi_{11}, \phi_{12}, \phi_{21}, \phi_{22}]^T, \\
\mathbf{A}(t) &= [A_1^H(t), A_2^H(t), A_3^H(t), A_4^H(t)]^T, \\
\mathbf{H}(t) &= [H_1^H(t), H_2^H(t), H_3^H(t), H_4^H(t)]^T,
\end{aligned}$$

and $\mathbb{T}$ is tensor of rank four (with dimension $4 \times 4 \times 4 \times 4$), which depends only on the dimension of the system and is defined as

$$\mathbb{T}\boldsymbol{\phi}\boldsymbol{\phi}\mathbf{A}(t)\mathbf{H}(t) = \sum_{i,j,m,n} \mathbb{T}_{ijmn}\boldsymbol{\phi}_n\boldsymbol{\phi}_m A_j(t)H_i(t). \tag{4.26}$$

The left side of Eq. (4.24) contains the derivatives of the measurements, which can be determined from the observations. On the right side, $\mathbf{A}(t)$ and $\boldsymbol{\phi}$ are known, so the unknown quantity is $\mathbf{H}(t)$.

In the HENN, we set the input dimension as one and the output is $\mathbf{H}(t)$. We choose the batch size to be the number of measurement points times the number of different initial states. The loss function is

$$\mathcal{L} = \sum_{A=\sigma_x,\sigma_y,\sigma_z} \left|\langle \dot{A}(t)\rangle_{\text{real}} - \mathbb{T}\boldsymbol{\phi}\boldsymbol{\phi}\mathbf{A}(t)\mathbf{H}(t)\right|^2. \tag{4.27}$$

We build the neural network from the Keras Tensorflow package [255], where the input is connected to two dense layers. Constructing this customized loss function is equivalent to designating a loss function with different weights.

Step 3: After the HENN is trained, we input a time series from 0 to 5 and predict the Hamiltonian. The prediction is carried out in the Heisenberg picture, which can be converted into the corresponding Hamiltonian in the Schrödinger picture. This can be done through an iteration process.

From the Hamiltonian in the Schrödinger picture, we can get the coefficients in each base through

$$H(t) = c_0(t)\mathbb{I}_2 + c_1(t)\sigma_x + c_2(t)\sigma_y + c_3(t)\sigma_z. \tag{4.28}$$

Writing it in a general form $H(t) = \sum_i S_i c_i(t)$, where $S = \mathbb{I}_2, \sigma_x, \sigma_y, \sigma_z$, we have that $c_i(t)$'s contain all the information about the Hamiltonian. The predicted $c_i(t)$ in a given basis are shown Fig. 4.7(b). The agreement between the solid (predicted) and dashed (true) curves is proof that the HENN can recover the Hamiltonian of the original system through observations.

The coupling configuration can be determined through

$$\overline{c}_i = \int |c_i(t)| dt. \tag{4.29}$$

If there exists a coupling between the spin components, the corresponding coefficient $c_i(t)$ should be non-zero, giving rise to a non-zero value of $\overline{c}_i$. Figure 4.7(c) shows the time averaged result of $c_i(t)$, which has a pronounced value in σ_x, in agreement with the original Hamiltonian (4.16).

4.5.2 A Three-spin Chain

We consider three interacting spins on a chain, with measurements taken from the first spin, as shown in Fig. 4.8. The Hamiltonian is

$$H(t) = \sin(t)\left(\sum_{i=1}^{3}\sum_{j=1}^{2}\sigma_j^i + \sum_{i=1}^{2}\sum_{l=1}^{3}\sum_{m=1}^{2}\sigma_l^i\sigma_m^{i+1}\right). \tag{4.30}$$

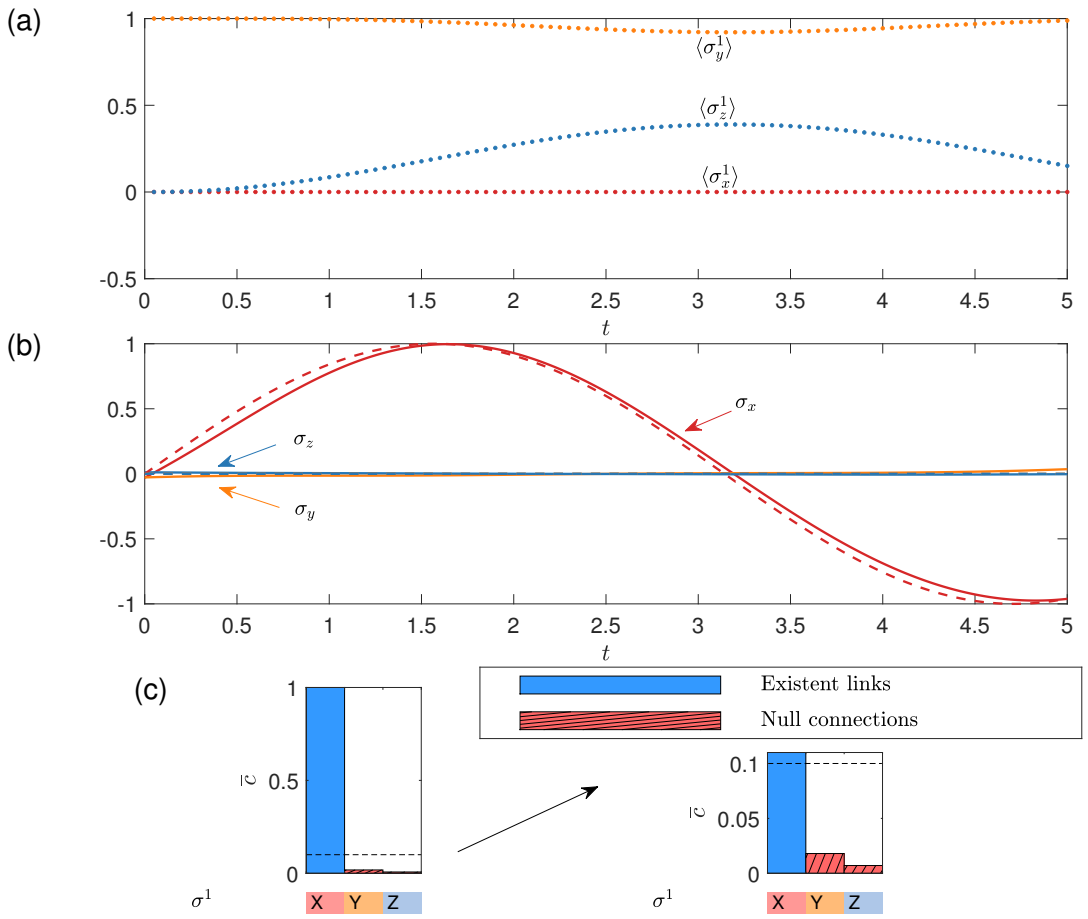

Figure 4.7: HENN based tomography of a single-spin system. (a) Time series of σ_x, σ_y and σ_z, where the initial state is $(|1\rangle + i|0\rangle)/\sqrt{2}$. The dots correspond to the sampled measurements. (b) The predicted Hamiltonian in the Schrödinger picture, where the three curves correspond to the decomposition in the three base states and the dashed curves are the true values. (c) Time average of the absolute value of $c_i(t)$ in different base states. The existent links are marked by the blue color, and the nonexistent ones by red, where the threshold for determining the existent links is 10% of the maximum value.

There are couplings between spins one and two, and between spins two and three. For each coupling, there are six links.

The HENN can be constructed following the three steps.

Step 1: The system contains 2^3 independent states: $|111\rangle, |110\rangle, |011\rangle \cdots |000\rangle$. The initial conditions are chosen according to Eq. (4.18), with the difference that here

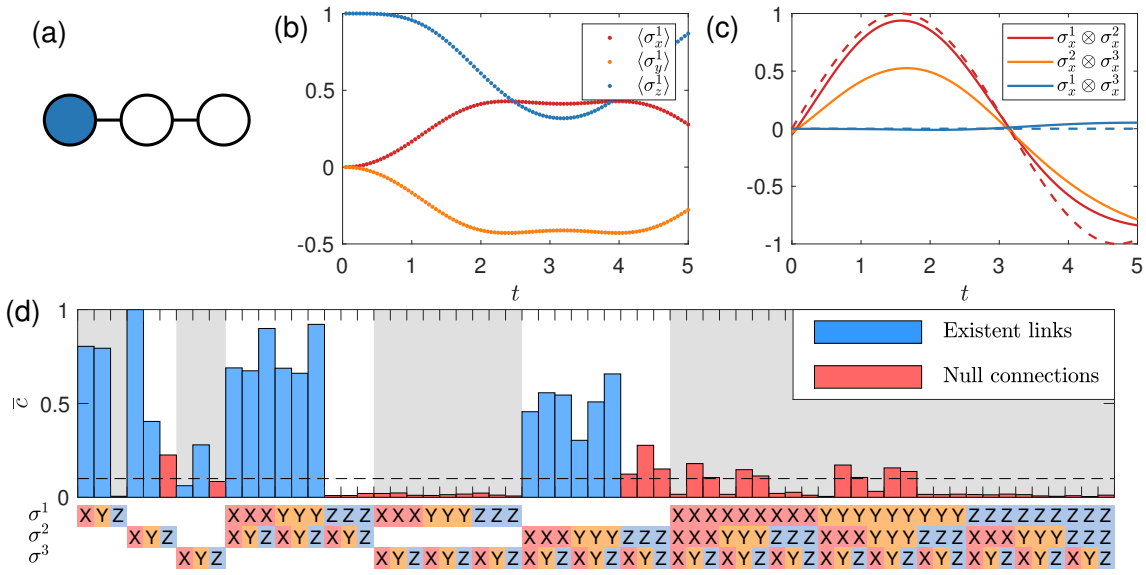

Figure 4.8: HENN based tomography of a three-spin chain system. (a) Schematic illustration of the chain, where measurements are taken from the left spin. (b) Time series of σ_x^1, σ_y^1 and σ_z^1 from the initial state $\frac{|1\rangle+i|0\rangle}{\sqrt{2}}|11\rangle$, where the dots correspond to the sampled measurement time series. (c) Predicted Hamiltonian in the Schrödinger picture after decomposition into different interaction terms, where $\sigma_x^1\sigma_x^2$ and $\sigma_x^1\sigma_x^3$ are the couplings between the observed node and the nonobservable nodes, which agree with the true Hamiltonian terms (dashed traces). The term $\sigma_x^2\sigma_x^3$ specifies the coupling between the two nonobservable nodes, and the predicted Hamiltonian is not similar to the true function $\sin(t)$ but not zero either, so the HENN does predict the existence of this interaction. (d) Time average of the absolute value of the coefficients $c_i(t)$ associated with different interaction terms, where the existent links are marked blue and the nonexistent ones red. The threshold for distinguishing the existent from nonexistent links is set to be 10% of the maximum coefficient value, as indicated by the horizontal dashed line.

there are eight dimensions. Since observations are taken from the first spin, we write $\sigma_x^1 = \sigma_+^1 + \sigma_-^1$, where the creation and annihilation operators act only on the first spin. The matrix expression for σ_x^1 is

$$\sigma_x^1 = \sigma_x \otimes \mathbb{I}_4.$$

Similarly, we can get the matrices for σ_y^1 and σ_z^1. For a given initial state, we calculate the expectation values of the three observables on the first spin, as shown in Fig. 4.8(b).

Step 2: Similar to the one-spin system, the Heisenberg equation is

$$\frac{d}{dt}\langle A\rangle_t = \mathbb{T}\boldsymbol{\phi}\boldsymbol{\phi}\mathbf{A}(t)\mathbf{H}(t), \tag{4.31}$$

where $\boldsymbol{\phi}$ is a vector of $2\times 2^3 = 16$ elements, $\mathbf{A}(t)$ and $\mathbf{H}(t)$ contain $4^3 = 64$ elements, and $\mathbb{T}$ is a tensor of rank four with the dimension $64\times 64\times 16\times 16$:

$$\mathbb{T}\boldsymbol{\phi}\boldsymbol{\phi}\mathbf{A}(t)\mathbf{H}(t) = \sum_{i,j,m,n} \mathbb{T}_{ijmn}\boldsymbol{\phi}_n\boldsymbol{\phi}_m A_j(t)H_i(t). \tag{4.32}$$

We build up the HENN according to the same loss function as in the case of a single-spin system, predict the Hamiltonian, and convert it to the Schrödinger picture. The Hamiltonian can be decomposed as

$$\begin{aligned}\mathbb{H}(t) = c_0(t)\mathbb{I} &+ \sum_{i,j} c_{i,j}(t)\sigma_j^i \\ &+ \sum_{i,j,m,n} c_{ijmn}(t)\sigma_j^i\sigma_n^m \\ &+ \sum_{i,j,m,n,k,l} c_{ijmnkl}(t)\sigma_j^i\sigma_n^m\sigma_l^k.\end{aligned} \tag{4.33}$$

The decomposition becomes cumbersome for systems with more than one spin. we thus write this as the direct product of the Pauli matrices plus the identity matrix. For example, the two-body coupling $\sigma_1^1\sigma_2^2$ can be written as $\sigma_1^1\otimes\sigma_2^2\otimes\sigma_0^3$.

Figure 4.8(c) shows the predicted Hamiltonian in several base states. The Hamiltonian for the coupling $\sigma_x^1\sigma_x^2$ can be compared with the sinusoidal function $\sin(t)$. The coupling term $\sigma_x^1\sigma_x^3$ is non-existent in the original system, so it should be compared with zero. The agreement indicates that the local Hamiltonian between the observed spin and the nonobservable spins can be recovered. Note that $\sigma_x^2\sigma_x^3$ represents a coupling between the two nonobservable nodes, whose true value is $\sin(t)$, but the predicted Hamiltonian is not close to it. Nonetheless, the non-zero value of the predicted term indicates the existence of the coupling term $\sigma_x^2\sigma_x^3$.

Step 3: After decomposing the Hamiltonian in different terms, we take the time average of each and normalize them, as shown in Fig. 4.8(d). The ideal case is that all the blue points have the value one and all the red points are zero. First, the Hamiltonian Eq. (4.30) contains self-couplings in x and y but not in z, which are indicated by the first three bars in Fig. 4.8(d). For the nonobservable spins, there are errors in predicting the self-coupling terms. Second, there are two-body interactions between spins one and two, and between spin two and three, but not between one and three, where each existent interaction has six terms of coupling. The predicted results for the couplings involving the first spin are more accurate than those between the nonobservable spins. Third, the true Hamiltonian does not include any three-body interactions, so all such terms should be zero.

The results in Fig. 4.8 indicate that our HENN can perform accurate tomography of the three spin chain.

4.6 Time-independent Toffoli and Fredkin Gates

The Hamiltonian for the time-independent Toffoli gate is

$$H_{\text{Toffoli}} = \frac{\pi}{8}(\mathbb{I}_2 - \sigma_3^1)(\mathbb{I}_2 - \sigma_3^2)(\mathbb{I}_2 - \sigma_1^3). \tag{4.34}$$

The corresponding time evolution operator is

$$U_{\text{Toffoli}} = \begin{pmatrix} \mathbb{I}_6 & \\ & \mathbb{X}_0(t) \end{pmatrix}, \tag{4.35}$$

where

$$\mathbb{X}_0(t) = \frac{1}{2}\begin{pmatrix} 1+\exp(i\pi t) & 1-\exp(i\pi t) \\ 1-\exp(i\pi t) & 1+\exp(i\pi t) \end{pmatrix}. \tag{4.36}$$

The Hamiltonian for the time-independent Fredkin gate is

$$H_{\text{Fredkin}} = \frac{\pi}{8}(\mathbb{I}_2 - \sigma_1)\left[\mathbb{I}_4 - \sum_{\alpha=1}^{3}\sigma_2^\alpha\sigma_3^\alpha\right], \tag{4.37}$$

with

$$U_{\text{Fredkin}} = \begin{pmatrix} \mathbb{I}_5 & & \\ & \mathbb{X}_0(t) & \\ & & 1 \end{pmatrix}. \tag{4.38}$$

4.7 Appendix

Following are the relevant publication on the topics presented in this chapter

C.-D. Han, B. Glaz, M. Haile, and Y.-C. Lai*, "Tomography of time-dependent quantum Hamiltonians with machine learning," Physical Review A **104**, 062404, (2021).

Chapter 5

META MACHINE-LEARNING BASED QUANTUM SCAR DETECTOR

A remarkable phenomenon in contemporary physics is quantum scarring in systems whose classical dynamics are chaotic, where certain wavefunctions tend to concentrate on classical periodic orbits of low periods. Quantum scarring has been studied for more than four decades, but detecting quantum scars still mostly relies on human visualization of the wavefunction patterns. The widespread and successful applications of machine learning in many branches of physics suggest the possibility of using artificial neural networks for automated detection of quantum scars. Conventional machine learning often requires substantial training data, but for quantum scars this poses a significant challenge: in typical systems the available distinct quantum scarring states are rare. We develop a Meta machine-learning approach to accurately detecting quantum scars in a fully automated and highly efficient fashion. In particular, taking advantage of some standard large dataset such as Omniglot from the field of image classification, we train a "preliminary" version of the neural network that has the ability to distinguish different classes of noisy images. We then perform few-shot classification to further train the neural network but with a small number of quantum scars. We demonstrate that the Meta learning scheme can find the correct quantum scars from thousands images of wavefunctions without any human intervention, regardless of the symmetry of the underlying system. Our success opens the door to exploiting Meta learning to solve challenging image detection and classification problems in other fields of science and engineering.

5.1 Introduction

In the field of quantum chaos that studies the quantum manifestations of classical chaos [3, 2], wavefunction scarring is a fundamental phenomenon that has been extensively investigated. Here the term "scarring" is referred to as the unusually high concentrations of the wavefunctions about certain classical periodic orbits. In this regard, if the underlying classical dynamics are integrable with stable periodic orbits, non-uniformly distributed wavefunctions about these orbits are expected. What is surprising is quantum scarring in systems whose classical dynamics are fully chaotic, where all periodic orbits are unstable so, intuitively, it seems not possible for the wavefunctions to concentrate about them. From another viewpoint, because of the intrinsic ergodicity associated with chaos, a classical trajectory generates a uniform distribution in the phase space (accordingly in the physical space as well), so the intuition would be that the quantum wavefunctions should also have a uniform distribution in space. It was first discovered by McDonald and Kaufman [7] when solving the Helmholtz equation in the classically chaotic stadium billiard that there are eigenstates whose wavefunctions are highly non-uniform and in fact tend to concentrate on some classical unstable periodic orbits. A more detailed study by Heller [8] confirmed the phenomenon, who gave the name "quantum scars" to the non-uniform wavefunctions. A theory for quantum scars based on the semiclassical Green's function was developed by Bogomolny [337] and Berry [338].

There are a variety of physical systems in which the phenomenon of quantum scarring can occur. For example, in graphene, the low energy excitations are governed by the Dirac equation [15, 17, 19] and correspond to massless particles, where relativistic quantum scars can arise [207, 11, 12]. Solutions of the Dirac equation in classically chaotic billiards also revealed a distinct class of relativistic quantum scars

- chiral scars that require two complete cycles for their wavefunctions to return to their original values [135, 208, 209, 339]. Quantum scarring can also occur in open or transport systems such as quantum dots, where the scarred states are referred to as quantum pointer states [340, 341, 342, 343, 344, 345, 346]. In micro cavity lasing systems, quasibound states in a deformed cavity are scarring states with strong directional emission [181, 347, 129, 348]. Quite recently, scarring has been uncovered in quantum many-body systems, where certain special eigenstates were found to concentrate in certain parts of the Hilbert space [349, 350, 351, 352, 353, 354, 355], which were named as quantum many-body scars that are highly relevant to multiple qubit systems in quantum information science and technology.

In the study of quantum chaos, to identify quantum scarring states has been a challenging problem. The conventional approach has been "manual," where one first generates a large number of eigenstates and then visually check to see if an eigenstate is a scarring state based on information about the classical periodic orbits. This can be an extremely difficult task, for two reasons. First, scarring states are rare in a typical quantum system. For example, in a chaotic billiard, only approximately 5% to 10% of the eigenstates are quantum scars. Compounding the difficulty is the different classes of quantum scars, where each class corresponds to a different group of classical periodic orbits. For a given class, quantum scarring states are even more rare. For example, when chiral scars were first discovered [135], more than 10^4 eigenstates were examined visually by human eyes. The second reason is that the wavefunction patterns associated with quantum scars can be complicated with a significant random or noisy component. As a result, labeling the quantum scars is a hard task with significant uncertainties for the human eyes.

In recent years, deep learning has enjoyed great success in visual object recognition, object detection, and many other domains [216]. Deep Convolutional Neural

Networks (DCNNs) were introduced for image recognition, the training of which often requires very large datasets [356]. For example, in Ref. [356], the authors considered 1000 classes, which required more than one million images as the training data. In physics, DCNNs have been adopted in different areas to solve problems associated with a large number of images [158], such as jet tagging [357, 358], neutrino event classifier [359], gravity lenses [360], and identification of symmetry-breaking states from SEM images [361]. A convolution-like neural network structure has also been developed to solve problems in quantum many-body physics [248]. Quite recently, a machine learning based quantum chaos detector was proposed [362]. In applications where such massive datasets are not available, machine learning based on small datasets has also been developed, especially for the classification task [363, 162, 364] where only a small number of labeled examples per class (a few shots) are required. Since only small training data are required for classification, such machine-learning schemes find broad applications, e.g., learning and detecting rare events [365, 366].

The remarkable success of machine learning in image recognition and pattern classification naturally leads to the idea of developing a neural-network based quantum scar detector. However, even as the image classification algorithms have become more powerful, detecting and classifying quantum scars remain to be an open and difficult problem, due to the availability of a small number of quantum scarring patterns and the lack of distinction between quantum scars and conventional wavefunctions. Intuitively, since each type of quantum scars corresponds to a unique class of classical periodic orbits, if we select some scarring states as the training data, the neural network so trained should be able to make the correct classification when an image contains a similar pattern is presented as the input to the network. However, this simple approach may not be workable for detecting quantum scars because it is difficult to find them in the first place and obtaining a large number of training images

is practically impossible. To develop a machine-learning based automated quantum scar detector, this difficulty must be overcome.

In this paper, we exploit Meta learning [161] to develop an automated and efficient quantum scar detector. The essence of Meta machine learning is to encode "previous experience" into a pretrained neural network so that it can quickly adapt to new input. Meta learning has found broad applications in regression, classification and reinforcement learning. Our idea is to exploit some existing data sets to train the neural network, which can be completely unrelated to quantum systems but with images similar to scarring patterns. In this regard, the Omniglot dataset [162] widely used in the field of image classification stands out as an excellent choice. Our strategy is to train a "preliminary" version of the neural network that has the ability to distinguish different classes of noisy images from the Omniglot dataset and then perform few-shot classification to further train the neural network with a small number of quantum scars. A pertinent issue is that standard Meta learning algorithms can only be used to classify but not detect patterns. Our solution is to use an ensemble of neural networks. In particular, non-scarring states as input to the neural networks can be used as references with significantly different statistical features than those associated with scarring states, thereby accomplishing the detection task. We demonstrate that the neural network so trained can detect and classify quantum scars with remarkable efficiency and accuracy even with limited training data of actual quantum scars. Our Meta-learning based approach not only solves a long-standing problem in the field of quantum chaos, but can also be generalized to address challenging image detection and classification problems in other fields.

In Sec. 5.2, we describe a prototypical relativistic quantum billiard system and articulate the basic training procedures for neural networks. In particular, we detail the construction of our quantum-scar detector based on Meta learning (Sec. 5.2.2)

and few-shot classification (Sec. 5.2.3). In Sec. 5.3, we use two types of chaotic Dirac billiard systems to demonstrate the feasibility and power of Meta learning in detecting and classifying relativistic quantum scars. Conclusions and discussions are presented in Sec. 5.4.

5.2 Methods: Relativistic Quantum Scars and Machine Learning Algorithms

5.2.1 Relativistic Quantum Scarring in Chaotic Billiard Systems

We use the standard setting of 2D billiard systems [7, 367, 1] to generate quantum scars. In the classical limit, a particle moves freely inside the billiard, experiencing elastic reflections from the hard-wall boundary. The geometric shape of the boundary determines the nature of the classical dynamics. For example, the dynamics in a circular billiard are integrable but those in a stadium, Africa, or heart-shaped billiard are chaotic. To be concrete, we focus on the relativistic quantum regime as described by the massless Dirac equation. For a chaotic billiard, there are standard methods to calculate the relativistic quantum eigenvalues and eigenstates such as those based on the boundary integrals [368, 6] and conformal mapping [369, 135].

We consider the situation where the particle has spin-1/2 and a vertical magnetic flux [208] is applied through the billiard region, whose strength α is effectively an externally adjustable parameter: different sets of spinor wavefunctions can be generated by changing the value of α. Possessing such an experimentally controllable parameter has two advantages. First, from the point of view of machine learning, the neural network can be trained with one set of the wavefunction patterns and then be tested using datasets from different parameter values, enabling the full power of the learning process to be revealed and exploited. Second, through testing the datasets from different values of α, the statistical properties of the detected quantum scars

can be calculated and compared with the semiclassical prediction [208].

Our working examples are the class of billiards with fully chaotic dynamics in the classical limit [369, 135], whose boundary is defined by the following conformal mapping of a unit circle in the complex plane $z \equiv x + iy$ to the complex $w \equiv u + iv$ plane:

$$w(z) = \frac{z + bz^2 + ce^{i\delta}z^3}{\sqrt{1 + 2b^2 + 3c^2}}, \tag{5.1}$$

where b, c, and δ are parameters, and the the origin in the z plane maps to the origin in the w plane. For example, for $b = 0.49$, $c = \delta = 0$, the billiard has the shape of a heart, which possesses an inverse symmetry with respect to the x axis. For $b = c = 0.2$ and $\delta = \pi/3$, the billiard has the shape of Africa. For a massless relativistic spin-1/2 particle inside the billiard, the Hamiltonian is [208]

$$\hat{H} = v_F \hat{\boldsymbol{\sigma}} \cdot \left(\hat{\mathbf{p}} - \frac{q}{c}\mathbf{A}\right) + V(u, v)\hat{\sigma}_z, \tag{5.2}$$

where v_F is the Fermi velocity, $\boldsymbol{\sigma} = (\sigma_x, \sigma_y)$ and σ_z are the Pauli matrices, the hard-wall confining potential $V(u, v)$ is zero inside and infinite outside the billiard region, and $\mathbf{A}$ is the vector potential. If the magnetic flux is Φ, the vector potential is given by

$$\mathbf{A}(u, v) = \frac{\Phi}{2\pi}\left(\frac{\partial F}{\partial v}, -\frac{\partial F}{\partial u}\right), \tag{5.3}$$

where $\alpha = q\Phi/(hc)$ is the normalized flux strength and the function $F(u, v)$ is determined by the flux profile in the (u, v) plane. For example, for a singular flux through the origin, $F(u, v)$ can be solved from the Poisson-like equation [370]:

$$\nabla^2_{\mu\nu} F = -2\pi\delta(u, v).$$

The eigenvalue problem $H\boldsymbol{\psi} = E\boldsymbol{\psi}$, where $\boldsymbol{\psi}$ is the two-component spinor, can be solved by employing the mapping method [135, 208], where the analytically solvable spinor wavefunctions in the hard-wall circular billiard in the z plane are transformed

to those in the w plane through the conformal mapping. The accuracy of this method depends on the size of the wavefunctions base. In particular, an arbitrarily large number of analytic eigenstates can be written down in the z plane. When transformed into the w plane, a smaller but still large number of eigenstates can be obtained with a preset, desired accuracy. In our work, for each α value in the unit interval, we use 40000 base eigenstates in the circular billiard in the z plane. For the heart and Africa billiards in the w plane, we use the first $15,000$ and the first $10,000$ eigenstates, respectively. The eigenstates are labeled by the integer n.

We represent each eigenstate $\boldsymbol{\psi}(n)$ by an image, where the probability distribution for a uniformly spaced, rectangular grid of points is evaluated and the probability is set to zero for points outside the billiard region. For the heart-shaped and Africa billiards, the image sizes are 202×232 and 201×148, respectively. From the point of view of pattern recognition, the eigenstates can be quite distinct. For example, quantum scarring states, the detection of which is the aim of this work, are those whose probability distribution is concentrated about certain classical periodic orbits and are rare. There are also quantum states with certain patterns that do not correspond to any periodic orbits, such as the boundary states. In fact, the majority of the eigenstates do not have any recognizable patterns.

Quantum scarring states are relatively more pronounced in the semiclassical limit $n \gg 1$. Numerically, for both types of chaotic billiards, we find that the first scarring states, which correspond to classical periodic orbits of period two, emerge at $n \approx 200$. As the period increases, the corresponding scarring states occur in a more semiclassical regime. For example, scarring states corresponding to classical periodic orbits of period four or five begin to arise for $n \approx 1000$. Figure 5.1(a) presents, for the heart billiard, the quantum states with $n = 8480$ and $n = 8523$ for $\alpha = 0.5$, where the $n = 8480$ eigenstate is localized about a square-like periodic orbit and belongs to a

scarring state of period four, while the $n = 8523$ eigenstate is approximately uniform in the billiard with no apparent pattern that can be deemed as a non-scarring state. A few distinct classes of quantum scarring states are shown in Fig. 5.1(b).

It should be noted that, in a classical chaotic system, the number of unstable periodic orbits increases exponentially with the period at the rate of the topological entropy [371]. Within a finite and relatively large period, the number of unstable periodic orbits can be quite large. However, quantum scarring states corresponding to periodic orbits of even moderately large periods tend to be unstable [372]. As a result, the study of quantum scars has mostly been limited to the quantum states that concentrate on classical periodic orbits of low periods.

5.2.2 Meta Learning

In the physical space (i.e., the billiard domain), the probability distribution of a quantum scarring state focusing on a classical periodic orbit of a low period resembles the image of a simple language character. To develop a machine-learning based quantum scar detector, we take advantage of this resemblance by exploiting Meta learning [161], an image classification algorithm that has attracted a great deal of recent interest. In the field of image classification, an often used and somewhat standard dataset is Omniglot [162], which contains more than one thousand handwritten characters taken from different languages. The basic principle underlying our work is then that a machine-learning algorithm that is able to distinguish the characters in Omniglot can be effectively transferred to detecting quantum scars.

Our specific idea is as follows. Since it is practically infeasible to have a training dataset that contains a large number of quantum scarring states, the problem of identifying quantum scars is similar to the tasks that Meta learning algorithms are designed to solve, e.g., image classification for the Omniglot dataset [161]. A

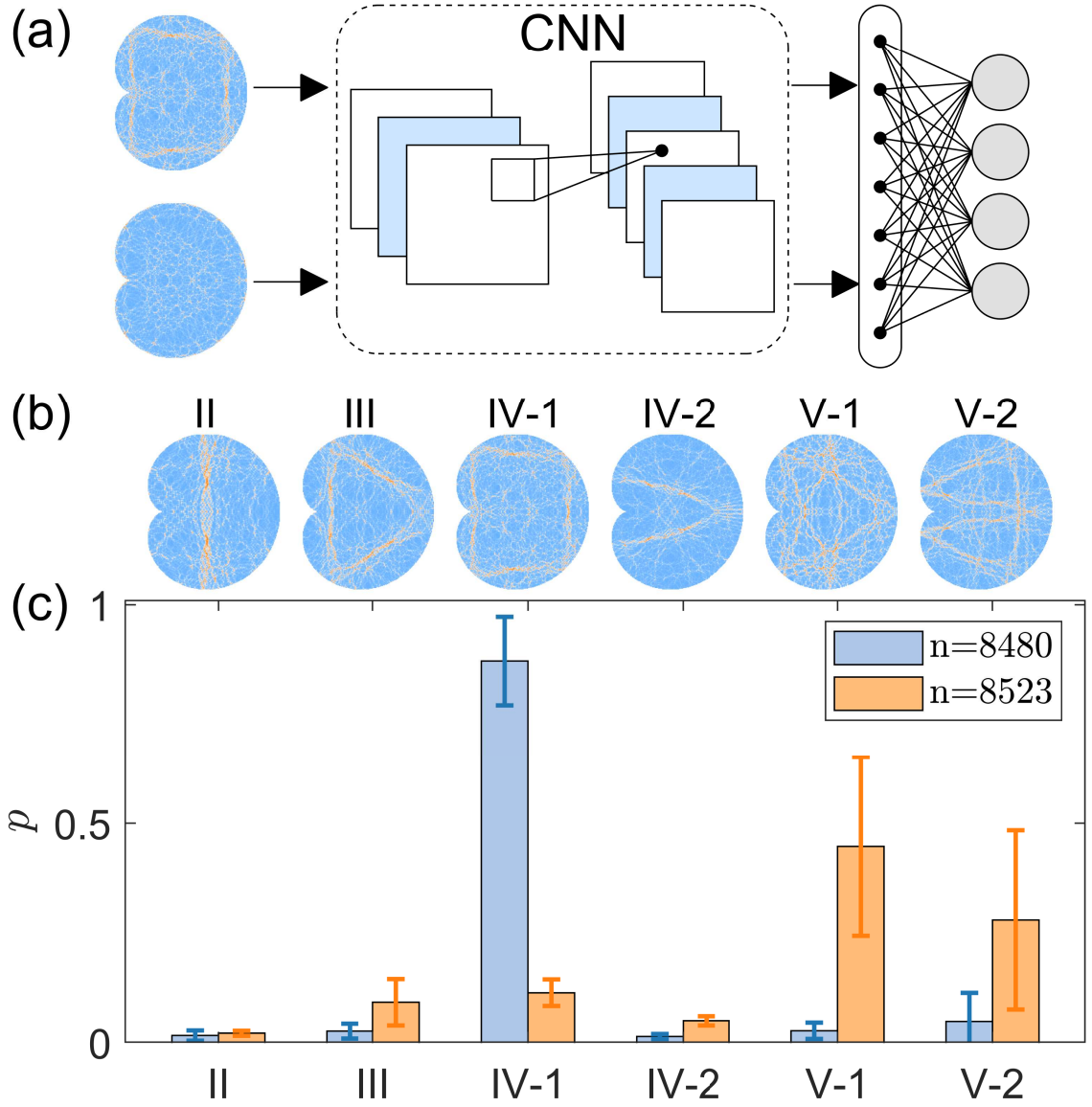

Figure 5.1: Convolutional Neural Network (CNN) based quantum scar detector. (a) Two specific eigenstates (images) from the chaotic heart billiard as the input to the CNN. The upper and lower left states correspond to a scarring and non-scarring state, respectively. The CNN contains convolutional layers and fully connected layers, and the output is an N-dimensional vector whose elements represent the probabilities for the input wavefunction to belong to the corresponding class. (b) Six types of scarring states in a heart billiard, where the corresponding periods of the classical periodic orbits are denoted by Roman numbers. (c) The averages and variances of the output probabilities from a small ensemble of ten neural network realizations.

Meta learning algorithm aims to encode "previous experience" in a pretrained neural network such that it can quickly adapt to new images. To encode the "previous experience," training from a large dataset containing many different patterns is necessary. We can thus use the images in Omniglot as the "substituting" training set. Since there is a unique pattern associated with each type of quantum scars, the problem is essentially one of supervised learning. Utilizing certain images from Omniglot as the training data, we can find the images whose patterns correspond to a specific type of quantum scars.

Our Meta-learning based training process consists of three major components.

The first component is image processing as described in Appendix 5.5, where we add noise to each image in the Omniglot dataset to generate a set of images whose patterns resemble those of quantum scars. We then perform image processing for some quantum eigenstates from a chaotic billiard, after which each eigenstate is represented by a matrix of dimension 100×100 with binary elements. The images so created are used as the inputs to the CNN.

The second component is to conduct training based on the Omniglot dataset to determine the weights and biases of the CNN. In particular, our CNN consists of four convolutional and two fully connected layers. Each convolution layer creates a kernel that is convoluted with the layer input to produce a tensor of outputs. The activation function is chosen to be ReLU - the rectified linear unit. After the convolution stage, we apply layer flattening and change the tensor to a large vector. A linear transformation takes this vector to the final output $\mathbf{v}$, whose dimension is equal to the total number of classes in the system. We use the softmax function to normalized v:

$$p_i = \frac{\exp(v_i)}{\sum_i \exp(v_i)}, \tag{5.4}$$

where p_i is the normalized probability for the input that belongs to class i. The loss

function in the training process is taken to be the cross entropy. The whole process can be implemented in Python with the open source package Tensorflow [254].

Initially, the weights and biases are randomly chosen from a normal distribution labeled as $\theta(w, b)$. In the training process, we first use the modified, noise-contaminated Omniglot dataset to find a pretrained neural network, denoted as $\theta_0(w, b)$, which readily results in different image classes. We use the Reptile algorithm to renew the weights and biases in each step until some criterion is met [373] (see Appendix 5.6 for the algorithmic details). To avoid overfitting for the Omniglot dataset, we monitor the test error from each step, as detailed in Appendix 5.8. The training based on the Omniglot dataset is deemed complete when the global minimum is reached.

The third component is training based on certain actual quantum scarring states. In particular, after training with Omniglot which yields a neural network $\theta_0(w, b)$, we perform further training by using the few-shot classification algorithm with a small number of actual images of quantum scars as the training dataset (see Sec. 5.2.3 below). Since $\theta_0(w, b)$ already has the ability to distinguish different classes of noisy images in the Omniglot dataset, the neural network can quickly adapt to quantum scarring states.

When all three components of the training process are completed, the CNN is effectively a mapping from the wavefunctions to the probability for each class of scars. Figure 5.1(c) shows the prediction probability for the two images in Fig. 5.1(a), where the average is taken from an ensemble of ten neural network realizations (see Sec. 5.3.1 for detail) with different combinations of quantum scars as the training data. Since the particular scarring pattern ($n = 8480$) corresponds to a period-4 orbit, a high value of the probability for the class IV-1 is achieved. For the non-scarring states ($n = 8523$), while there is a relatively large probability to associate it with a period-5 scarring state, the variance is large, making the association highly uncertain. The

distinct statistical behaviors for scarring and non-scarring states are initial indication that our Meta-learning based method has the ability to detect quantum scars.

5.2.3 Few Shot Classification

Physically significant quantum scarring states that have been studied in the past are typically those corresponding to classical periodic orbits of low periods and thus consist of a few classes. From the point of view of image processing, the number of classes to be classified is small and there are only a few images available in each class. The few-shot classification algorithm is designed to deal with such problems [363, 162, 364]. Say the target contains N different classes and each class has at least $K+1$ images with K being a small integer. We randomly choose $K+1$ images in each class, train the neural network using the first K images, and then use the trained network to classify the last image. For the Omniglot dataset, a 95% accuracy can be achieved even for $N = 5$ and $K = 1$, and the accuracy can reach 99% if K is increased to $K = 5$. In general, for a fixed N value, increasing the value of K can improve the accuracy. However, for a fixed K value, a larger value of N leads to a lower accuracy as more classes require more images to train the network.

Table 5.1: Few shot classification accuracy of detecting relativistic quantum scars in the chaotic heart billiard

Shot number	Accuracy
$K = 1$	$90.16\% \pm 1.41\%$
$K = 2$	$95.30\% \pm 0.68\%$
$K = 5$	$98.58\% \pm 0.32\%$

To test the power of few-shot classification in detecting quantum scars, we take the heart billiard as an example, where six distinct classes of scars are shown in Fig. 5.1(b). We create an image dataset where each class contains 10 images taken

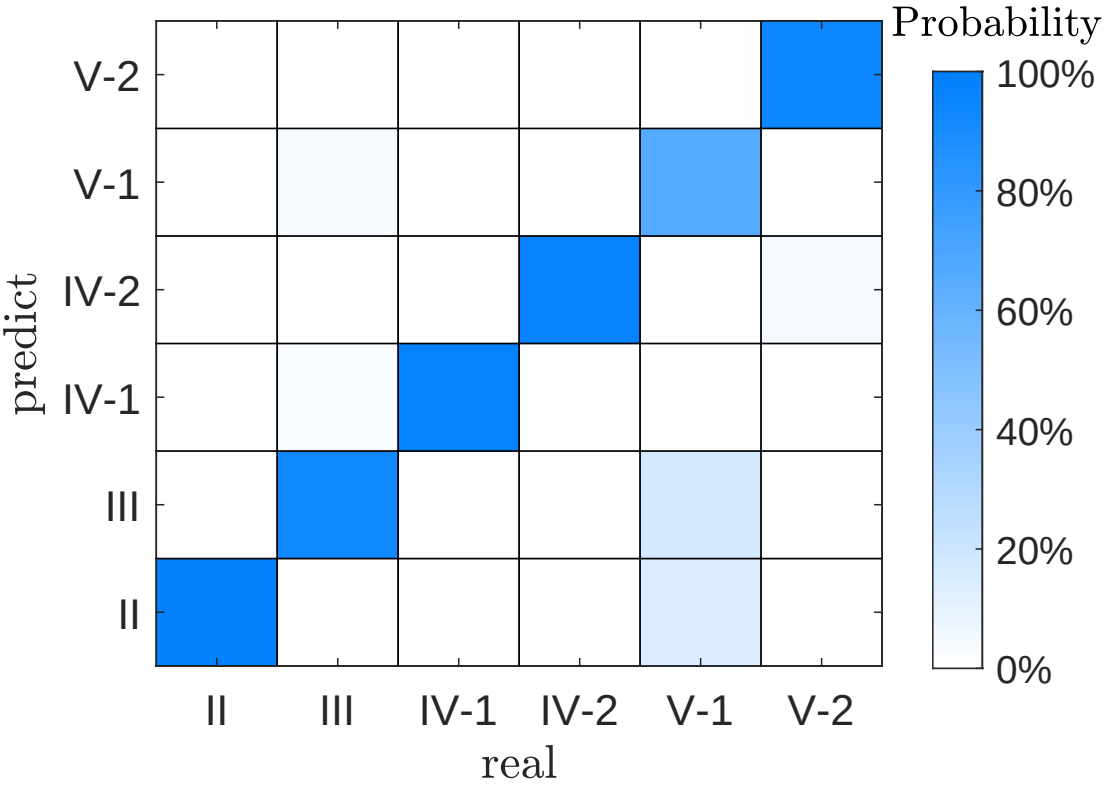

Figure 5.2: Confusion matrix associated with detection of quantum scars in the chaotic heart billiard. The horizontal and vertical axes are the labels of the real and detected classes, respectively. The various probabilities are color-coded as defined by the color bar on the right, where the diagonal and off-diagonal elements are the probabilities of correct and incorrect detection of different classes of quantum scars, respectively. The detection accuracies (the values of the diagonal elements) are generally close to 100% except for the V-1 scars - see text for an explanation.

from the eigenstates whose energy index $n \in [1000, 15000]$. We obtain the pretrained neural network θ_0 by using Meta learning and conduct further training using the actual quantum-scar images. To be concrete, we fix $N = 6$ and vary K. For $K = 1$, we randomly choose two image from each class, where the first and second images are used for training and testing, respectively. The training and testing datasets are thus extremely small. To obtain the accuracy, we repeat this process (i.e., training θ_0) 1000 times and calculate the ratio of the number of correctly labeled images to the total number of test images in all the trials. The average ratio and the corresponding

variance are obtained by using ten realizations of the neural network θ_0. Table 5.1 lists the accuracy versus the shot number K, where a 90% accuracy is achieved even for $K = 1$. The somewhat lower accuracy than that of classifying the characters in the Omniglot dataset is due to the "noisy" nature of the images of quantum scars, where the scarring patterns are not sharp but have local fluctuations. As the shot number K increases, the scar-detection accuracy improves.

Another quantity characterizing the detection performance is the confusion matrix whose diagonal and off-diagonal elements are the probabilities of correct and incorrect detection of different classes of quantum scars, respectively. Figure 5.2 shows, for $K = 1$, the confusion matrix for the six classes of quantum scars from the chaotic heart billiard, where the Roman letters indicate the periods of the underlying classical periodic orbits and the arabic numbers represent distinct configurations of the orbits. For five out of the six classes (except type V-1), the detection accuracy is close to 100%. The accuracy for type V-I scars is somewhat lower, where occasionally the machine-learning algorithm would erroneously classify such a scar as belonging to class II or III. From Fig. 5.1(b), it can be seen that this error may be expected for two reasons: (1) the relatively long period of the corresponding classical periodic orbits leads to a relatively low concentration of the spinor wavefunction around the orbits and (2) the type V-1 scars have approximately the same edges with those in class II or III.

5.3 Results

We use two types of chaotic Dirac billiard systems to demonstrate the feasibility and power of Meta learning in detecting and classifying relativistic quantum scars.

5.3.1 Chaotic Heart Billiard

In Sec. 5.2.3, we articulate the idea of adapting the method of few-shot classification to detect quantum scars from a chaotic billiard. For a target set of six distinct scars in the chaotic heart billiard, the one-shot accuracy is close to 90%. Here we extend the method to finding desired quantum scars from a large number of wavefunctions.

There can be a large variety of spatial distributions of the eigenstates of a chaotic Dirac billiard. Visually, most eigenstates have a random distribution and they do not exhibit any discernible pattern. For those eigen wavefunctions that do, they can be quantum scars associated with periodic orbits of high periods or boundary states. From the point of view of machine-learning based detection of scars, such a dataset is unbalanced because the number of non-scarring states is much larger than that of the actual scarring states. In computer science, this problem associated with object detection is well known [374], where the number of "real" objects in the training dataset is extraordinarily small. An intuitive solution is to treat the non-scarring states as a new class of objects when implementing the few-shot classification algorithm, but there are limitations due to the unbalanced nature of the dataset.

We develop an alternative approach to addressing this issue of unbalanced dataset. After the CNN is properly trained, we input an additional quantum scarring pattern to the neural network. Ideally the prediction should be robust, meaning that the normalized probability should have a high concentration about this pattern. Since this non-scarring image does not belong to any of the classes that the neural network has learned from the training process, when other eigenstates are inputted to the neural network, the overall prediction probability will be lower. Another feature that can be exploited, as can be seen from Fig. 5.1(c), is that the variances of the output

probabilities with nonscarring states as the input are typically large. Our general quantum-scar detector is built on these two features.

For the chaotic heart billiard, the training data are the eigenwavefunctions generated at $\alpha = 0$, where two quantum scars are chosen for each training class: one at low energy ($n \approx 2000$) and another at high energy ($n \approx 10^4$). We perform one-shot classification and train the neural networks. Since each class contains two images, there are 2^6 possible neural networks. The testing data consist of eigenwavefunctions from $\alpha = 0.25$ or $\alpha = 0.5$. For a quantum state n inputted to the ith neural network, the output is denoted as p_j, where j is the scar index. We multiply the results from different neural networks to find the scarring state that maximizes the product:

$$\xi(n) = \max_j \left(\prod_i p_j^{(i)}(n) \right). \tag{5.5}$$

As an illustrative example, we consider two quantum states: $n = 8480$ (a scarring state) and $n = 8523$ (a nonscarring state) for $\alpha = 0.5$. The average and variance of p for the distinct scar indices are shown in Fig. 5.1(c). We use an ensemble of ten neural networks for the one-shot ($K = 1$) scheme. For the scarring state $n = 8480$, the predicted p value concentrates on the corresponding type of scar, where the product in Eq. (5.5) achieves maximum for the IV-1 scar class with $\xi(n = 8480) = 0.23$. For the nonscarring state, the p value does not concentrate at any scar index and the variance from different neural networks is large. We get $\xi(n = 8523) = 4.3 \times 10^{-5}$. There is then a four orders-of-magnitude difference in the ξ values for the scarring and nonscarring states, rendering accurate and reliable detection of the scarring state.

To detect the scarring states from all available quantum states for $\alpha = 0.5$, we use the ensemble of neural networks and sweep through the states. Figure 5.3(a) shows $\xi(n)$ versus n for $n \in [4000, 4100]$, where a nearly eight-orders-of-magnitude difference in the $\xi(n)$ values emerges among the quantum states. Setting a threshold,

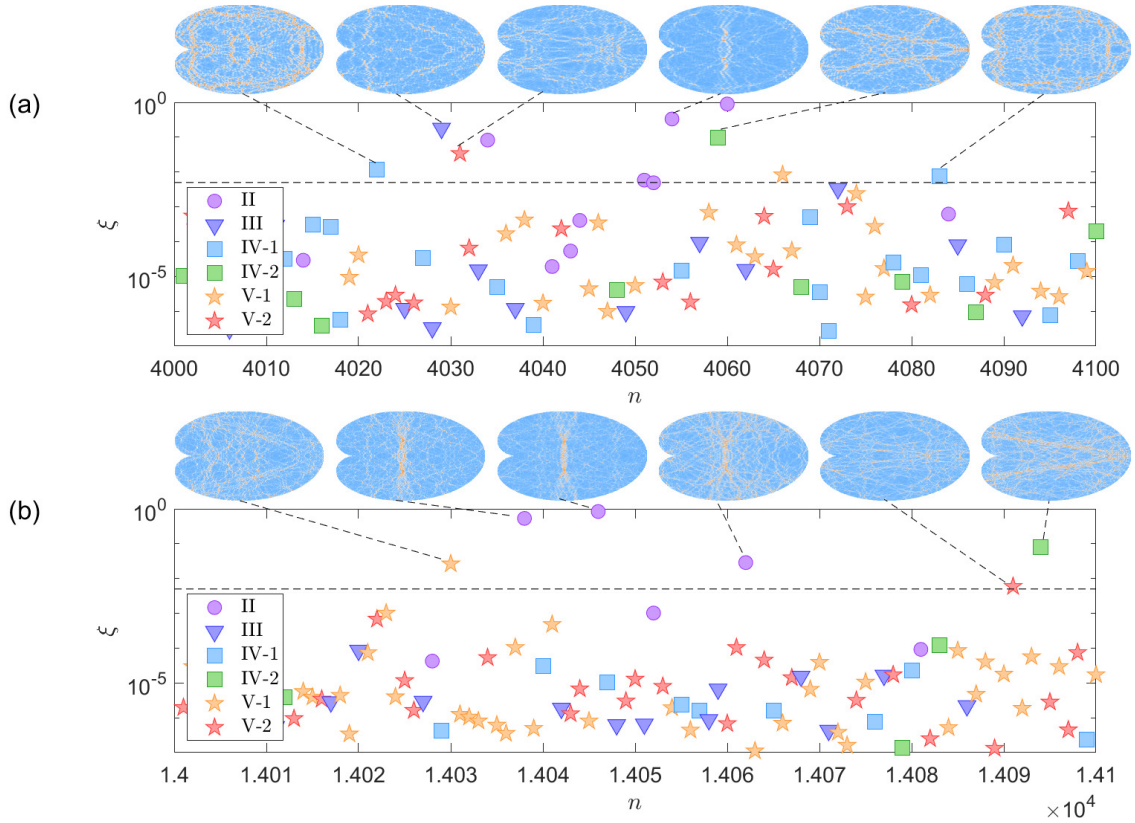

Figure 5.3: Emergence of a Meta-learning based quantum scar detector. Shown is $\xi(n)$ versus the eigenstate index n for the chaotic heart billiard for $\alpha = 0.5$. (a) The results are for $n \in [4000, 4100]$ (a relatively low energy interval), where the scar types are distinguished by the different colored symbols and the horizontal dashed line at 5×10^{-3} indicates the detection threshold. The eigenstates with $\xi(n)$ values above the threshold are deemed to be quantum scars. Some representative scarring states detected are displayed above the panel. The agreement between the detected states and the ground truth is perfect. (b) Same as (a) but for $n \in [14000, 14100]$ - a higher-energy interval.

e.g., at 5×10^{-3} (the horizontal dashed line), we deem those states whose $\xi(n)$ above the threshold as quantum scars, as verified by the corresponding distinct scarring patterns above Fig. 5.3(a). Similar results have been obtained for $n \in [14000, 14100]$, as shown in Fig. 5.3(b). The results in Fig. 5.3 thus indicate the viability of a Meta-learning quantum scar detector based on evaluating the values of $\xi(n)$.

To verify if the scarring states found are truly relativistic quantum scars, we exploit a criterion from semiclassical theory. In particular, the semiclassical theory

for relativistic quantum billiard systems [135, 208] predicts that the recurrent interval in the wavevector must obey the following rule:

$$\Delta k = \begin{cases} 2\pi(\Delta n - 2W\alpha)/L, & \text{even bounces,} \\ 2\pi(\Delta n - 2W\alpha + \Delta\beta)/L, & \text{odd bounces,} \end{cases} \tag{5.6}$$

where L is the length of the classical orbit associated with the scarring state and W is the winding number. For the scarring states corresponding to classical periodic orbits with an odd number of bounces or reflections from the billiard boundary, the quantity $\Delta\beta$ takes on the value of $1/2$, so the spacing in the wavevector for such scarring states to occur is π/L.

Table 5.2 presents the statistics of different types of quantum scarring states detected. The simulation setting is as follows. The training datasets are the same as those described in Sec. 5.2.3, which are associated with zero magnetic flux $\alpha = 0$. Ten neural networks are trained for one-shot classification. The test datasets are associated with $\alpha = 0.25$ and $\alpha = 0.5$ with states whose level index ranges from $n = 4000$ to $n = 15000$. The types of scarring states in Tab. 5.2 are those corresponding to classical periodic orbits II, III, and IV-1. The reference wavevector k_0 is chosen from a scarring state with the maximum value of $\xi(n)$ for $n \in [9000, 10000]$. For example, for orbits II, there are approximately 300 scarring states. Out of the $11,000$ available eigenstates in the chaotic heart billiard, approximately 5% are quantum scars. To verify that the machine-detected scarring states are true quantum scars, we select a few energy intervals and identify visually the scarring states. The percentage of the scarring states identified this way is consistent with that of the scarring states found by machine learning.

A convenient semiclassical quantity to characterize the recurrence of a quantum

Table 5.2: Statistics of quantum scars in the heart billiard system detected by machine learning

Scar index	Δk	k_0	Number
II ($\alpha = 0.25$)	1.4810	190.9758	271
II ($\alpha = 0.5$)	1.4810	199.8387	303
III ($\alpha = 0.25$)	1.1687	197.3900	61
III ($\alpha = 0.5$)	1.1687	199.4852	51
IV-1 ($\alpha = 0.25$)	1.0843	197.4324	26
IV-1 ($\alpha = 0.5$)	1.0843	192.9035	40

scarring state is the following winding number [135]

$$\eta = \frac{|k_n - k_0|}{\delta k} - \left[\frac{|k_n - k_0|}{\delta k}\right], \tag{5.7}$$

where $\delta k = 2\pi/L$, $[x]$ denotes the largest integer less than or equal to x, k_0 is the wavevector of a reference scarring state (usually of high energy). Figure 5.4 shows the value of η for a large number of scarring states of three types: type II, type III, and type IV-1, where type-II scars are the most abundant. As shown in Figs. 5.4(a) and 5.4(b) for $\alpha = 0.25$ and $\alpha = 0.5$, respectively, the winding number of type-II scars are either zero or one are zero. For this type of scars, because of the even bounce numbers of the underling classical orbits off the billiard wall, it is not possible for the quantum states to have $\eta = 1/2$. For type-III scars, the winding numbers are zero or one are one for $\alpha = 0.25$, as shown in Fig. 5.4(c). However, because of their odd bounces with the billiard wall, scarring states with $\eta = 0.5$ can arise for $\alpha = 0.5$, as shown in Fig. 5.4(d). For type-IV-1 scars, $\eta = 0.5$ occurs for magnetic flux value $\alpha = 0.25$ but not for $\alpha = 0.5$, as shown in Figs. 5.4(e) and 5.4(f), respectively. These features of the relativistic quantum scars have been understood theoretically based on a detailed analysis of the geometric phases [208]. It is remarkable that our Meta-learning based

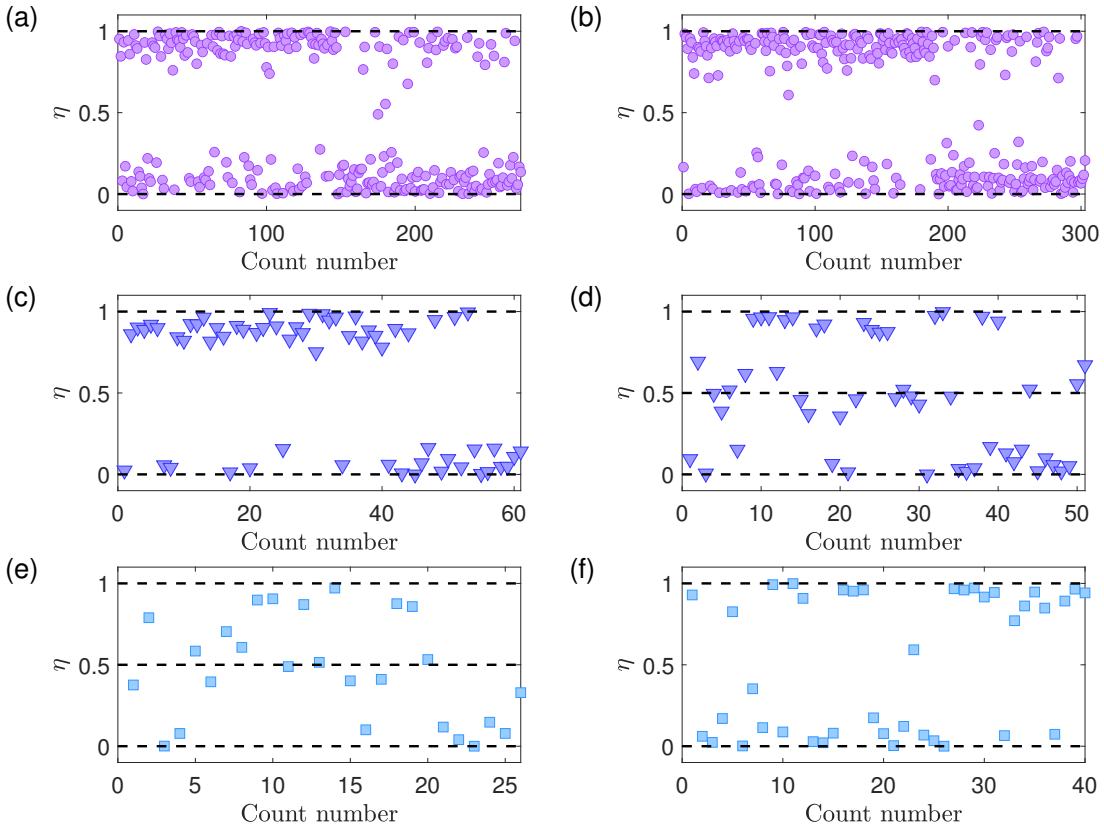

Figure 5.4: Semiclassical characterization of the relativistic quantum scarring states detected by the Meta-learning algorithm. Shown are the values of the semiclassical winding number η for three types of scars, where the top, middle, and bottom rows correspond to type-II, type-III, and type-IV-1 scars, respectively. A magnetic flux is present: $\alpha = 0.25$ (the left column) and $\alpha = 0.5$ (right column). A geometric-phase based semiclassical theory [208] stipulates that for all types of scars, the values of η can be zero or one, and $\eta = 0.5$ is not possible for type-II scars, but can occur for type-III (type IV-1) scars under magnetic flux $\alpha = 0.5$ ($\alpha = 0.25$). The various quantum scars detected by the machine-learning algorithm obey these rules remarkably well, signifying high detection accuracy.

algorithm can reliably detect the different types of the scarring states with the correct semiclassical characteristic features.

A few remarks are in order. First, from Fig. 5.3, it can be seen that most detected quantum scars are associated with classical periodic orbits of low periods, due to the fact that these scarring states are more pronounced than those corresponding to periodic orbits of high periods. This is consistent with the semiclassical theory of

quantum scars [337, 338].

Second, we use an ensemble of neural networks to combine the accuracy measures, where each is trained based on Meta-learning with one-shot classification. The training thus requires more than one image per class. For two images per class, we can generate 2^6 combinations so the training data is sufficient for scar detection. In simulations, we find that choosing two scars with one at relatively low energy and another at high energy can help diversify the neural-network outputs to increase the detection accuracy.

Third, Fig. 5.3 demonstrates the working of the scar detector in the energy interval with the level index ranging from 4000 to 15000. In a lower energy interval, e.g., $n \approx 1000$, the degree of quantum scarring in terms of the wavefunction concentration is relatively weak. Especially, the difference in the values of $\xi(n)$ between scarring and nonscarring states is small, making detecting scars difficult. However, this is expected as quantum scars are more pronounced in the relatively high energy regime, the so-called semiclassical regime where both quantum and classical behaviors are relevant [337, 338].

Fourth, an advantage of the neural-network based scar detector is that all types of scars can be detected based on a single threshold, as exemplified in Fig. 5.3. A large threshold means that a small number of states can be found with high accuracy. Decreasing the threshold can lead to more "scarring" states but the accuracy may be compromised. In principle, we can still pick only those states with the highest ξ values as the scarring states.

5.3.2 The Africa Billiard

The Africa billiard system has fully chaotic dynamics in the classical limit and exhibits relatively more complicated quantum scarring states than the chaotic heart

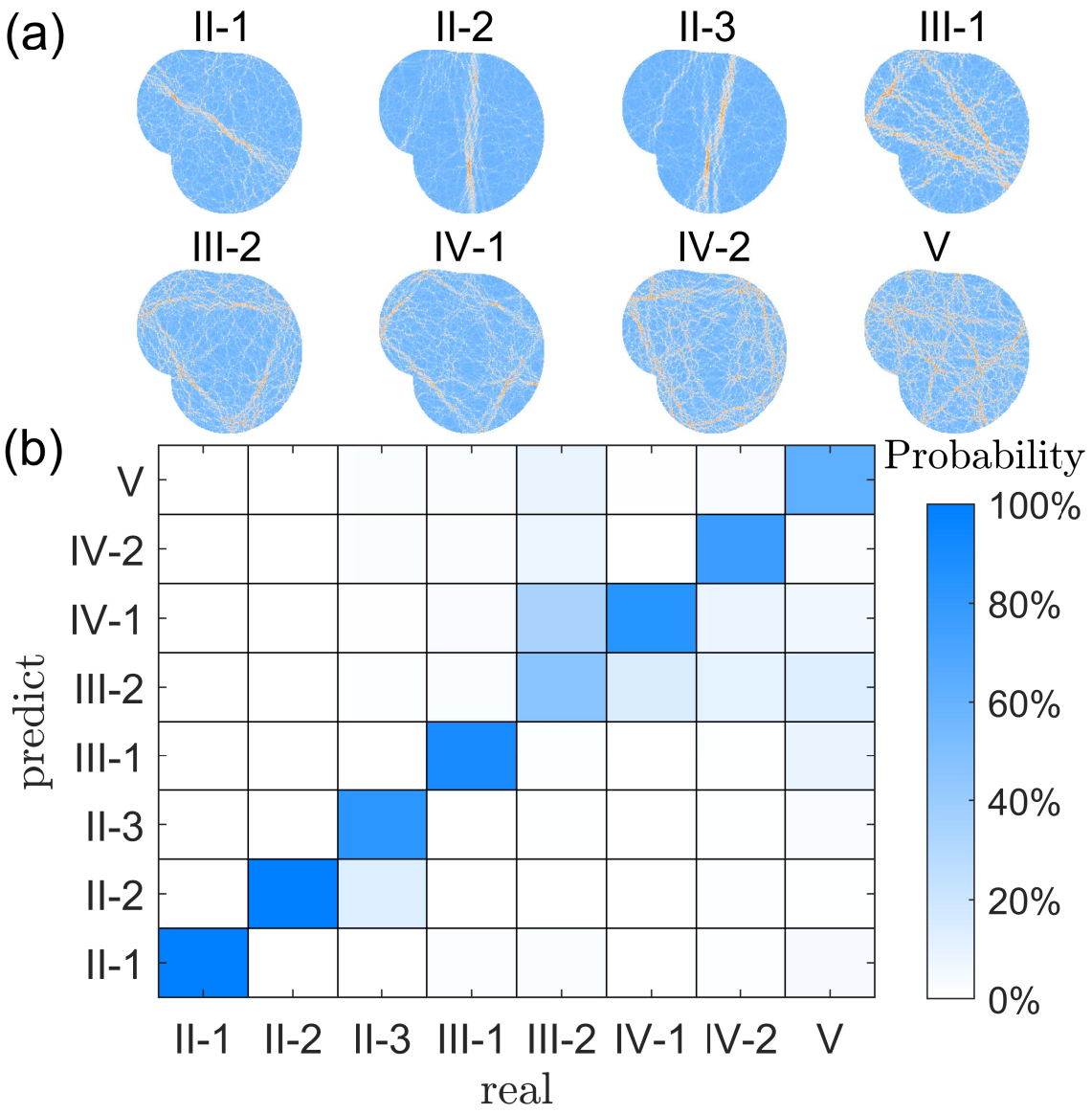

Figure 5.5: Types of quantum scarring states in the chaotic Africa billiard and confusion matrix of classification. (a) Eight different types of scars in the Africa billiard and (b) confusion matrix between the true and predicted labels from one shot classification. The diagonal elements have large values, indicating high overall accuracies of the predictions of the neural network. Some off diagonal elements are non-negligible, e.g., those between type-II-2 and type-II-3 scars.

billiard. Figure 5.5(a) shows eight types of scars. We extend the few-shot classification algorithm by setting $N = 8$ and selecting 10 images for each class for $\alpha \in [0, 0.1]$ from our quantum-scar dataset. By varying K and repeating the process 1000 times, we obtain the average accuracy, as listed in Tab. 5.3. Due to the large number of

Table 5.3: Few shot classification accuracy for scarring states in the chaotic Africa billiard

Shot number	Accuracy
$K = 1$	$78.14\% \pm 1.64\%$
$K = 2$	$85.44\% \pm 1.15\%$
$K = 5$	$93.00\% \pm 0.96\%$

Table 5.4: Characteristics of representative relativistic quantum scars in the Africa billiard

Scar index	Δk	k_0	Collected number
II-2	1.6558	198.1712	41
II-3	1.6427	195.5084	89
III-1	1.2313	191.8086	59
IV-2	1.0699	191.5694	25

scarring states and the relatively more sophisticated geometric shape of the Africa billiard (than the chaotic heart billiard), the accuracy for one shot classification is somewhat lower: less than 80%. The accuracy can be improved by increasing the number of images for training.

Figure 5.5(b) shows the confusion matrix under one-shot classification, where the large diagonal values indicate that most predictions are correct. Nonetheless, some off-diagonal values are not negligible. For example, the algorithm appears to be "easily confused" by type-II-2 and type-II-3 scars, because the concentration patterns of their spinor wavefunctions are similar, as shown in Fig. 5.5(a). In fact, the different kinds of type-II scars all have a straight line pattern and this leads to the confusion. We also note that the accuracy for type-III-2 scars is the lowest, due to their similarity to the type-V scars.

We build a scar detector based on two-shot classification, due to its reasonable

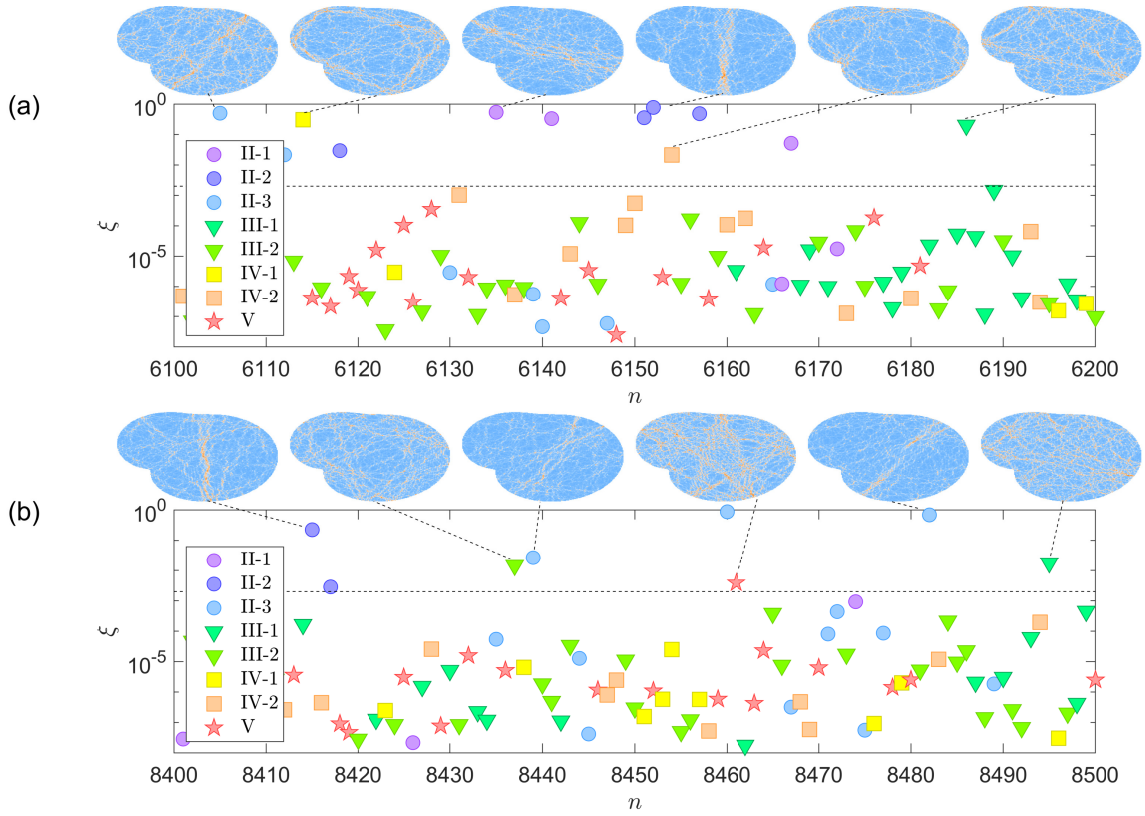

Figure 5.6: Machine-learning detection of quantum scars in the chaotic Africa billiard. Shown is ξ versus the mode index for $\alpha = 0.25$ for (a) $n \in [6100, 6200]$ and (b) $n \in [8400, 8500]$, where the horizontal dashed line indicates the detection threshold 2×10^{-3}. Points above the threshold correspond to the detected scarring states with some representatives shown above the panel.

accuracy, as shown in Tab. 5.3. The number of images per class in the training set is set to be three and the scarring states used for training are associated with the magnetic flux in the interval $\alpha \in [0, 0.1]$. Figure 5.6 shows $\xi(n)$ for $n \in [6100, 6200]$ and $\alpha = 0.25$. We set the threshold to be 2×10^{-3} and those quantum states whose ξ values are larger than the threshold are deemed as scarring states. The detection results for $n \in [8400, 8500]$ are shown in Fig. 5.6(b). As in the case of the chaotic heart billiard, most detected quantum scars correspond to classical periodic orbits of low periods. We then calculate the η values for each type of scars from Eq. (5.7), where k_0 is chosen to be associated with the eigenstate with the largest $\xi(n)$ value

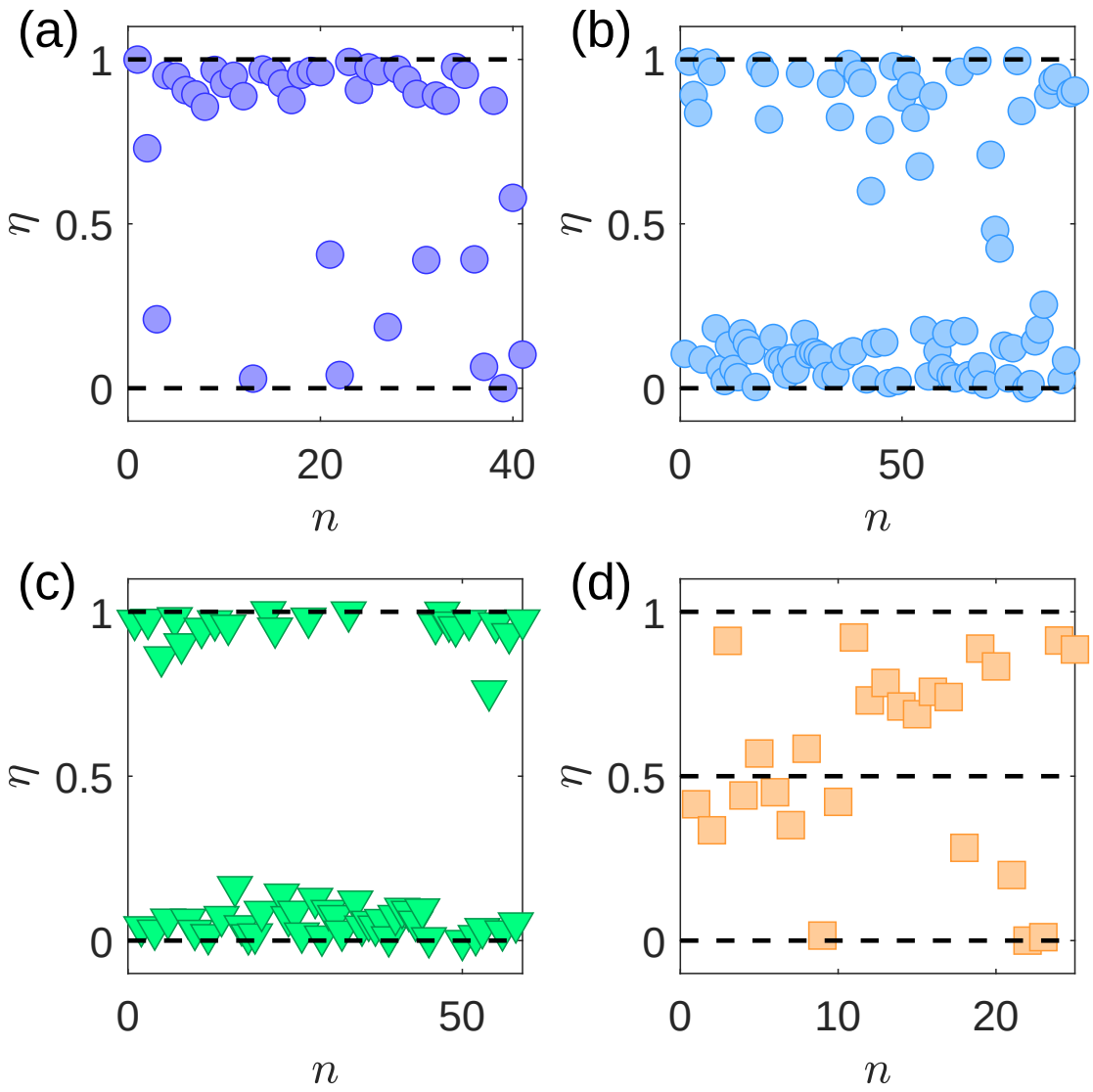

Figure 5.7: Semiclassical quantification of the machine-learning detected relativistic quantum scars in the chaotic Africa billiard. Shown in the semiclassical winding number η for scars of types (a) II-2, (b) II-3, (c) III-1, and (d) blue IV-2. Due to the presence of a magnetic flux, only blue IV-2 scars can have $\eta = 0.5$. The statistical behaviors of the detected quantum scars agree well with the semiclassical predictions.

for $n \in [9000, 10000]$. The results for some representative scars detected from the energy-level range $n \in [4000, 10000]$ are shown in Fig. 5.7 with their basic properties listed in Tab. 5.4. For all scars displayed (for $\alpha = 0.25$), the η values are either zero or one, except the type-IV-2 scars where $\eta = 1/2$ can occur, which agrees well with

the semiclassical prediction [208].

5.4 Discussion

In spite of the fundamental importance of the phenomenon of quantum scarring, identifying such states from a large number of eigenstates has been a laborious task since the beginning of the field of quantum chaos, relying mostly on going through all the available wavefunction patterns one after another and performing a visual check of each pattern. This task has been well known to practitioners of quantum chaos. To develop a fully automated method to accurately detect quantum scars is thus highly desired, as it will enable a more systematic and comprehensive study of the exotic quantum states. This paper accomplishes this goal by developing a machine-learning based quantum-scar detector. While it may seem straightforward to adopt some conventional deep neural networks for image recognition to the problem of quantum scar detection, a significant challenge lies in the requirement of extensive training data, as quantum scattering states are rare. We have met this challenge by articulating a Meta-learning approach based on few-shot classification of quantum scarring states, which allows the neural machine to learn the hidden structure for each class and quickly adapt to the new class even with a quite small number of images of the scarring states. We have tested this approach using two paradigmatic relativistic quantum billiard systems subject to a magnetic flux, which exhibit fully developed chaos in the classical limit. Using a previously developed conformal-mapping method [135, 208] to calculate a large number of quantum states in each system, we have demonstrated the power of our Meta-learning method for accurate detection of quantum scars.

Our fully automated quantum-scar detection framework consists of two major steps. The first step is to use neural networks to classify quantum scars. In particular, we apply few-shot Meta learning by designating a number of images per class in

the training dataset, e.g., a few different classes of quantum scarring states in a chaotic Dirac billiard. For one-shot classification, the achieved accuracy can already be about 80%, which can be improved by increasing the number of shots. The overall classification accuracy can be characterized by the confusion matrix. In general, the accuracies for quantum scars associated with classical periodic orbits of short periods are higher than those with longer periods. The next step is to exploit the few shot classification algorithm to detect quantum scars by monitoring the predicted probability difference between scarring and non-scarring states. For quantum scars, due to the high few-shot classification accuracy, the probability distributions are well localized and large with small variances, but the opposite occur for non-scarring states. Utilizing an ensemble of neural networks allows us to define a statistical measure, whereas a quantum eigenstate can be faithfully deemed as a scarring state if this measure exceeds a well defined threshold. Utilizing a quantity derived from the semiclassical theory, we have demonstrated that the quantum scars detected are the correct scarring states. Our Meta-learning based quantum-scar detector is thus accurate, efficient, and fully automated.

We discuss a few pertinent issues. First, to build a general machine-learning based quantum-scar detector, it is necessary to choose some representative quantum scars as the training dataset. The training accuracy can depend on the specific scarring states, e.g., whether they are from the same or different energy regimes. For example, for the chaotic heart billiard system, two scars per class are needed for training. We find that, if one scar is chosen from the low-energy regime and another from the high-energy regime, high accuracy can be achieved. The intuitive reason is that quantum scarring states from drastically energy regimes tend to be maximally distinct in their patterns, thereby enhancing the learning capability of the neural network.

The second issue concerns about the number of images per class required for the

few-shot classification algorithm. Accuracy can be improved by increasing the number of such images but the training dataset also becomes larger, thereby increasing the computational complexity. Our empirical experience is that the number of images should be chosen such that the classification accuracy is about 90%. For example, for the chaotic heart billiard with six classes of quantum scars, using one image per class in the training dataset already suffices. For the chaotic African billiard, two-shot classification is needed to achieve the desired accuracy. For an ensemble of neural networks, various combinations of the training data are needed. As a result, the number of images per class should be at least one more than the shot number.

The third issue is about the non-scarring states. An alternative idea is to treat them as another class of "scars." However, non-scarring quantum states tend to be uniform and typically do not possess any unique feature. Treating them as a class of scarring states can lead to large classification errors.

The fourth issue is about overfitting in the Meta-learning algorithm. In our work, the neural networks are trained based on the Omniglot dataset and transferred to quantum scars. If a neural network performs well in Omniglot dataset, when executing transfer learning to quantum scarring states, overfitting can arise, reducing the classification accuracy. Our empirical method is to monitor the test error to ensure that it does not exceed a certain (small) threshold.

The fifth issue is about the necessity to use machine learning for automated detection of quantum scars. Can some filtering techniques be used, e.g., based on counting the probability value of the quantum wavefunction along a classical periodic orbit? This is in general infeasible because, from an image processing point of view, quantum scars are extremely noisy. Moreover, setting a proper region to estimate the probabilities requires a threshold, which can be difficult as the degree of localization of the scarring wavefunction depends on the energy. Even worse, the wavefunction

associated with a quantum scarring state may not follow a classical periodic orbit exactly. To our knowledge, exploiting machine learning for quantum-scar detection represents the best strategy at the present.

Taken together, in the field of quantum chaos, finding quantum scars usually relies on human visualization. To develop a machine-learning framework typically requires extensive training data, but quantum scarring states are typically rare, posing a significant challenge. Taking advantage of Meta learning, a special class of machine learning for image recognition and classification, we have developed a fully automated quantum scar detector and demonstrated its working with a remarkably small number of scar images for training. Our detection algorithm requires no fine tuning of the parameters of the neural networks and is effective in a broad energy range that contains many thousands of eigenstates. The framework developed in this paper can be readily generalized and extended to solving difficult image processing problems in other disciplines of science and engineering.

5.5 Data Processing

We use the Omniglot dataset to obtain the basic neural network before training with quantum scars, which contains more than one thousand handwritten characters from different languages [162]. While some quantum scars resemble certain characters, their details can be quite different. The left panel in Fig. 5.8(a) shows the Greek character α in the Omniglot dataset, which has a well-defined structure in that each stroke has a similar width. The left panel of Fig. 5.8(b) shows a period-4 quantum scar, which differs significantly from the α character and is noisy. To use the Omniglot dataset to train the neural networks for detecting quantum scars, image pre-processing is necessary to reduce the difference in details.

An image is represented by a matrix defined on a grid of pixels, so a white pixel

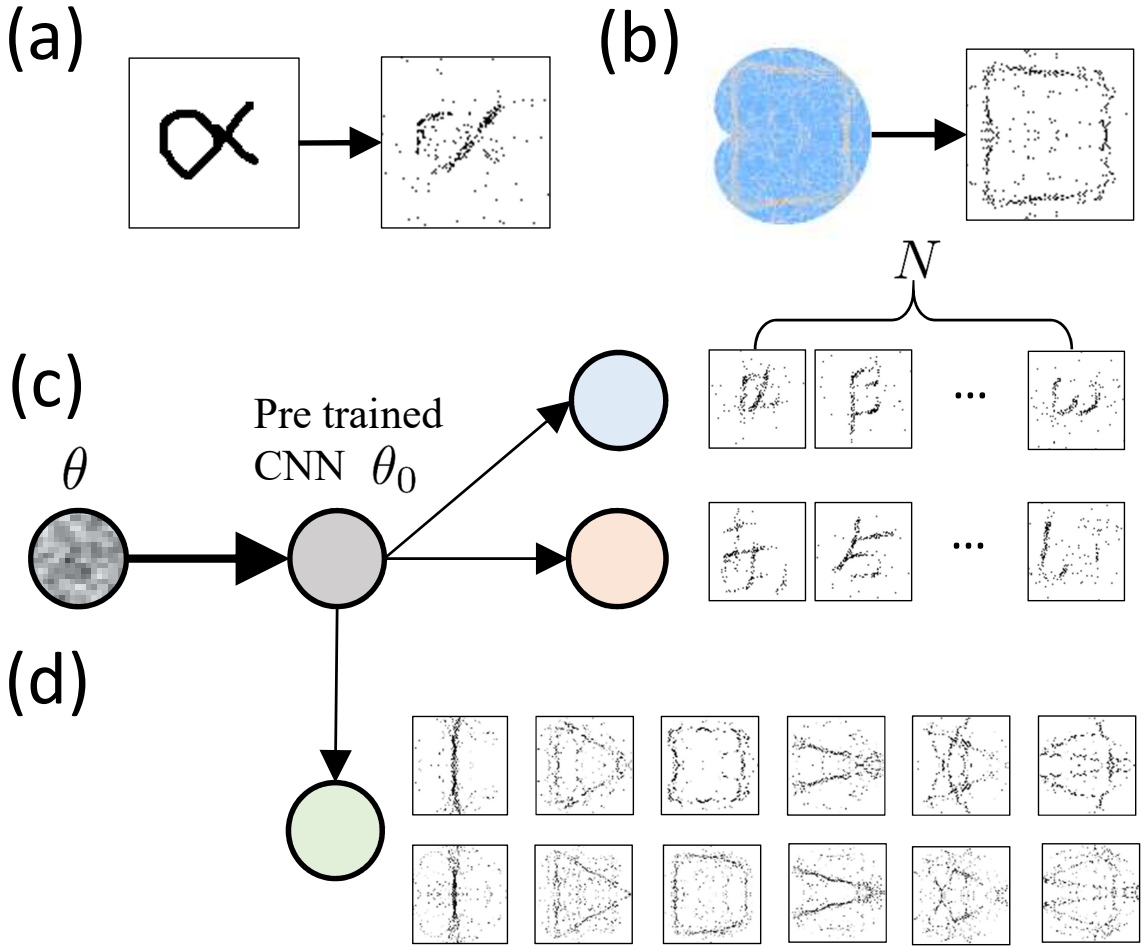

Figure 5.8: Schematic illustration of image processing. (a) Left panel: a representative image in the Omniglot dataset with the Greek character α of size 105×105 pixels. Right panel: the image after the transform. (b) Left: a quantum scar image of size 202×234 from the chaotic heart billiard. Right: the image after data processing and resizing. With image processing, the detailed patterns the Omniglot and quantum scar images become similar. (c,d) Two steps of the Meta learning algorithm, respectively. The first step (c) is to start from a neural network with random weights and biases θ to find the optimal network denoted as θ_0 that quickly adapts to randomly selected classes in the Omniglot dataset. The second step (d) is to perform training with quantum scars. Due to the similarity between the processed Omniglot character and quantum scar images, adaptation of the neural network θ_0 to quantum scars can be achieved with only a small number of images.

can be defined as zero and a black pixel is represented by one. Our data processing for each image in the Omniglot dataset consists of the following steps: (i) converting an image to a matrix of dimension 108×108, (ii) adding a uniform noise between

0 to 1 for each element, (iii) smoothing the image with low-frequency filtering, (iv) setting the maximum 4% of points in the matrix as black and others as white, and (v) using the PIL package in Python to resize the image to 100×100. For low-pass filtering, we use Fast Fourier Transform (FFT) to transform the matrix into a coefficient vector, one component for each frequency, set the coefficients for the top 10% of the frequency range to zero, then perform the inverse FFT. To process the quantum scarring images, we first calculate the density $\psi = \psi_1^2 + \psi_2^2$ of the spinor wavefunction, assuming that the wavefunction outside the cavity is zero. We then set the points with the top 4% values of the density to be one (black) and others to zero (white), and use the PIL package in Python to resize the image to 100×100. The right panels in Figs. 5.8(a) and 5.8(b) show the images after processing for the α character and the period-4 quantum scars, respectively. It can be seen that the processing has resulted in similar patterns for the Omniglot character and quantum scar. A heuristic reason for this similarity is that the neural network trained with the Omniglot dataset performs classification by extracting line segments as features. With image processing, both the Omniglot characters and quantum scars contain line segments as features.

5.6 Meta Learning Algorithm

There are two categories of Meta learning algorithms: metric [364] and optimization [161, 373] based. We adopt the optimization-based approach, where θ denotes the training parameter set in the convolutional neural network. We aim to find a pre-trained neural network, denoted as θ_0, such that the loss is minimized after operation $U_\tau^k(\theta)$ - k steps of gradient descent based on task τ. Let S be the cross entropy loss. The goal can then be formulated as $\min_\theta S[U_\tau^k(\theta)]$ for all τ.

Different methods are available to solve the optimization problem. One is Model-

Agnostic Meta-Learning (MAML) [161], which can give higher accuracy. The method requires computing the gradient for the iterative equation $U^k(\theta)$, which is time consuming when $k > 1$. An approximate method is Reptile [373]. For the Omniglot and Mini-imagenet datasets, Reptile is computationally efficient but the accuracy is slightly compromised in comparison with that of MAML.

Figure 5.8(c) and 5.8(d) illustrate, respectively, the two basic steps involved in our Meta-learning procedure. Starting from a neural network with randomly generated weights and biases denoted as θ, we use the images from the Omniglot dataset to train the network to find an optimal network denoted as θ_0, as shown in Fig. 5.8(c). In the second step [Fig. 5.8(d)], we start from θ_0 and perform training with quantum scars. Since the Omniglot dataset is large, θ_0 that has been trained to adapt to any image class in the dataset, can quickly adapt to quantum scars.

Algorithm 1 Reptile - batched version

Initialized θ

for iteration = 1, 2, ... **do**

  Sample task τ_1, τ_2, ..., τ_m

  for i = 1,2, ... **do**

    compute $W_i = U^k_{\tau_i}(\theta)$

  end for

  Update: $\theta \leftarrow \theta + \beta \frac{1}{m} \sum_{i=1}^{m} (W_i - \theta)$

end for

The Reptile method is illustrated in Algorithm 1. Initially, all the weights and biases are randomly chosen. A loop is employed to update θ until the desired neural network θ_0 is found. In each epoch, the network θ is updated, as follows. The first step is to randomly generate N classes from the processed Omniglot dataset, where each class contains K images, so the labeled dataset has NK images. The second step

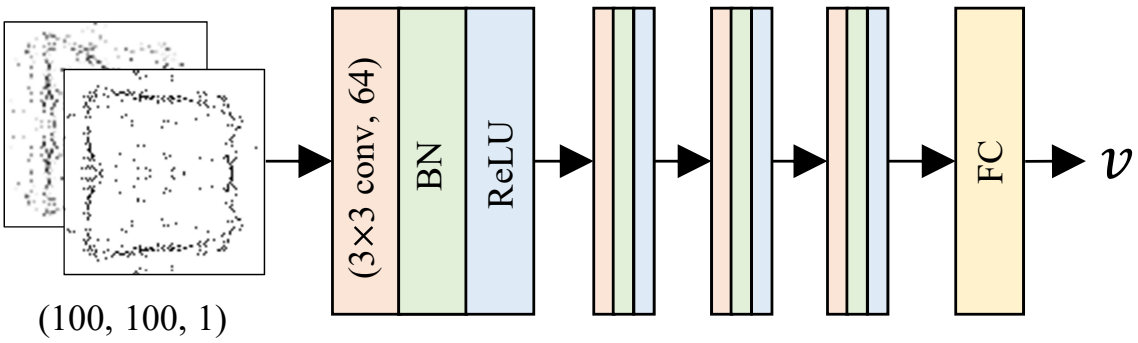

Figure 5.9: Detailed structure of the convolutional neural network (CNN) used in our study. The CNN contains four convolutional layers, each layer having a 3×3 convolutional kernel with stride equal to 2 and zero padding, followed by ReLU and batch normalization.

is to generate five tasks, each containing ten images that are randomly chosen and can be repeated. The third step is to perform stochastic gradient descent (SGD) and Adam for each task. Let $U^k(\theta)$ denote the stochastic gradient updating for k steps on cross-entropy loss with the initial parameter set θ, which returns the final parameter set. During this training process, we set $k = 5$. The fourth step is to update θ based on the average. The learning rate β decays linearly with the number of epochs, e.g., $\beta = 10^{-3}$ for the first epoch and $\beta = 0$ at the last epoch. After each epoch, θ is saved and SGD is performed for quantum scar images. The testing accuracy can then be calculated. The cutoff point can be found based on hyperparameter optimization as described in Appendix 5.8.

5.7 Neural Network Architecture

The Convolutional neural network used in our study is illustrated in Fig. 5.9, which contains four convolutional layers. Each layer is followed by Batch Normalization (BN) to avoid overfitting [375]. We use ReLU as the nonlinear activation function, which is applied after the batch normalization. Each convolutional layers contains 64 kernels of size 3×3. When applying convolutional kernels, we use stride 2 and zero

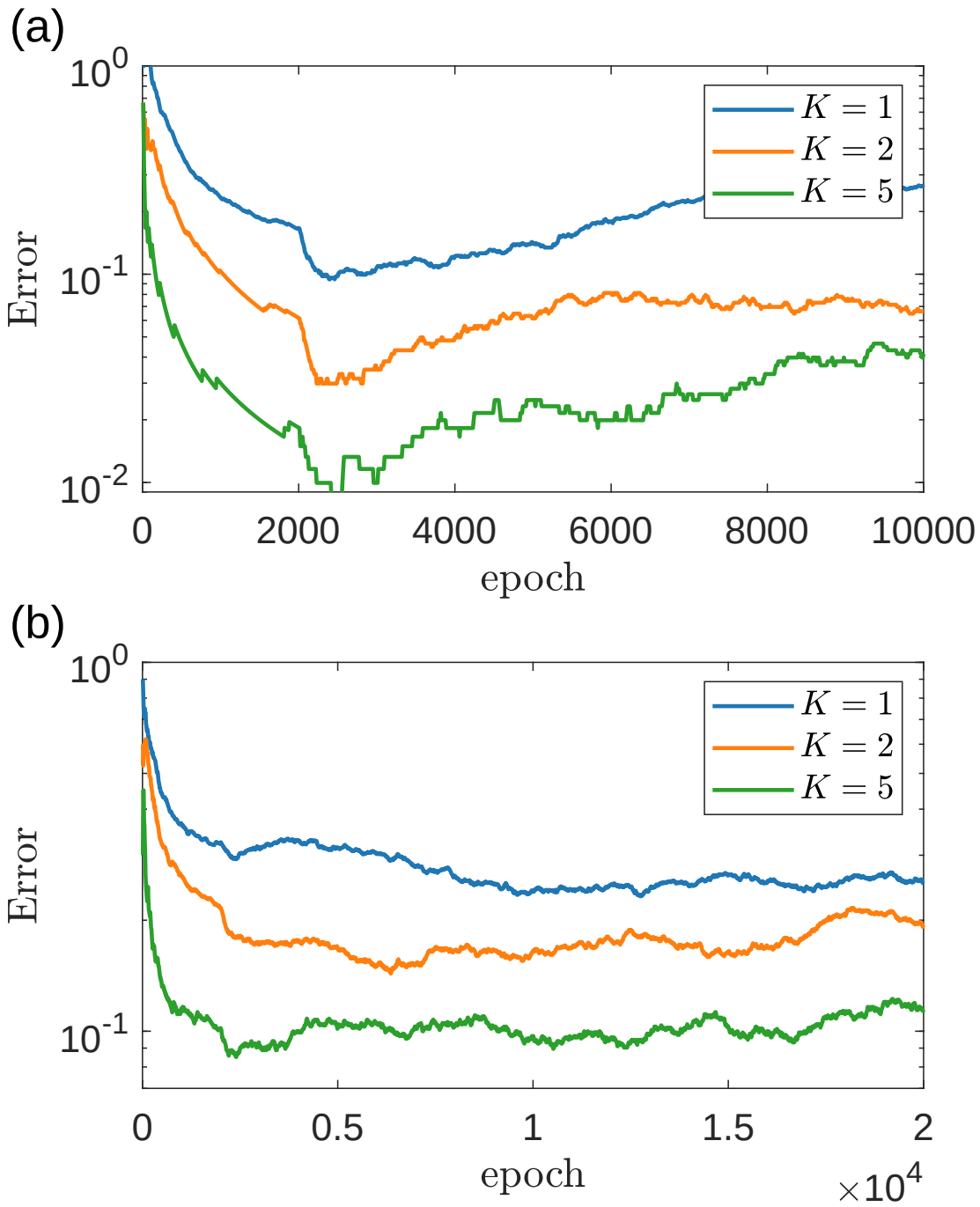

Figure 5.10: Few shot classification errors versus the number of training epochs. The error is the average from a moving window of 200 epochs. (a,b) The results for the quantum scars from the chaotic heart and Africa billiards, respectively. The cutoff point is determined by the empirical criterion that the testing error reaches the minimum. Further updating the neural network can lead to overfitting.

padding. The whole neural network contains approximately 10^5 training parameters.

5.8 Hyperparameter Optimization

Since the neural network with the parameter set θ is optimized with the processed Omniglot dataset, too many training steps can lead to overfitting for the Omniglot dataset and reduce the training accuracy for quantum scars. Our solution is to monitor the test error at each epoch. Specifically, for each epoch, we save θ, perform a few-shot classification for quantum scars, and calculate the classification error. The error can have large fluctuations due to the small quantum-scar testing set, so we take the average over 200 steps. Figure 5.10(a) and 5.10(b) show the classification error versus the number of epochs for the quantum scars in the chaotic heart and African billiards, respectively. For quantum scars from the heart billiard, the number of classes is 6, the total number of iteration is 10^4, and the error reaches a minimum for about 2000 steps for all K values. The cutoff point is then set to be 2000. For quantum scars from the Africa billiard, the number of classes is 8, the total number of iterations is 2×10^4, and the cutoff points for $K = 1, 2$ and 5 are 10^4, 5000 and 2000, respectively. When performing the hyperparameter optimization, the number of available quantum scars is also limited. In our computations, the number of images per class in the quantum scar dataset is fixed to be $K+1$ to achieve data consistency.

[1] Martin C Gutzwiller. *Chaos in Classical and Quantum Mechanics*. Springer-Verlag, New York, 1990.

[2] H.-J. Stöckmann. *Quantum Chaos: An Introduction*. Cambridge University Press, New York, 1999.

[3] F. Haake. *Quantum Signatures of Chaos*. Springer series in synergetics. Springer-Verlag, Berlin, 3rd edition, 2010.

[4] Oriol Bohigas, Marie-Joya Giannoni, and Charles Schmit. Characterization of chaotic quantum spectra and universality of level fluctuation laws. *Phys. Rev. Lett.*, 52(1):1–4, 1984.

[5] M. V. Berry. Semiclassical theory of spectral rigidity. *Proc. R. Soc. London Series A Math. Phys. Eng. Sci.*, 400(1819):229–251, 1985.

[6] Michael Victor Berry and RJ Mondragon. Neutrino billiards: time-reversal symmetry-breaking without magnetic fields. *Proc. R. Soc. Lond. A*, 412(1842):53–74, 1987.

[7] S. W. Mcdonald and A. N. Kaufman. Spectrum and eigenfunctions for a Hamiltonian with stochastic trajectories. *Phys. Rev. Lett.*, 42(18):1189–1191, 1979.

[8] E. J. Heller. Bound-state eigenfunctions of classically chaotic Hamiltonian systems - scars of periodic orbits. *Phys. Rev. Lett.*, 53(16):1515–1518, 1984.

[9] R. A. Jalabert., H. U. Baranger, and A. D. Stone. Conductance fluctuations in the ballistic regime - a probe of quantum chaos. *Phys. Rev. Lett.*, 65(19):2442–2445, 1990.

[10] Shmuel Fishman, D. R. Grempel, and R. E. Prange. Chaos, quantum recurrences, and Anderson localization. *Phys. Rev. Lett.*, 49:509–512, Aug 1982.

[11] Ying-Cheng Lai, Hong-Ya Xu, Liang Huang, and Celso Grebogi. Relativistic quantum chaos: an emergent interdisciplinary field. *Chaos*, 28(5):052101, 2018.

[12] Liang Huang, Hong-Ya Xu, Celso Grebogi, and Ying-Cheng Lai. Relativistic quantum chaos. *Phys. Rep.*, 753:1–128, 2018.

[13] A. H. Castro Neto and K. Novoselov. Two-dimensional crystals: beyond graphene. *Mater. Exp.*, 1:10–17, 2011.

[14] P. Ajayan, P. Kim, and K. Banerjee. Two-dimensional van der Waals materials. *Phys. Today*, 69(9):38–44, 2016.

[15] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov. Electric field effect in atomically thin carbon films. *Science*, 306(5696):666–669, 2004.

[16] C. Berger, Z. M. Song., T. B. Li., X. B. Li, A. Y. Ogbazghi, R. Fengand Z. T. Dai, A. N. Marchenkov., E. H. Conrad, P. N. First, and W. A. de Heer. Ultrathin epitaxial graphite: 2D electron gas properties and a route toward graphene-based nanoelectronics. *J. Phys. Chem. B*, 108(52):19912–19916, 2004.

[17] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, and A. A. Firsov. Two-dimensional gas of massless Dirac fermions in graphene. *Nature*, 438(7065):197–200, 2005.

[18] Y. B. Zhang, Y. W. Tan, H. L. Stormer, and P. Kim. Experimental observation of the quantum Hall effect and Berry's phase in graphene. *Nature*, 438(7065):201–204, 2005.

[19] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim. The electronic properties of graphene. *Rev. Mod. Phys.*, 81(1):109–162, 2009.

[20] N. M. R. Peres. Colloquium: The transport properties of graphene: An introduction. *Rev. Mod. Phys.*, 82(3):2673–2700, 2010.

[21] S. Das Sarma, S. Adam, E. H. Hwang, and E. Rossi. Electronic transport in two-dimensional graphene. *Rev. Mod. Phy.*, 83(2):407–470, 2011.

[22] M. Z. Hasan and C. L. Kane. Colloquium: Topological insulators. *Rev. Mod. Phys.*, 82:3045–3067, Nov 2010.

[23] R. Yang, L. Huang, Y.-C. Lai, and C. Grebogi. Quantum chaotic scattering in graphene systems. *Europhys. Lett.*, 94:40004, 2011.

[24] R. Yang, L. Huang, Y.-C. Lai, and L. M. Pecora. Modulating quantum transport by transient chaos. *Appl. Phys. Lett.*, 100(9):093105, 2012.

[25] R. Yang, L. Huang, Y.-C. Lai, C. Grebogi, and L. M. Pecora. Harnessing quantum transport by transient chaos. *Chaos*, 23(1):013125, 2013.

[26] M. Büttiker, Y. Imry, and R Landauer. Josephson behavior in small normal one-dimensional rings. *Phys. Lett. A*, 96:365¡§C–367, July 1983.

[27] Ho-Fai Cheung, Eberhard K. Riedel, and Yuval Gefen. Persistent currents in mesoscopic rings and cylinders. *Phys. Rev. Lett.*, 62:587–590, Jan 1989.

[28] Albert Schmid. Persistent currents in mesoscopic rings by suppression of charge fluctuations. *Phys. Rev. Lett.*, 66:80–83, Jan 1991.

[29] H Bouchiat. New clues in the mystery of persistent currents. *Physics*, 1:7, July 2008.

[30] L. P. Lévy, G. Dolan, J. Dunsmuir, and H. Bouchiat. Magnetization of mesoscopic copper rings: Evidence for persistent currents. *Phys. Rev. Lett.*, 64:2074–2077, Apr 1990.

[31] V. Chandrasekhar, R. A. Webb, M. J. Brady, M. B. Ketchen, W. J. Gallagher, and A. Kleinsasser. Magnetic response of a single, isolated gold loop. *Phys. Rev. Lett.*, 67:3578–3581, Dec 1991.

[32] D. Mailly, C. Chapelier, and A. Benoit. Experimental observation of persistent currents in GaAs-AlGaAs single loop. *Phys. Rev. Lett.*, 70:2020–2023, Mar 1993.

[33] W. Rabaud, L. Saminadayar, D. Mailly, K. Hasselbach, A. Benoît, and B. Etienne. Persistent currents in mesoscopic connected rings. *Phys. Rev. Lett.*, 86:3124–3127, Apr 2001.

[34] N. A. J. M. Kleemans, I. M. A. Bominaar-Silkens, V. M. Fomin, V. N. Gladilin, D. Granados, A. G. Taboada, J. M. García, P. Offermans, U. Zeitler, P. C. M. Christianen, J. C. Maan, J. T. Devreese, and P. M. Koenraad. Oscillatory persistent currents in self-assembled quantum rings. *Phys. Rev. Lett.*, 99:146808, Oct 2007.

[35] A. C. Bleszynski-Jayich1, W. E. Shanks, B. Peaudecerf, E. Ginossar, F. von Oppen, L. Glazman, and J. G. E. Harris. Persistent currents in normal metal rings. *Science*, 326:272–275, Oct 2009.

[36] Hendrik Bluhm, Nicholas C. Koshnick, Julie A. Bert, Martin E. Huber, and Kathryn A. Moler. Persistent currents in normal metal rings. *Phys. Rev. Lett.*, 102:136802, Mar 2009.

[37] M. A. Castellanos-Beltran, D. Q. Ngo, W. E. Shanks, A. B. Jayich, and J. G. E. Harris. Measurement of the full distribution of persistent current in normal-metal rings. *Phys. Rev. Lett.*, 110:156801, Apr 2013.

[38] Ho-Fai Cheung, Yuval Gefen, Eberhard K. Riedel, and Wei-Heng Shih. Persistent currents in small one-dimensional metal rings. *Phys. Rev. B*, 37:6050–6062, Apr 1988.

[39] Felix von Oppen and Eberhard K. Riedel. Average persistent current in a mesoscopic ring. *Phys. Rev. Lett.*, 66:84–87, Jan 1991.

[40] J. F. Weisz, R. Kishore, and F. V. Kusmartsev. Persistent current in isolated mesoscopic rings. *Phys. Rev. B*, 49:8126–8131, Mar 1994.

[41] Tapash Chakraborty and Pekka Pietiläinen. Persistent currents in a quantum ring: Effects of impurities and interactions. *Phys. Rev. B*, 52:1932–1935, Jul 1995.

[42] Yuriy V. Pershin and Carlo Piermarocchi. Persistent and radiation-induced currents in distorted quantum rings. *Phys. Rev. B*, 72:125348, Sep 2005.

[43] A. Bruno-Alfonso and A. Latgé. Quantum rings of arbitrary shape and non-uniform width in a threading magnetic field. *Phys. Rev. B*, 77:205303, May 2008.

[44] Hamutal Bary-Soroker, Ora Entin-Wohlman, and Yoseph Imry. Persistent currents of noninteracting electrons in one-, two-, and three-dimensional thin rings. *Phys. Rev. B*, 82:144202, Oct 2010.

[45] Lei Ying and Ying-Cheng Lai. Robustness of persistent currents in two-dimensional dirac systems with disorder. *Phys. Rev. B*, 96:165407, Oct 2017.

[46] H.-Y. Xu, L. Huang, Y.-C. Lai, and C. Grebogi. Superpersistent currents and whispering gallery modes in relativistic quantum chaotic systems. *Sci. Rep.*, 5:8963, 2015.

[47] Hongya Xu, Liang Huang, and Ying-Cheng Lai. A robust relativistic quantum two-level system with edge-dependent currents and spin polarization. *Europhys. Lett.*, 115:20005, 2016.

[48] A. H. C. Neto, F. Guinea, and N. M. R. Peres. Edge and surface states in the quantum Hall effect in graphene. *Phys. Rev. B*, 73:205408, May 2006.

[49] D. S. L. Abergel, Vadim M. Apalkov, and Tapash Chakraborty. Interplay between valley polarization and electron-electron interaction in a graphene ring. *Phys. Rev. B*, 78:193405, Nov 2008.

[50] C. W. J. Beenakker, A. R. Akhmerov, P. Recher, and J. Tworzydło. Correspondence between Andreev reflection and Klein tunneling in bipolar graphene. *Phys. Rev. B*, 77:075409, Feb 2008.

[51] M. Zarenia, J. M. Pereira, F. M. Peeters, and G. A. Farias. Electrostatically confined quantum rings in bilayer graphene. *Nano Lett.*, 9(12):4088–4092, 2009.

[52] R. Jackiw, A. I. Milstein, S.-Y. Pi, and I. S. Terekhov. Induced current and Aharonov-Bohm effect in graphene. *Phys. Rev. B*, 80:033413, Jul 2009.

[53] M. M. Ma, J. W. Ding, and N. Xu. Odd-even width effect on persistent current in zigzag hexagonal graphene rings. *Nanoscale*, 1:387–390, 2009.

[54] M. Zarenia, J. Milton Pereira, A. Chaves, F. M. Peeters, and G. A. Farias. Simplified model for the energy levels of quantum rings in single layer and bilayer graphene. *Phys. Rev. B*, 81:045431, Jan 2010.

[55] D. Soriano and J. Fernández-Rossier. Spontaneous persistent currents in a quantum spin Hall insulator. *Phys. Rev. B*, 82:161302, Oct 2010.

[56] Paolo Michetti and Patrik Recher. Bound states and persistent currents in topological insulator rings. *Phys. Rev. B*, 83:125420, Mar 2011.

[57] Bor-Luen Huang, Ming-Che Chang, and Chung-Yu Mou. Persistent currents in a graphene ring with armchair edges. *J. Phys.: Condens. Matter*, 24:245304, Jun 2012.

[58] D. Faria, A. Latgé, S. E. Ulloa, and N. Sandler. Currents and pseudomagnetic fields in strained graphene rings. *Phys. Rev. B*, 87:241403, Jun 2013.

[59] Doru Sticlet, Balázs Dóra, and Jérôme Cayssol. Persistent currents in Dirac fermion rings. *Phys. Rev. B*, 88:205401, Nov 2013.

[60] N. Bolívar, E. Medina, and B. Berche. Persistent charge and spin currents in the long-wavelength regime for graphene rings. *Phys. Rev. B*, 89:125413, Mar 2014.

[61] Chen-Di Han, Hong-Ya Xu, Liang Huang, and Ying-Cheng Lai. Manifestations of chaos in relativistic quantum systems-a study based on out-of-time-order correlator. *Physics Open*, 1:100001, 2019.

[62] AI Larkin and Yu N Ovchinnikov. Quasiclassical method in the theory of superconductivity. *Sov. Phys. JETP*, 28(6):1200–1205, 1969.

[63] CW von Keyserlingk, Tibor Rakovszky, Frank Pollmann, and SL Sondhi. Operator hydrodynamics, otocs, and entanglement growth in systems without conservation laws. *Phys. Rev. X*, 8(2):021013, 2018.

[64] Adam Nahum, Sagar Vijay, and Jeongwan Haah. Operator spreading in random unitary circuits. *Phys. Rev. X*, 8(2):021014, 2018.

[65] Pavan Hosur, Xiao-Liang Qi, Daniel A Roberts, and Beni Yoshida. Chaos in quantum channels. *J. High Energy Phys.*, 2016(2):4, 2016.

[66] Adam Nahum, Jonathan Ruhman, Sagar Vijay, and Jeongwan Haah. Quantum entanglement growth under random unitary dynamics. *Phys. Rev. X*, 7(3):031016, 2017.

[67] Jun Li, Ruihua Fan, Hengyan Wang, Bingtian Ye, Bei Zeng, Hui Zhai, Xinhua Peng, and Jiangfeng Du. Measuring out-of-time-order correlators on a nuclear magnetic resonance quantum simulator. *Phys. Rev. X*, 7(3):031011, 2017.

[68] Ivan Kukuljan, Sašo Grozdanov, and Tomaž Prosen. Weak quantum chaos. *Phys. Rev. B*, 96(6):060301, 2017.

[69] Cheng-Ju Lin and Olexei I Motrunich. Out-of-time-ordered correlators in a quantum ising chain. *Phys. Rev. B*, 97(14):144304, 2018.

[70] Kristan Jensen. Chaos in ads 2 holography. *Phys. Rev. Lett.*, 117(11):111601, 2016.

[71] Antonio M García-García, Bruno Loureiro, Aurelio Romero-Bermúdez, and Masaki Tezuka. Chaotic-integrable transition in the sachdev-ye-kitaev model. *Phys. Rev. Lett.*, 120(24):241603, 2018.

[72] Daniel A Roberts and Douglas Stanford. Diagnosing chaos using four-point functions in two-dimensional conformal field theory. *Phys. Rev. Lett.*, 115(13):131603, 2015.

[73] Juan Maldacena, Stephen H Shenker, and Douglas Stanford. A bound on chaos. *J. High Energy Phys.*, 2016(8):106, 2016.

[74] Koji Hashimoto, Keiju Murata, and Kentaroh Yoshida. Chaos in chiral condensates in gauge theories. *Phys. Rev. Lett.*, 117(23):231602, 2016.

[75] Koji Hashimoto, Keiju Murata, and Ryosuke Yoshii. Out-of-time-order correlators in quantum mechanics. *J. High Energy Phys.*, 2017(10):138, 2017.

[76] Balázs Dóra, Miklós Antal Werner, and Cătălin Paşcu Moca. Information scrambling at an impurity quantum critical point. *Phys. Rev. B*, 96(15):155116, 2017.

[77] Efim B. Rozenbaum, Sriram Ganeshan, and Victor Galitski. Lyapunov exponent and out-of-time-ordered correlator's growth rate in a chaotic system. *Phys. Rev. Lett.*, 118:086801, Feb 2017.

[78] Chen-Di Han, Cheng-Zhen Wang, Hong-Ya Xu, Danhong Huang, and Ying-Cheng Lai. Decay of semiclassical massless Dirac fermions from integrable and chaotic cavities. *Phys. Rev. B*, 98(10):104308, 2018.

[79] Tsuneya Ando, Takeshi Nakanishi, and Riichiro Saito. Berry's phase and absence of back scattering in carbon nanotubes. *J. Phys. Soc. Jpn.*, 67(8):2857–2862, 1998.

[80] DS Novikov. Elastic scattering theory and transport in graphene. *Phys. Rev. B*, 76(24):245435, 2007.

[81] Oskar Klein. Die reflexion von elektronen an einem potentialsprung nach der relativistischen dynamik von dirac. *Z. Phys.*, 53(3):157–165, 1929.

[82] UK Cambridge. *Relativistic Quantum Mechanics: with applications in condensed matter and atomic physics*. Cambridge University Press, 1998.

[83] M. I. Katsnelson, K. S. Novoselov, and A. K. Geim. Chiral tunnelling and the Klein paradox in graphene. *Nat. Phys.*, 2(9):620–625, Sep 2006.

[84] Chen-Di Han, Hong-Ya Xu, Danhong Huang, and Ying-Cheng Lai. Atomic collapse in pseudospin-1 systems. *Phys. Rev. B*, 99:245413, Jun 2019.

[85] Li-Li Ye, Chen-Di Han, Liang Huang, and Ying-Cheng Lai. Geometry-induced wave-function collapse. *Phys. Rev. A*, 106(2):022207, 2022.

[86] Chen-Di Han, Hong-Ya Xu, and Ying-Cheng Lai. Electrical confinement in a spectrum of two-dimensional dirac materials with classically integrable, mixed, and chaotic dynamics. *Phys. Rev. Research*, 2:013116, Feb 2020.

[87] Yue Zhao, Jonathan Wyrick, Fabian D Natterer, Joaquin F Rodriguez-Nieva, Cyprian Lewandowski, Kenji Watanabe, Takashi Taniguchi, Leonid S Levitov, Nikolai B Zhitenev, and Joseph A Stroscio. Creating and probing electron whispering-gallery modes in graphene. *Science*, 348(6235):672–675, 2015.

[88] Juwon Lee, Dillon Wong, Jairo Velasco Jr, Joaquin F Rodriguez-Nieva, Salman Kahn, Hsin-Zon Tsai, Takashi Taniguchi, Kenji Watanabe, Alex Zettl, Feng Wang, et al. Imaging electrostatically confined Dirac fermions in graphene quantum dots. *Nat. Phys.*, 12(11):1032, 2016.

[89] Christopher Gutierrez, Lola Brown, Cheol-Joo Kim, Jiwoong Park, and Abhay N. Pasupathy. Klein tunnelling and electron trapping in nanometre-scale graphene quantum dots. *Nat. Phys.*, 12:1069–1075, 2016.

[90] Fereshte Ghahari, Daniel Walkup, Christopher Gutiérrez, Joaquin F Rodriguez-Nieva, Yue Zhao, Jonathan Wyrick, Fabian D Natterer, William G Cullen, Kenji Watanabe, Takashi Taniguchi, et al. An on/off Berry phase switch in circular graphene resonators. *Science*, 356(6340):845–849, 2017.

[91] Jairo Velasco Jr, Long Ju, Dillon Wong, Salman Kahn, Juwon Lee, Hsin-Zon Tsai, Chad Germany, Sebastian Wickenburg, Jiong Lu, Takashi Taniguchi, et al. Nanoscale control of rewriteable doping patterns in pristine graphene/boron nitride heterostructures. *Nano Lett.*, 16(3):1620–1625, 2016.

[92] Andrei V. Shytov, Mark S. Rudner, and Leonid S. Levitov. Klein backscattering and Fabry-Pérot interference in graphene heterojunctions. *Phys. Rev. Lett.*, 101:156804, Oct 2008.

[93] N. Stander, B. Huard, and D. Goldhaber-Gordon. Evidence for Klein tunneling in graphene $p-n$ junctions. *Phys. Rev. Lett.*, 102:026807, Jan 2009.

[94] Andrea F. Young and Philip Kim. Quantum interference and Klein tunnelling in graphene heterojunctions. *Nat. Phys.*, 5:222–226, 2009.

[95] Rui Yang, Liang Huang, Ying-Cheng Lai, and Celso Grebogi. Abnormal electron paths induced by Klein tunneling in graphene quantum point contacts. *Phys. Rev. B*, 84:035426, Jul 2011.

[96] Hong-Ya Xu and Ying-Cheng Lai. Revival resonant scattering, perfect caustics, and isotropic transport of pseudospin-1 particles. *Phys. Rev. B*, 94:165405, Oct 2016.

[97] Pierre E Allain and Jean-Noel Fuchs. Klein tunneling in graphene: optics with massless electrons. *Eur. Phys. J. B*, 83(3):301, 2011.

[98] Cheng-Zhen Wang, Chen-Di Han, Hong-Ya Xu, and Ying-Cheng Lai. Chaos-based Berry phase detector. *Phys. Rev. B*, 99(14):144302, 2019.

[99] József Cserti, András Pályi, and Csaba Péterfalvi. Caustics due to a negative refractive index in circular graphene p-n junctions. *Phys. Rev. Lett.*, 99(24):246801, 2007.

[100] Vadim V Cheianov, Vladimir Fal'ko, and BL Altshuler. The focusing of electron flow and a Veselago lens in graphene pn junctions. *Science*, 315(5816):1252–1255, 2007.

[101] Xu Du, Ivan Skachko, Anthony Barker, and Eva Y Andrei. Approaching ballistic transport in suspended graphene. *Nat. Nanotechnol.*, 3(8):491, 2008.

[102] CWJ Beenakker, RA Sepkhanov, AR Akhmerov, and J Tworzydło. Quantum Goos-Hänchen effect in graphene. *Phys. Rev. Lett.*, 102(14):146804, 2009.

[103] Ali G Moghaddam and Malek Zareyan. Graphene-based electronic spin lenses. *Phys. Rev. Lett.*, 105(14):146803, 2010.

[104] Nan Gu, Mark Rudner, and Leonid Levitov. Chirality-assisted electronic cloaking of confined states in bilayer graphene. *Phys. Rev. Lett.*, 107(15):156603, 2011.

[105] JR Williams, Tony Low, MS Lundstrom, and CM Marcus. Gate-controlled guiding of electrons in graphene. *Nat. Nanotechnol.*, 6(4):222, 2011.

[106] Peter Rickhaus, Romain Maurand, Ming-Hao Liu, Markus Weiss, Klaus Richter, and Christian Schönenberger. Ballistic interferences in suspended graphene. *Nat. Commun.*, 4:2342, 2013.

[107] Bolin Liao, Mona Zebarjadi, Keivan Esfarjani, and Gang Chen. Isotropic and energy-selective electron cloaks on graphene. *Phys. Rev. B*, 88(15):155432, 2013.

[108] RL Heinisch, FX Bronold, and H Fehske. Mie scattering analog in graphene: Lensing, particle confinement, and depletion of Klein tunneling. *Phys. Rev. B*, 87(15):155409, 2013.

[109] Mahmoud M Asmar and Sergio E Ulloa. Rashba spin-orbit interaction and birefringent electron optics in graphene. *Phys. Rev. B*, 87(7):075420, 2013.

[110] Jhih-Sheng Wu and Michael M. Fogler. Scattering of two-dimensional massless Dirac electrons by a circular potential barrier. *Phys. Rev. B*, 90:235402, Dec 2014.

[111] Peter Rickhaus, Ming-Hao Liu, Péter Makk, Romain Maurand, Samuel Hess, Simon Zihlmann, Markus Weiss, Klaus Richter, and Christian Schönenberger. Guiding of electrons in a few-mode ballistic graphene channel. *Nano Lett.*, 15(9):5819–5825, 2015.

[112] Gil-Ho Lee, Geon-Hyoung Park, and Hu-Jong Lee. Observation of negative refraction of Dirac fermions in graphene. *Nat. Phys.*, 11(11):925, 2015.

[113] Peter Rickhaus, Péter Makk, Ming-Hao Liu, Klaus Richter, and Christian Schönenberger. Gate tuneable beamsplitter in ballistic graphene. *Appl. Phys. Lett.*, 107(25):251901, 2015.

[114] Jamie D Walls and Daniel Hadad. The Talbot effect for two-dimensional massless Dirac fermions. *Sci. Rep.*, 6:26698, 2016.

[115] José M Caridad, Stephen Connaughton, Christian Ott, Heiko B Weber, and Vojislav Krstić. An electrical analogy to Mie scattering. *Nat. Commun.*, 7:12894, 2016.

[116] Shaowen Chen, Zheng Han, Mirza M Elahi, KM Masum Habib, Lei Wang, Bo Wen, Yuanda Gao, Takashi Taniguchi, Kenji Watanabe, James Hone, et al. Electron optics with pn junctions in ballistic graphene. *Science*, 353(6307):1522–1525, 2016.

[117] Mikkel Settnes, Stephen R Power, Mads Brandbyge, and Antti-Pekka Jauho. Graphene nanobubbles as valley filters and beam splitters. *Phys. Rev. Lett.*, 117(27):276801, 2016.

[118] Ming-Hao Liu, Cosimo Gorini, and Klaus Richter. Creating and steering highly directional electron beams in graphene. *Phys. Rev. Lett.*, 118:066801, Feb 2017.

[119] Arthur W Barnard, Alex Hughes, Aaron L Sharpe, Kenji Watanabe, Takashi Taniguchi, and David Goldhaber-Gordon. Absorptive pinhole collimators for ballistic Dirac fermions in graphene. *Nat. Commun.*, 8:15418, 2017.

[120] Yuhang Jiang, Jinhai Mao, Dean Moldovan, Massoud Ramezani Masir, Guohong Li, Kenji Watanabe, Takashi Taniguchi, Francois M Peeters, and Eva Y Andrei. Tuning a circular p-n junction in graphene from quantum confinement to optical guiding. *Nat. Nanotechnol.*, 12(11):1045, 2017.

[121] Shu-Hui Zhang, Jia-Ji Zhu, Wen Yang, and Kai Chang. Focusing RKKY interaction by graphene p-n junction. *2D Mater.*, 4(3):035005, 2017.

[122] Peter Bøggild, José M Caridad, Christoph Stampfer, Gaetano Calogero, Nick Rübner Papior, and Mads Brandbyge. A two-dimensional Dirac fermion microscope. *Nat. Commun.*, 8:15783, 2017.

[123] Hong-Ya Xu, Guang-Lei Wang, Liang Huang, and Ying-Cheng Lai. Chaos in Dirac electron optics: Emergence of a relativistic quantum chimera. *Phys. Rev. Lett.*, 120(12):124101, 2018.

[124] Hong-Ya Xu, Liang Huang, Danhong Huang, and Ying-Cheng Lai. Geometric valley Hall effect and valley filtering through a singular Berry flux. *Phys. Rev. B*, 96(4):045412, 2017.

[125] Chen-Di Han, Hong-Ya Xu, and Ying-Cheng Lai. Pseudospin modulation in coupled graphene systems. *Phys. Rev. Res.*, 2(3):033406, 2020.

[126] Shanhui Fan. Sharp asymmetric line shapes in side-coupled waveguide-cavity systems. *Appl. Phys. Lett.*, 80(6):908–910, 2002.

[127] Andrey E. Miroshnichenko, Sergej Flach, and Yuri S. Kivshar. Fano resonances in nanoscale structures. *Rev. Mod. Phys.*, 82:2257–2298, Aug 2010.

[128] Alessandro Chiasera, Yannick Dumeige, Patrice Feron, Maurizio Ferrari, Yoann Jestin, Gualtiero Nunzi Conti, Stefano Pelli, Silvia Soria, and Giancarlo C Righini. Spherical whispering-gallery-mode microresonators. *Laser Photon. Rev*, 4(3):457–482, 2010.

[129] Hui Cao and Jan Wiersig. Dielectric microcavities: Model systems for wave chaos and non-Hermitian physics. *Rev. Mod. Phys.*, 87:61–111, Jan 2015.

[130] Fan-Ming Zhang, Ying He, and Xi Chen. Guided modes in graphene waveguides. *Appl. Phys. Lett.*, 94(21):212105, 2009.

[131] Richard R Hartmann, NJ Robinson, and ME Portnoi. Smooth electron waveguides in graphene. *Phys. Rev. B*, 81(24):245431, 2010.

[132] Nojoon Myoung, G Ihm, and SJ Lee. Magnetically induced waveguide in graphene. *Phys. Rev. B*, 83(11):113407, 2011.

[133] Richard R Hartmann and ME Portnoi. Quasi-exact solution to the dirac equation for the hyperbolic-secant potential. *Phys. Rev. A*, 89(1):012101, 2014.

[134] CA Downing, AR Pearce, RJ Churchill, and ME Portnoi. Optimal traps in graphene. *Phys. Rev. B*, 92(16):165401, 2015.

[135] Hongya Xu, Liang Huang, Ying-Cheng Lai, and Celso Grebogi. Chiral scars in chaotic dirac fermion systems. *Phys. Rev. Lett.*, 110(6):064102, 2013.

[136] Zhong-Qiu Fu, Yue-Ting Pan, Jiao-Jiao Zhou, Dong-Lin Ma, Yu Zhang, Jia-Bin Qiao, Haiwen Liu, Hua Jiang, and Lin He. Relativistic artificial molecules realized by two coupled graphene quantum dots. *arXiv preprint arXiv:1908.06580*, 2019.

[137] Mahdiyeh Sadrara and MirFaez Miri. Dirac electron scattering from a cluster of electrostatically defined quantum dots in graphene. *Phys. Rev. B*, 99(15):155432, 2019.

[138] Chen-Di Han and Ying-Cheng Lai. Optical response of two-dimensional dirac materials with a flat band. *Phys. Rev. B*, 105(15):155405, 2022.

[139] D. Bercioux, D. F. Urban, H. Grabert, and W. Häusler. Massless dirac-weyl fermions in a $\sqcup_3$ optical lattice. *Phys. Rev. A*, 80(6):063603, 2009.

[140] A Raoux, M Morigi, J-N Fuchs, F Piéchon, and G Montambaux. From dia- to paramagnetic orbital susceptibility of massless fermions. *Phys. Rev. Lett.*, 112(2):026402, 2014.

[141] E Illes, J P Carbotte, and E J Nicol. Hall quantization and optical conductivity evolution with variable Berry phase in the α-T_3 model. *Phys. Rev. B*, 92(24):245410, 2015.

[142] Máté Vigh, László Oroszlány, Szabolcs Vajna, Pablo San-Jose, Gyula Dávid, József Cserti, and Balázs Dóra. Diverging dc conductivity due to a flat band in a disordered system of pseudospin-1 dirac-weyl fermions. *Phys. Rev. B*, 88(16):161413, 2013.

[143] Thibaud Louvet, Pierre Delplace, Andrei A Fedorenko, and David Carpentier. On the origin of minimal conductivity at a band crossing. *Phys. Rev. B*, 92(15):155116, 2015.

[144] CJ Tabert, JP Carbotte, and EJ Nicol. Optical and transport properties in three-dimensional Dirac and Weyl semimetals. *Phys. Rev. B*, 93(8):085426, 2016.

[145] Áron Dániel Kovács, Gyula Dávid, Balázs Dóra, and József Cserti. Frequency-dependent magneto-optical conductivity in the generalized α-T_3 model. *Phys. Rev. B*, 95(3):035414, 2017.

[146] Yan-Ru Chen, Yong Xu, Jun Wang, Jun-Feng Liu, and Zhongshui Ma. Enhanced magneto-optical response due to the flat band in nanoribbons made from the α-$\mathcal{T}_3$ lattice. *Phys. Rev. B*, 99(4):045420, 2019.

[147] Chen-Di Han, Bryan Glaz, Mulugeta Haile, and Ying-Cheng Lai. Adaptable hamiltonian neural networks. *Phys. Rev. Res.*, 3(2):023156, 2021.

[148] Razieh Falahian, Maryam Mehdizadeh Dastjerdi, Malihe Molaie, Sajad Jafari, and Shahriar Gharibzadeh. Artificial neural network-based modeling of brain response to flicker light. *Nonlinear Dyn.*, 81(4):1951–1967, 2015.

[149] Rok Cestnik and Markus Abel. Inferring the dynamics of oscillatory systems using recurrent neural networks. *Chaos*, 29(6):063128, 2019.

[150] Jason Z Kim, Zhixin Lu, Erfan Nozari, George J Pappas, and Danielle S Bassett. Teaching recurrent neural networks to modify chaotic memories by example. *arXiv preprint arXiv:2005.01186*, 2020.

[151] Christian Klos, Yaroslav Felipe Kalle Kossio, Sven Goedeke, Aditya Gilra, and Raoul-Martin Memmesheimer. Dynamical learning of dynamics. *Phys. Rev. Lett.*, 125(8):088103, 2020.

[152] Ling-Wei Kong, Hua-Wei Fan, Celso Grebogi, and Ying-Cheng Lai. Machine learning prediction of critical transition and system collapse. *Phys. Rev. Research*, 3:013090, Jan 2021.

[153] Sam Greydanus, Misko Dzamba, and Jason Yosinski. Hamiltonian neural networks. *arXiv:1906.01563*, 2019.

[154] Peter Toth, Danilo Jimenez Rezende, Andrew Jaegle, Sébastien Racanière, Aleksandar Botev, and Irina Higgins. Hamiltonian generative networks. *arXiv:1909.13789*, 2019.

[155] Tom Bertalan, Felix Dietrich, Igor Mezić, and Ioannis G Kevrekidis. On learning Hamiltonian systems from data. *Chaos*, 29(12):121107, 2019.

[156] Anshul Choudhary, John F. Lindner, Elliott G. Holliday, Scott T. Miller, Sudeshna Sinha, and William L. Ditto. Physics-enhanced neural networks learn order and chaos. *Phys. Rev. E*, 101:062207, 2020.

[157] Ayush Garg and Sammed Shantinath Kagi. Hamiltonian neural networks, 2019.

[158] Giuseppe Carleo, Ignacio Cirac, Kyle Cranmer, Laurent Daudet, Maria Schuld, Naftali Tishby, Leslie Vogt-Maranto, and Lenka Zdeborová. Machine learning and the physical sciences. *Rev. Mod. Phys.*, 91(4):045002, 2019.

[159] Chen-Di Han, Bryan Glaz, Mulugeta Haile, and Ying-Cheng Lai. Tomography of time-dependent quantum hamiltonians with machine learning. *Phys. Rev. A*, 104(6):062404, 2021.

[160] Philippe De Wilde. Class of Hamiltonian neural networks. *Phys. Rev. E*, 47(2):1392, 1993.

[161] Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *International Conference on Machine Learning*, pages 1126–1135. PMLR, 2017.

[162] Brenden M Lake, Ruslan Salakhutdinov, and Joshua B Tenenbaum. Human-level concept learning through probabilistic program induction. *Science*, 350(6266):1332–1338, 2015.

[163] K S Novoselov, A Mishchenko, A Carvalho, and A H Castro Neto. 2D materials and van der Waals heterostructures. *Science*, 353(6298):aac9439, 2016.

[164] E Illes and EJ Nicol. Magnetic properties of the α- t 3 model: Magneto-optical conductivity and the hofstadter butterfly. *Phys. Rev. B*, 94(12):125435, 2016.

[165] Tutul Biswas and Tarun Kanti Ghosh. Magnetotransport properties of the α-T_3 model. *J. Phys.: Condens. Matter*, 28(49):495302, 2016.

[166] E Illes and E J Nicol. Klein tunneling in the α-T_3 model. *Phys. Rev. B*, 95(23):235432, 2017.

[167] Yong Xu and L-M Duan. Unconventional quantum Hall effects in two-dimensional massive spin-1 fermion systems. *Phys. Rev. B*, 96(15):155301, 2017.

[168] Bashab Dey and Tarun Kanti Ghosh. Photoinduced valley and electron-hole symmetry breaking in α-T_3 lattice: The role of a variable Berry phase. *Phys. Rev. B*, 98(7):075422, 2018.

[169] E V Gorbar, V P Gusynin, and D O Oriekhov. Electron states for gapped pseudospin-1 fermions in the field of a charged impurity. *Phys. Rev. B*, 99(15):155124, 2019.

[170] Matteo Rizzi, Vittorio Cataudella, and Rosario Fazio. Phase diagram of the Bose-Hubbard model with T_3 symmetry. *Phys. Rev. B*, 73(14):144511, 2006.

[171] Daniel Leykam and Sergej Flach. Perspective: Photonic flatbands. *APL Photon.*, 3(7):070901, 2018.

[172] Sebabrata Mukherjee, Marco Di Liberto, Patrik Öhberg, Robert R Thomson, and Nathan Goldman. Experimental observation of aharonov-bohm cages in photonic lattices. *Phys. Rev. Lett.*, 121(7):075502, 2018.

[173] A. Fang, Z. Q. Zhang, Steven G. Louie, and C. T. Chan. Klein tunneling and supercollimation of pseudospin-1 electromagnetic waves. *Phys. Rev. B*, 93:035422, Jan 2016.

[174] Fa Wang and Ying Ran. Nearly flat band with Chern number $c = 2$ on the dice lattice. *Phys. Rev. B*, 84:241103, Dec 2011.

[175] Judit Romhányi, Karlo Penc, and R. Ganesh. Hall effect of triplons in a dimerized quantum magnet. *Nat. Commun.*, 6:6805, 2015.

[176] Gianluca Giovannetti, Massimo Capone, Jeroen van den Brink, and Carmine Ortix. Kekulé textures, pseudospin-one Dirac cones, and quadratic band crossings in a graphene-hexagonal indium chalcogenide bilayer. *Phys. Rev. B*, 91:121417, Mar 2015.

[177] Kerry J Vahala. Optical microcavities. *Nature*, 424(6950):839, 2003.

[178] Wei Han, Roland K Kawakami, Martin Gmitra, and Jaroslav Fabian. Graphene spintronics. *Nat. Nanotechnol.*, 9(10):794, 2014.

[179] A Rycerz, J Tworzydło, and CWJ Beenakker. Valley filter and valley valve in graphene. *Nat. Phys.*, 3(3):172, 2007.

[180] A. Mekis, J. U. Nöckel, G. Chen, A. D. Stone, and R. K. Chang. Ray chaos and Q spoiling in lasing droplets. *Phys. Rev. Lett.*, 75:2682–2685, Oct 1995.

[181] Jens U. Nöckel and A. Douglas Stone. Ray and wave chaos in asymmetric resonant optical cavities. *Nature*, 385(6611):45–47, Jan 1997.

[182] Hongya Xu and Ying-Cheng Lai. Superscattering of a pseudospin-1 wave in a photonic lattice. *Phys. Rev. A*, 95(1):012119, 2017.

[183] Eugene P Wigner. Lower limit for the energy derivative of the scattering phase shift. *Phys. Rev.*, 98(1):145, 1955.

[184] Felix T Smith. Lifetime matrix in collision theory. *Phys. Rev.*, 118(1):349, 1960.

[185] Stefan Rotter and Sylvain Gigan. Light fields in complex media: Mesoscopic scattering meets wave control. *Rev. Mod. Phys.*, 89(1):015005, 2017.

[186] DM Basko. Resonant low-energy electron scattering on short-range impurities in graphene. *Phys. Rev. B*, 78(11):115432, 2008.

[187] M Titov, PM Ostrovsky, IV Gornyi, A Schuessler, and AD Mirlin. Charge transport in graphene with resonant scatterers. *Phys. Rev. Lett.*, 104(7):076802, 2010.

[188] Martina Hentschel and Klaus Richter. Quantum chaos in optical systems: The annular billiard. *Phys. Rev. E*, 66(5):056207, 2002.

[189] Arnd Bäcker, Roland Ketzmerick, Steffen Löck, Jan Wiersig, and Martina Hentschel. Quality factors and dynamical tunneling in annular microcavities. *Phys. Rev. A*, 79(6):063804, 2009.

[190] Alexey Yamilov and Hui Cao. Density of resonant states and a manifestation of photonic band structure in small clusters of spherical particles. *Phys. Rev. B*, 68(8):085111, 2003.

[191] Matthieu Davy, Zhou Shi, Jing Wang, Xiaojun Cheng, and Azriel Z Genack. Transmission eigenchannels and the densities of states of random media. *Phys. Rev. Lett.*, 114(3):033901, 2015.

[192] R. Blümel and U. Smilansky. Classical irregular scattering and its quantum-mechanical implications. *Phys. Rev. Lett.*, 60(6):477–480, 1988.

[193] R. Blümel and U. Smilansky. A simple model for chaotic scattering: II. Quantum mechanical theory. *Physica D*, 36(1-2):111–136, 1989.

[194] Y.-C. Lai, R. Blümel, E. Ott, and C. Grebogi. Quantum manifestations of chaotic scattering. *Phys. Rev. Lett.*, 68(24):3491–3494, 1992.

[195] R. Ketzmerick. Fractal conductance fluctuations in generic chaotic cavities. *Phys. Rev. B*, 54(15):10841–10844, 1996.

[196] A. P. S. de Moura, Y.-C. Lai., R. Akis, J. P. Bird, and D. K. Ferry. Tunneling and nonhyperbolicity in quantum dots. *Phys. Rev. Lett.*, 88(23):236804, 2002.

[197] Jens H Bardarson, M Titov, and PW Brouwer. Electrostatic confinement of electrons in an integrable graphene quantum dot. *Phys. Rev. Lett.*, 102(22):226803, 2009.

[198] Martin Schneider and Piet W Brouwer. Resonant scattering in graphene with a gate-defined chaotic quantum dot. *Phys. Rev. B*, 84(11):115440, 2011.

[199] Julia Heinl, Martin Schneider, and Piet W Brouwer. Interplay of Aharonov-Bohm and Berry phases in gate-defined graphene quantum dots. *Phys. Rev. B*, 87(24):245426, 2013.

[200] Martin Schneider and Piet W Brouwer. Density of states as a probe of electrostatic confinement in graphene. *Phys. Rev. B*, 89(20):205437, 2014.

[201] Hong-Ya Xu and Ying-Cheng Lai. Pseudospin-1 wave scattering that defies chaos Q-spoiling and Klein tunneling. *Phys. Rev. B*, 99:235403, Jun 2019.

[202] Y. Leviatan and A. Boag. Analysis of electromagnetic scattering from dielectric cylinders using a multifilament current model. *IEEE Trans. Anten. Propa.*, 35(10):1119–1127, 1987.

[203] Matthias G. Imhof. Multiple multipole expansions for elastic scattering. *J. Acous. Soc. Am.*, 100(5):2969–2979, 1996.

[204] D. I. Kaklamani and H. T. Anastassiu. Aspects of the method of auxiliary sources (MAS) in computational electromagnetics. *IEEE Anten. Propag. Maga.*, 44(3):48–64, 2002.

[205] Esteban Moreno, Daniel Erni, Christian Hafner, and Rüdiger Vahldieck. Multiple multipole method with automatic multipole setting applied to the simulation of surface plasmons in metallic nanostructures. *J. Opt. Soc. Am. A*, 19(1):101–111, Jan 2002.

[206] Gérard Tayeb and Stefan Enoch. Combined fictitious-sources–scattering-matrix method. *J. Opt. Soc. Am. A*, 21(8):1417–1423, Aug 2004.

[207] L. Huang, Y.-C. Lai, D. K. Ferry, S. M. Goodnick, and R. Akis. Relativistic quantum scars. *Phys. Rev. Lett.*, 103(5):054101, 2009.

[208] Cheng-Zhen Wang, Liang Huang, and Kai Chang. Scars in Dirac fermion systems: The influence of an Aharonov-Bohm flux. *New J. Phys.*, 19(1):013018, 2017.

[209] Min-Yue Song, Zi-Yuan Li, Hong-Ya Xu, Liang Huang, and Ying-Cheng Lai. Quantization of massive dirac billiards and unification of nonrelativistic and relativistic chiral quantum scars. *Phys. Rev. Research*, 1:033008, Oct 2019.

[210] Jan Wiersig and Martina Hentschel. Combining directional light output and ultralow loss in deformed microdisks. *Phys. Rev. Lett.*, 100(3):033901, 2008.

[211] Susumu Shinohara, Takahisa Harayama, Takehiro Fukushima, Martina Hentschel, Satoshi Sunada, and Evgenii E Narimanov. Chaos-assisted emission from asymmetric resonant cavity microlasers. *Phys. Rev. A*, 83(5):053837, 2011.

[212] Kai Chang and Wen-Kai Lou. Helical quantum states in HgTe quantum dots with inverted band structures. *Phys. Rev. Lett.*, 106(20):206802, 2011.

[213] Athmane Tadjine and Christophe Delerue. Topological protection of electronic states against disorder probed by their magnetic moment. *Phys. Rev. B*, 95(23):235426, 2017.

[214] Bartłomiej Szafran, Alina Mreńca-Kolasińska, Bartłomiej Rzeszotarski, and Dariusz Żebrowski. Electrical control of a confined electron spin in a silicene quantum dot. *Phys. Rev. B*, 97(16):165303, 2018.

[215] Frank W Olver, Daniel W Lozier, Ronald F Boisvert, and Charles W Clark. *NIST Handbook of Mathematical Functions*. Cambridge university press, New York, NY, 2010.

[216] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[217] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT, Cambridge, MA, 2016.

[218] Herbert Jaeger. The "echo state" approach to analysing and training recurrent neural networks-with an erratum note. *German National Research Center for Information Technology GMD Technical Report*, 148(34):13, 2001.

[219] W. Mass, T. Nachtschlaeger, and H. Markram. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neur. Comp.*, 14:2531–2560, 2002.

[220] H. Jaeger and H. Haas. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication. *Science*, 304:78–80, 2004.

[221] G. Manjunath and H. Jaeger. Echo state property linked to an input: Exploring a fundamental characteristic of recurrent neural networks. *Neur. Comp.*, 25:671–696, 2013.

[222] Nicholas D. Haynes, Miguel C. Soriano, David P. Rosin, Ingo Fischer, and Daniel J. Gauthier. Reservoir computing with a single time-delay autonomous Boolean node. *Phys. Rev. E*, 91:020801, Feb 2015.

[223] Laurent Larger, Antonio Baylón-Fuentes, Romain Martinenghi, Vladimir S. Udaltsov, Yanne K. Chembo, and Maxime Jacquot. High-speed photonic reservoir computing using a time-delay-based architecture: Million words per second classification. *Phys. Rev. X*, 7:011015, Feb 2017.

[224] Jaideep Pathak, Zhixin Lu, Brian Hunt, Michelle Girvan, and Edward Ott. Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data. *Chaos*, 27:121102, 2017.

[225] Zhixin Lu, Jaideep Pathak, Brian Hunt, Michelle Girvan, Roger Brockett, and Edward Ott. Reservoir observers: Model-free inference of unmeasured variables in chaotic systems. *Chaos*, 27:041102, 2017.

[226] Thomas Duriez, Steven L Brunton, and Bernd R Noack. *Machine Learning Control - Taming Nonlinear Dynamics and Turbulence*. Springer, 2017.

[227] Zhixin Lu, Brian R Hunt, and Edward Ott. Attractor reconstruction by machine learning. *Chaos*, 28(6):061104, 2018.

[228] Jaideep Pathak, Alexander Wilner, Rebeckah Fussell, Sarthak Chandra, Brian Hunt, Michelle Girvan, Zhixin Lu, and Edward Ott. Hybrid forecasting of chaotic processes: Using machine learning in conjunction with a knowledge-based model. *Chaos*, 28:041101, 2018.

[229] Jaideep Pathak, Brian Hunt, Michelle Girvan, Zhixin Lu, and Edward Ott. Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach. *Phys. Rev. Lett.*, 120:024102, Jan 2018.

[230] T. L. Carroll. Using reservoir computers to distinguish chaotic signals. *Phys. Rev. E*, 98:052209, Nov 2018.

[231] Kengo Nakai and Yoshitaka Saiki. Machine-learning inference of fluid variables from data using reservoir computing. *Phys. Rev. E*, 98:023111, Aug 2018.

[232] Zimmermann S. Roland and Ulrich Parlitz. Observing spatio-temporal dynamics of excitable media using reservoir computing. *Chaos*, 28:043118, 2018.

[233] Tongfeng Weng, Huijie Yang, Changgui Gu, Jie Zhang, and Michael Small. Synchronization of chaotic systems and their machine-learning models. *Phys. Rev. E*, 99:042203, Apr 2019.

[234] Aaron Griffith, Andrew Pomerance, and Daniel J. Gauthier. Forecasting chaotic systems with very low connectivity reservoir computers. *Chaos*, 29:123108, 2019.

[235] Junjie Jiang and Ying-Cheng Lai. Model-free prediction of spatiotemporal dynamical systems with recurrent neural networks: Role of network spectral radius. *Phys. Rev. Research*, 1:033056, Oct 2019.

[236] Pantelis R Vlachas, Jaideep Pathak, Brian R Hunt, Themistoklis P Sapsis, Michelle Girvan, Edward Ott, and Petros Koumoutsakos. Forecasting of spatiotemporal chaotic dynamics with recurrent neural networks: A comparative study of reservoir computing and backpropagation algorithms. *arXiv preprint arXiv:1910.05266*, 2019.

[237] Huawei Fan, Junjie Jiang, Chun Zhang, Xingang Wang, and Ying-Cheng Lai. Long-term prediction of chaotic systems with machine learning. *Phys. Rev. Research*, 2:012080, Mar 2020.

[238] Chun Zhang, Junjie Jiang, Si-Xian Qu, and Ying-Cheng Lai. Predicting phase and sensing phase coherence in chaotic systems with machine learning. *Chaos*, 30:083114, 2020.

[239] Zhengdao Chen, Jianyu Zhang, Martin Arjovsky, and Léon Bottou. Symplectic recurrent neural networks. *arXiv:1909.13334*, 2019.

[240] Miles Cranmer, Sam Greydanus, Stephan Hoyer, Peter Battaglia, David Spergel, and Shirley Ho. Lagrangian neural networks. *arXiv preprint arXiv:2003.04630*, 2020.

[241] Alvaro Sanchez-Gonzalez, Victor Bapst, Kyle Cranmer, and Peter Battaglia. Hamiltonian graph networks with ode integrators. *arXiv preprint arXiv:1909.12790*, 2019.

[242] Zachary Dulberg and Jonathan Cohen. Learning canonical transformations. *arXiv preprint arXiv:2011.08822*, 2020.

[243] Anshul Choudhary, John F Lindner, Elliott G Holliday, Scott T Miller, Sudeshna Sinha, and William L Ditto. Forecasting Hamiltonian dynamics without canonical coordinates. *arXiv preprint arXiv:2010.15201*, 2020.

[244] Michael Lutter, Christian Ritter, and Jan Peters. Deep Lagrangian networks: Using physics as model prior for deep learning. *arXiv preprint arXiv:1907.04490*, 2019.

[245] Ioannis Havoutis and Subramanian Ramamoorthy. Geodesic trajectory generation on learnt skill manifolds. In *2010 IEEE International Conference on Robotics and Automation*, pages 2946–2952. IEEE, 2010.

[246] Dan Guest, Kyle Cranmer, and Daniel Whiteson. Deep learning and its application to LHC physics. *Annu. Rev. Nucl. Part. Sci.*, 68:161–181, 2018.

[247] Alexander Radovic, Mike Williams, David Rousseau, Michael Kagan, Daniele Bonacorsi, Alexander Himmel, Adam Aurisano, Kazuhiro Terao, and Taritree Wongjirad. Machine learning at the energy and intensity frontiers of particle physics. *Nature*, 560(7716):41–48, 2018.

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

[249] John Peurifoy, Yichen Shen, Li Jing, Yi Yang, Fidel Cano-Renteria, Brendan G DeLacy, John D Joannopoulos, Max Tegmark, and Marin Soljačić. Nanophotonic particle simulation and inverse design using artificial neural networks. *Sci. Adv.*, 4(6):eaar4206, 2018.

[250] Vedran Dunjko and Hans J Briegel. Machine learning & artificial intelligence in the quantum domain: a review of recent progress. *Rep. Prog. Phys.*, 81(7):074001, 2018.

[251] Raban Iten, Tony Metger, Henrik Wilming, Lídia Del Rio, and Renato Renner. Discovering physical concepts with neural networks. *Phys. Rev. Lett.*, 124(1):010508, 2020.

[252] Skander Karkar, Ibrahhim Ayed, Emmanuel de Bézenac, and Patrick Gallinari. A principle of least action for the training of neural networks. *arXiv preprint arXiv:2009.08372*, 2020.

[253] Shuqi Yang, Xingzhe He, and Bo Zhu. Learning physical constraints with neural projections. *Adv. Neu. Info. Process. Sys.*, 33, 2020.

[254] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. Tensorflow: A system for large-scale machine learning. In *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, pages 265–283, 2016.

[255] François Chollet et al. Keras, 2015.

[256] Michel Hénon and Carl Heiles. The applicability of the third integral of motion: some numerical experiments. *Astron. J.*, 69:73, 1964.

[257] Boyd A Waite and William H Miller. Mode specificity in unimolecular reaction dynamics: The Henon–Heiles potential energy surface. *J. Chem. Phys.*, 74(7):3910–3915, 1981.

[258] MD Feit and JA Fleck Jr. Wave packet dynamics and chaos in the Hénon–Heiles system. *J. Chem. Phys.*, 80(6):2578–2584, 1984.

[259] Oriol Vendrell and Hans-Dieter Meyer. Multilayer multiconfiguration time-dependent Hartree method: Implementation and applications to a Henon–Heiles Hamiltonian and to pyrazine. *J. Chem. Phys.*, 134(4):044135, 2011.

[260] Alessandro PS de Moura and Patricio S Letelier. Fractal basins in Hénon–Heiles and other polynomial potentials. *Phys. Lett. A*, 256(5-6):362–368, 1999.

[261] Jesús M Seoane, Jacobo Aguirre, Miguel AF Sanjuán, and Ying-Cheng Lai. Basin topology in dissipative chaotic scattering. *Chaos*, 16(2):023101, 2006.

[262] Jesús M Seoane, Miguel AF Sanjuán, and Ying-Cheng Lai. Fractal dimension in dissipative chaotic scattering. *Phys. Rev. E*, 76(1):016208, 2007.

[263] L. D. Landau and E. M. Lifshitz. *Quantum Mechanics: Non-Relativistic Theory.* Pergamon, New York, 1981.

[264] Ch Skokos. Alignment indices: a new, simple method for determining the ordered or chaotic nature of orbits. *J. Phys. A Math. Gen.*, 34(47):10029, 2001.

[265] Euaggelos E Zotos. Classifying orbits in the classical Hénon–Heiles Hamiltonian system. *Nonlinear Dyn.*, 79(3):1665–1677, 2015.

[266] Philip M Morse. Diatomic molecules according to the wave mechanics II. Vibrational levels. *Phys. Rev.*, 34(1):57, 1929.

[267] Ying-Cheng Lai. Abrupt bifurcation to chaotic scattering with discontinuous change in fractal dimension. *Phys. Rev. E*, 60:R6283–R6286, Dec 1999.

[268] Ying-Cheng Lai, Alessandro PS De Moura, and Celso Grebogi. Topology of high-dimensional chaotic scattering. *Phys. Rev. E*, 62(5):6421, 2000.

[269] Enrico Fermi, P Pasta, Stanislaw Ulam, and M Tsingou. Studies of the nonlinear problems. Technical report, Los Alamos National Laboratory, 1955.

[270] Lando Caiani, Lapo Casetti, Cecilia Clementi, Giulio Pettini, Marco Pettini, and Raoul Gatto. Geometry of dynamics and phase transitions in classical lattice φ-4 theories. *Phys. Rev. E*, 57(4):3886, 1998.

[271] R. B. Levien and S. M. Tan. Double pendulum: An experiment in chaos. *Am. J. Phys.*, 61(11):1038–1044, 1993.

[272] Michael A Nielsen and Isaac Chuang. *Quantum computation and quantum information.* Cambridge University Press, Cambridge, England, 2000.

[273] Charles Neill, Pedran Roushan, K Kechedzhi, Sergio Boixo, Sergei V Isakov, V Smelyanskiy, A Megrant, B Chiaro, A Dunsworth, K Arya, et al. A blueprint for demonstrating quantum supremacy with superconducting qubits. *Science*, 360(6385):195–199, 2018.

[274] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.

[275] Bernard Yurke and John S Denker. Quantum network theory. *Phys. Rev. A*, 29(3):1419, 1984.

[276] Juan Ignacio Cirac, Peter Zoller, H Jeff Kimble, and Hideo Mabuchi. Quantum state transfer and entanglement distribution among distant nodes in a quantum network. *Phys. Rev. Lett.*, 78(16):3221, 1997.

[277] Matthias Christandl, Nilanjana Datta, Artur Ekert, and Andrew J Landahl. Perfect state transfer in quantum spin networks. *Phys. Rev. Lett.*, 92(18):187902, 2004.

[278] AS Parkins, P Marte, P Zoller, and HJ Kimble. Synthesis of arbitrary quantum states via adiabatic transfer of zeeman coherence. *Phys. Rev. Lett.*, 71(19):3095, 1993.

[279] L-M Duan and Christopher Monroe. Colloquium: Quantum networks with trapped ions. *Rev. Mod. Phys.*, 82(2):1209, 2010.

[280] Adriano Barenco, Charles H Bennett, Richard Cleve, David P DiVincenzo, Norman Margolus, Peter Shor, Tycho Sleator, John A Smolin, and Harald Weinfurter. Elementary gates for quantum computation. *Phys. Rev. A*, 52(5):3457, 1995.

[281] Dorit Aharonov, Wim Van Dam, Julia Kempe, Zeph Landau, Seth Lloyd, and Oded Regev. Adiabatic quantum computation is equivalent to standard quantum computation. *SIAM J. Comput.*, 50(4):755–787, 2008.

[282] Simon C Benjamin and Sougato Bose. Quantum computing with an always-on Heisenberg interaction. *Phys. Rev. Lett.*, 90(24):247901, 2003.

[283] Daniel Burgarth, Koji Maruyama, Michael Murphy, Simone Montangero, Tommaso Calarco, Franco Nori, and Martin B Plenio. Scalable quantum computation via local control of only two qubits. *Phys. Rev. A*, 81(4):040303, 2010.

[284] Erik Sjöqvist, Dian-Min Tong, L Mauritz Andersson, Björn Hessmo, Markus Johansson, and Kuldip Singh. Non-adiabatic holonomic quantum computation. *New J. Phys.*, 14(10):103035, 2012.

[285] Steffen J Glaser, Ugo Boscain, Tommaso Calarco, Christiane P Koch, Walter Köckenberger, Ronnie Kosloff, Ilya Kuprov, Burkhard Luy, Sophie Schirmer, Thomas Schulte-Herbrüggen, et al. Training Schrödinger's cat: quantum optimal control. *Eur.Phys. J. D*, 69(12):279, 2015.

[286] Leonardo Banchi, Nicola Pancotti, and Sougato Bose. Quantum gate learning in qubit networks: Toffoli gate without time-dependent control. *NPJ Quantum Inf.*, 2(1):1–6, 2016.

[287] Luca Innocenti, Leonardo Banchi, Alessandro Ferraro, Sougato Bose, and Mauro Paternostro. Supervised learning of time-independent Hamiltonians for gate design. *New J. Phys.*, 2020.

[288] Shahnawaz Ahmed, Carlos Sánchez Munoz, Franco Nori, and Anton Frisk Kockum. Quantum state tomography with conditional generative adversarial networks. *arXiv:2008.03240*, 2020.

[289] Shahnawaz Ahmed, Carlos Sánchez Munoz, Franco Nori, and Anton Frisk Kockum. Classification and reconstruction of optical quantum states with deep neural networks. *arXiv:2012.02185*, 2020.

[290] Jun Zhang and Mohan Sarovar. Quantum Hamiltonian identification from measurement time traces. *Phys. Rev. Lett.*, 113(8):080401, 2014.

[291] Jun Zhang and Mohan Sarovar. Identification of open quantum systems from observable time traces. *Phys. Rev. A*, 91(5):052121, 2015.

[292] Alireza Shabani, Masoud Mohseni, Seth Lloyd, Robert L Kosut, and Herschel Rabitz. Estimation of many-body quantum Hamiltonians via compressive sensing. *Phys. Rev. A*, 84(1):012107, 2011.

[293] Easwar Magesan, Alexandre Cooper, and Paola Cappellaro. Compressing measurements in quantum dynamic parameter estimation. *Phys. Rev. A*, 88(6):062109, 2013.

[294] Zhi Li, Liujun Zou, and Timothy H Hsieh. Hamiltonian tomography via quantum quench. *Phys. Rev. Lett.*, 124(16):160502, 2020.

[295] Ludwig E De Clercq, Robin Oswald, Christa Flühmann, Ben Keitch, Daniel Kienzler, H-Y Lo, Matteo Marinelli, David Nadlinger, Vlad Negnevitsky, and Jonathan P Home. Estimation of a general time-dependent Hamiltonian for a single qubit. *Nat. Commun.*, 7(1):1–8, 2016.

[296] Eyal Bairey, Itai Arad, and Netanel H Lindner. Learning a local Hamiltonian from local measurements. *Phys. Rev. Lett.*, 122(2):020504, 2019.

[297] Jonathan Unger, S Dong, R Flores, Q Su, and R Grobe. Infinite-dimensional optimization applied to pair creation from the vacuum. *Phys. Rev. A*, 99(2):022128, 2019.

[298] Dafa Zhao, Chao Wei, Shunzhong Xue, Yulei Huang, Xinfang Nie, Jun Li, Dong Ruan, Dawei Lu, Tao Xin, and Guilu Long. Characterizing quantum simulations with quantum tomography on a spin quantum simulator. *Phys. Rev. A*, 103(5):052403, 2021.

[299] Antonio A Gentile, Brian Flynn, Sebastian Knauer, Nathan Wiebe, Stefano Paesani, Christopher E Granade, John G Rarity, Raffaele Santagati, and Anthony Laing. Learning models of quantum systems from experiments. *Nat. Phys.*, pages 1–7, 2021.

[300] Marc Timme. Revealing network connectivity from response dynamics. *Phys. Rev. Lett.*, 98(22):224101, 2007.

[301] Srinivas Gorur Shandilya and Marc Timme. Inferring network topology from complex dynamics. *New J. Phys.*, 13(1):013004, 2011.

[302] Ri-Qi Su, Wen-Xu Wang, and Ying-Cheng Lai. Detecting hidden nodes in complex networks from time series. *Phys. Rev. E*, 85(6):065201, 2012.

[303] Xiao Han, Zhesi Shen, Wen-Xu Wang, and Zengru Di. Robust reconstruction of complex networks from sparse data. *Phys. Rev. Lett.*, 114(2):028701, 2015.

[304] Wen-Xu Wang, Ying-Cheng Lai, and Celso Grebogi. Data based identification and prediction of nonlinear and complex dynamical systems. *Phys. Rep.*, 644:1–76, 2016.

[305] Yuzuru Kato and Naoki Yamamoto. Structure identification and state initialization of spin networks with limited access. *New J. Phys.*, 16(2):023024, 2014.

[306] C Di Franco, M Paternostro, and MS Kim. Hamiltonian tomography in an access-limited setting without state initialization. *Phys. Rev. Lett.*, 102(18):187203, 2009.

[307] Nicholas A Peters, Julio T Barreiro, Michael E Goggin, Tzu-Chieh Wei, and Paul G Kwiat. Remote state preparation: arbitrary remote control of photon polarization. *Phys. Rev. Lett.*, 94(15):150502, 2005.

[308] Xiao-Qi Xiao, Jin-Ming Liu, and Guihua Zeng. Joint remote state preparation of arbitrary two-and three-qubit states. *J. Phys. B: At. Mol. Opt. Phys.*, 44(7):075501, 2011.

[309] Peter J Shadbolt, Maria R Verde, Alberto Peruzzo, Alberto Politi, Anthony Laing, Mirko Lobino, Jonathan CF Matthews, Mark G Thompson, and Jeremy L O'Brien. Generating, manipulating and measuring entanglement and mixture with a reconfigurable photonic circuit. *Nat. Photonics*, 6(1):45–49, 2012.

[310] Isaac L Chuang and Michael A Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *J. Mod. Opt.*, 44(11-12):2455–2467, 1997.

[311] Seth T Merkel, Jay M Gambetta, John A Smolin, Stefano Poletto, Antonio D Córcoles, Blake R Johnson, Colm A Ryan, and Matthias Steffen. Self-consistent quantum process tomography. *Phys. Rev. A*, 87(6):062119, 2013.

[312] Jeremy L O'Brien, GJ Pryde, Alexei Gilchrist, DFV James, Nathan K Langford, TC Ralph, and AG White. Quantum process tomography of a controlled-not gate. *Phys. Rev. Lett.*, 93(8):080502, 2004.

[313] M Riebe, K Kim, P Schindler, T Monz, PO Schmidt, TK Körber, W Hänsel, H Häffner, CF Roos, and R Blatt. Process tomography of ion trap quantum gates. *Phys. Rev. Lett.*, 97(22):220407, 2006.

[314] Thomas Monz, Kihwan Kim, Wolfgang Hänsel, M Riebe, AS Villar, Philipp Schindler, Michael Chwalla, Markus Hennrich, and Rainer Blatt. Realization of the quantum toffoli gate with trapped ions. *Phys. Rev. Lett.*, 102(4):040501, 2009.

[315] Liangyu Che, Chao Wei, Yulei Huang, Dafa Zhao, Shunzhong Xue, Xinfang Nie, Jun Li, Dawei Lu, and Tao Xin. Learning quantum hamiltonians from single-qubit measurements. *Phys. Rev. Res.*, 3(2):023246, 2021.

[316] AH Myerson, DJ Szwer, SC Webster, DTC Allcock, MJ Curtis, G Imreh, JA Sherman, DN Stacey, AM Steane, and DM Lucas. High-fidelity readout of trapped-ion qubits. *Phys. Rev. Lett.*, 100(20):200502, 2008.

[317] JM Chow, L DiCarlo, JM Gambetta, A Nunnenkamp, Lev S Bishop, L Frunzio, MH Devoret, SM Girvin, and RJ Schoelkopf. Detecting highly entangled states with a joint qubit readout. *Phys. Rev. A*, 81(6):062325, 2010.

[318] Matthew Neeley, Radoslaw C Bialczak, M Lenander, Erik Lucero, Matteo Mariantoni, AD O'connell, D Sank, H Wang, M Weides, J Wenner, et al. Generation of three-qubit entangled states using superconducting phase qubits. *Nature*, 467(7315):570–573, 2010.

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

[320] Rodney J Baxter. *Exactly Solved Models in Statistical Mechanics*. Elsevier, 2016.

[321] Hidetoshi Nishimori. *Statistical Physics of Spin Glasses and Information Processing: An Introduction*. Oxford University Press, Oxford, 2001.

[322] Edward Fredkin and Tommaso Toffoli. Conservative logic. *Int. J. Theor. Phys.*, 21(3-4):219–253, 1982.

[323] Leonardo DiCarlo, Matthew D Reed, Luyan Sun, Blake R Johnson, Jerry M Chow, Jay M Gambetta, Luigi Frunzio, Steven M Girvin, Michel H Devoret, and Robert J Schoelkopf. Preparation and measurement of three-qubit entanglement in a superconducting circuit. *Nature*, 467(7315):574–578, 2010.

[324] Benjamin P Lanyon, Marco Barbieri, Marcelo P Almeida, Thomas Jennewein, Timothy C Ralph, Kevin J Resch, Geoff J Pryde, Jeremy L O'brien, Alexei Gilchrist, and Andrew G White. Simplifying quantum logic using higher-dimensional Hilbert spaces. *Nat. Phys.*, 5(2):134–140, 2009.

[325] Raj B Patel, Joseph Ho, Franck Ferreyrol, Timothy C Ralph, and Geoff J Pryde. A quantum Fredkin gate. *Sci. Adv.*, 2(3):e1501531, 2016.

[326] Arkady Fedorov, Lars Steffen, Matthias Baur, Marcus P da Silva, and Andreas Wallraff. Implementation of a Toffoli gate with superconducting circuits. *Nature*, 481(7380):170–172, 2012.

[327] Daniel Burgarth, Koji Maruyama, and Franco Nori. Coupling strength estimation for spin chains despite restricted access. *Phys. Rev. A*, 79(2):020305, 2009.

[328] Elham Hosseini Lapasar, Koji Maruyama, Daniel Burgarth, Takeji Takui, Yasushi Kondo, and Mikio Nakahara. Estimation of coupling constants of a three-spin chain: a case study of Hamiltonian tomography with nuclear magnetic resonance. *New J. Phys.*, 14(1):013043, 2012.

[329] Akira Sone and Paola Cappellaro. Exact dimension estimation of interacting qubit systems assisted by a single quantum probe. *Phys. Rev. A*, 96(6):062334, 2017.

[330] Hauke Haehne, Jose Casadiego, Joachim Peinke, and Marc Timme. Detecting hidden units and network size from perceptible dynamics. *Phys. Rev. Lett.*, 122(15):158301, 2019.

[331] Noam Shazeer, Azalia Mirhoseini, Krzysztof Maziarz, Andy Davis, Quoc Le, Geoffrey Hinton, and Jeff Dean. Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. *arXiv:1701.06538*, 2017.

[332] John Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, 2018.

[333] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S Kottmann, Tim Menke, et al. Noisy intermediate-scale quantum (nisq) algorithms. *arXiv:2101.08448*, 2021.

[334] John Denholm Lambert et al. *Numerical methods for ordinary differential systems*, volume 146. John Wiley & Sons, Inc., New York, USA, 1991.

[335] Benedict Leimkuhler and Sebastian Reich. *Simulating hamiltonian dynamics*. Number 14. Cambridge university press, Cambridge, England, 2004.

[336] Jun John Sakurai and Eugene D Commins. *Modern Quantum Mechanics*. Addison-Wesley,Reading, MA, 1995.

[337] EB Bogomolny. Smoothed wave functions of chaotic quantum systems. *Physica D*, 31(2):169–189, 1988.

[338] M. V. Berry. Quantum scars of classical closed orbits in phase-space. *Proc. R. Soc. London Series A Math. Phys. Eng. Sci.*, 423(1864):219–231, 1989.

[339] Zi-Yuan Li and Liang Huang. Quantization and interference of a quantum billiard with fourfold rotational symmetry. *Phys. Rev. E*, 101(6):062201, 2020.

[340] Richard Akis, DK Ferry, and JP Bird. Wave function scarring effects in open stadium shaped quantum dots. *Phys. Rev. Lett.*, 79(1):123, 1997.

[341] Ph Blanchard and R Olkiewicz. Decoherence-induced continuous pointer states. *Phys. Rev. Lett.*, 90(1):010403, 2003.

[342] Wojciech Hubert Zurek. Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.*, 75(3):715, 2003.

[343] DK Ferry, R Akis, and JP Bird. Einselection in action: Decoherence and pointer states in open quantum dots. *Phys. Rev. Lett.*, 93(2):026803, 2004.

[344] Roland Brunner, R Akis, DK Ferry, Friedemar Kuchar, and Ronald Meisels. Coupling-induced bipartite pointer states in arrays of electron billiards: Quantum darwinism in action? *Phys. Rev. Lett.*, 101(2):024102, 2008.

[345] DK Ferry, L Huang, R Yang, Ying-Cheng Lai, and R Akis. Open quantum dots in graphene: Scaling relativistic pointer states. *J. Phys. Conf. Ser.*, 220(1):012015, 2010.

[346] Michael Zwolak and Wojciech H Zurek. Complementarity of quantum discord and classically accessible information. *Sci. Rep.*, 3(1):1–8, 2013.

[347] Susumu Shinohara, Takahisa Harayama, Takehiro Fukushima, Martina Hentschel, Takahiko Sasaki, and Evgenii E Narimanov. Chaos-assisted directional light emission from microcavity lasers. *Phys. Rev. Lett.*, 104(16):163902, 2010.

[348] Qinghai Song, Li Ge, Brandon Redding, and Hui Cao. Channeling chaotic rays into waveguides for efficient collection of microcavity emission. *Phys. Rev. Lett.*, 108(24):243902, 2012.

[349] Hannes Bernien, Sylvain Schwartz, Alexander Keesling, Harry Levine, Ahmed Omran, Hannes Pichler, Soonwon Choi, Alexander S Zibrov, Manuel Endres, Markus Greiner, et al. Probing many-body dynamics on a 51-atom quantum simulator. *Nature*, 551(7682):579–584, 2017.

[350] Christopher J Turner, Alexios A Michailidis, Dmitry A Abanin, Maksym Serbyn, and Zlatko Papić. Weak ergodicity breaking from quantum many-body scars. *Nat. Phys.*, 14(7):745–749, 2018.

[351] Michael Schecter and Thomas Iadecola. Weak ergodicity breaking and quantum many-body scars in spin-1 XY magnets. *Phys. Rev. Lett.*, 123:147201, Oct 2019.

[352] Soonwon Choi, Christopher J Turner, Hannes Pichler, Wen Wei Ho, Alexios A Michailidis, Zlatko Papić, Maksym Serbyn, Mikhail D Lukin, and Dmitry A Abanin. Emergent su (2) dynamics and perfect quantum many-body scars. *Phys. Rev. Lett.*, 122(22):220603, 2019.

[353] Cheng-Ju Lin, Vladimir Calvera, and Timothy H. Hsieh. Quantum many-body scar states in two-dimensional Rydberg atom arrays. *Phys. Rev. B*, 101:220304, Jun 2020.

[354] Shane Dooley. Robust quantum sensing in strongly interacting systems with many-body scars. *PRX Quantum*, 2:020330, May 2021.

[355] Jie Ren, Chenguang Liang, and Chen Fang. Quasisymmetry groups and many-body scar dynamics. *Phys. Rev. Lett.*, 126:120604, Mar 2021.

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

[357] Pierre Baldi, Kevin Bauer, Clara Eng, Peter Sadowski, and Daniel Whiteson. Jet substructure classification in high-energy physics with deep neural networks. *Phys. Rev. D*, 93(9):094034, 2016.

[358] Luke de Oliveira, Michael Kagan, Lester Mackey, Benjamin Nachman, and Ariel Schwartzman. Jet-images—deep learning edition. *J. High Energy Phys.*, 2016(7):1–32, 2016.

[359] Adam Aurisano, Alexander Radovic, D Rocco, Alexander Himmel, MD Messier, E Niner, G Pawloski, Fernanda Psihas, Alexandre Sousa, and P Vahle. A convolutional neural network neutrino event classifier. *J. Instrum.*, 11(09):P09001, 2016.

[360] François Lanusse, Quanbin Ma, Nan Li, Thomas E Collett, Chun-Liang Li, Siamak Ravanbakhsh, Rachel Mandelbaum, and Barnabás Póczos. Cmu deeplens: deep learning for automatic image-based galaxy–galaxy strong lens finding. *Mon. Not. R. Astrom. Soc.*, 473(3):3895–3906, 2018.

[361] Yi Zhang, A Mesaros, Kazuhiro Fujita, SD Edkins, MH Hamidian, K Ch'ng, H Eisaki, S Uchida, JC Séamus Davis, Ehsan Khatami, et al. Machine learning in electronic-quantum-matter imaging experiments. *Nature*, 570(7762):484–490, 2019.

[362] Ya A Kharkov, VE Sotskov, AA Karazeev, Evgeniy O Kiktenko, and Aleksey K Fedorov. Revealing quantum chaos with machine learning. *Phys. Rev. B*, 101(6):064406, 2020.

[363] Li Fei-Fei, Rob Fergus, and Pietro Perona. One-shot learning of object categories. *IEEE transactions on pattern analysis and machine intelligence*, 28(4):594–611, 2006.

[364] Oriol Vinyals, Charles Blundell, Timothy Lillicrap, Daan Wierstra, et al. Matching networks for one shot learning. *Advances in neural information processing systems*, 29:3630–3638, 2016.

[365] Łukasz Kaiser, Ofir Nachum, Aurko Roy, and Samy Bengio. Learning to remember rare events. *arXiv:1703.03129*, 2017.

[366] Yu-Xiong Wang, Deva Ramanan, and Martial Hebert. Meta-learning to detect rare objects. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9925–9934, 2019.

[367] Michael V Berry. Quantum chaology (the bakerian lecture). In *A Half-Century of Physical Asymptotics and Other Diversions: Selected Works by Michael Berry*, pages 307–322. World Scientific, 1987.

[368] Michael Victor Berry and Mark Wilkinson. Diabolical points in the spectra of triangles. *Proc. R. Soc. A*, 392(1802):15–43, 1984.

[369] Marko Robnik. Quantising a generic family of billiards with analytic boundaries. *J. Phys. A*, 17(5):1049, 1984.

[370] Michael V Berry and Marko Robnik. Statistics of energy levels without time-reversal symmetry: Aharonov-bohm chaotic billiards. *J. Phys. A*, 19(5):649, 1986.

[371] E. Ott. *Chaos in Dynamical Systems*. Cambridge University Press, Cambridge, UK, second edition, 2002.

[372] T. M. Antonsen, E. Ott, Q. Chen, and R. N. Oerter. Statistics of wave-function scars. *Phys. Rev. E*, 51:111–121, Jan 1995.

[373] Alex Nichol, Joshua Achiam, and John Schulman. On first-order meta-learning algorithms. *arXiv:1803.02999*, 2018.

[374] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. pages 2980–2988, 2017.

[375] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. pages 448–456, 2015.

APPENDIX A

PUBLICATIONS

Following are the relevant publications on the topics presented in this dissertation

C.-D. Han, H.-Y. Xu, and Y.-C. Lai*, "Electrical confinement in a spectrum of two-dimensional Dirac materials with classically integrable, mixed, and chaotic dynamics," Physical Review Research **2**, 013116, (2020). [Chapter2]

C.-D. Han, B. Glaz, M. Haile, and Y.-C. Lai*, "Adaptable Hamiltonian neural networks," Physical Review Research **3**, 023156, (2021). [Chapter3]

C.-D. Han, B. Glaz, M. Haile, and Y.-C. Lai*, "Tomography of time-dependent quantum Hamiltonians with machine learning," Physical Review A **104**, 062404, (2021). [Chapter4]

C.-D. Han, C.-Z. Wang and Y.-C. Lai*, "Classifying and detecting quantum scars by machine learning," To be submitted. [Chapter5]