PHYSICAL ARTIFICIAL NEURAL NETWORKS

OPTICAL AND ATOMIC APPROACHES TO DEEP LEARNING

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Vom Fachbereich Physik
der Technischen Universität Darmstadt

zur Erlangung des Grades
eines Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation von
KAI NIKLAS HANSMANN

ERSTGUTACHTER: Prof. Dr. Reinhold Walser
ZWEITGUTACHTER: Prof. Dr. Enno Giese

Darmstadt 2024

Kai Niklas Hansmann:
Physical Artificial Neural Networks
Optical and Atomic Approaches to Deep Learning

ERSTGUTACHER: Prof. Dr. Reinhold Walser
ZWEITGUTACHTER: Prof. Dr. Enno Giese

TAG DER EINREICHUNG: 06.12.2023
TAG DER PRÜFUNG: 15.01.2024

DARMSTADT, Technische Universität Darmstadt
Jahr der Veröffentlichung der Dissertation auf TUprints: 2024
URN: urn:nbn:de:tuda-tuprints-269156
URI: *https://tuprints.ulb.tu-darmstadt.de/id/eprint/26915*

„So eine Arbeit wird eigentlich nie fertig, man muß sie für fertig erklären, wenn man nach Zeit und Umständen das Möglichste gethan hat."

JOHANN WOLFGANG VON GOETHE - ITALIENISCHE REISE

„Mach immer zuerst das, was du kannst."

DR. WERNER HANSMANN

ZUSAMMENFASSUNG

Maschinelles Lernen hat in den vergangenen Jahren zahllose Anwendungen in Industrie, Wissenschaft sowie im alltäglichen Leben gefunden. Die Umsetzung solcher Systeme stützt sich dabei vorwiegend auf die immer weiter steigenden Kapazitäten von digitalen Computern. Aber auch physikalische Systeme - und im Speziellen deren Dynamik - haben bereits ihr Potential bewiesen als mögliche Implementierungsplattformen für maschinelles Lernen zu dienen. Ein besonderer Schwerpunkt wurde in dieser Diskussion auf optische Systeme gelegt. Sowohl Freistrahlsysteme als auch integrierte photonische Plattformen wurden dabei als potentielle Realisierungen für optische neuronale Netze in Betracht gezogen.

Das Ziel dieser Arbeit ist es zur Suche nach neuen physikalischen Implementierungen von künstlichen neuronalen Netzwerken beizutragen. Auf dem Gebiet von optischen neuronalen Netzen wird dazu eine Beschreibung von temperaturabhängiger Unterdrückung von Intensitätsrauschen in Quantenpunkt-Superlumineszenzdioden entwickelt. Diese Lichtquellen emittieren gerichtete Strahlung mit hoher Ausgangsleistung und großer spektraler Breite. Typischerweise unterliegt das emittierte Licht solcher Dioden thermischen Fluktuationen, charakterisiert durch den zentralen Kohärenzgrad zweiter Ordnung $g^{(2)}(0) = 2.0$. Im Jahr 2011 demonstrierten Experimente in der Gruppe von Prof. Dr. W. Elsäßer an der Technischen Universität Darmstadt, dass beim Abkühlen der Diode auf $T = 190\,\mathrm{K}$ Intensitätsfluktuationen auf $g^{(2)}(0) = 1.33$ reduziert werden. Dieser Effekt kann durch eine Manipulation der Photonenstatistik der Diodenemission, verursacht durch die Wechselwirkung mit dem gepumpten Diodenmaterial, beschrieben werden. Dies erlaubt die Interpretation der experimentellen Beobachtungen als temperaturgetriebenen Sättigungseffekt.

Neben puren optischen Implementierungen von künstlichen neuronalen Netzwerken legt diese Arbeit einen speziellen Fokus auf atomare Systeme als potentielle Plattformen für maschinelles Lernen. Im Speziellen werden dazu die Bewegung thermischer Atome in optisch modellierten Potentiallandschaften sowie die nichtlineare Dynamik von kohärenten Materiewellen in Bose-Einstein Kondensaten genutzt. Durch thermische Atome innerhalb eines Boxpotentials kann ein Neuronenalgorithmus realisiert werden, indem optische Dipolpotentiale als Eingabe und Messungen von Teilchenzahlen am Ende der Prozedur als Ausgabe interpretiert werden. Dieses Neuron zeigt vergleichbare Eigenschaften zu konventionellen Implementierungen und ist in der Lage, Standardprobleme aus dem Bereich des maschinellen Lernens zu lösen.

Mit einem noch stärkeren Fokus auf den atomaren Aspekt einer möglichen Implementierung von neuronalen Netzwerken behandelt diese Arbeit die Verwendung von intrinsischen Nichtlinearitäten in kohärenten Materiewellen. Dazu wird eine detaillierte Beschreibung der Vierwellenmischung von ebenen Wellen in einem homogenen Bose-Einstein Kondensat präsentiert. Durch analytische und numerische Untersuchung der Dynamik dieses Prozesses offenbaren sich Josephson-ähnliche Oszillationen. Die Implementierung eines komplexwertigen Neurons basierend auf dem Vierwellenmischprozess wird demonstriert, indem die komplexen Amplituden dreier Impulszustände des Prozesses als Eingabe und die verbleibende vierte Amplitude als Ausgabe interpretiert werden. Ein einziges dieser Neuronen ist bereits in der Lage, das XOR-Problem zu lösen. Durch Parallelisierung und den Aufbau kommunizierender Schichten eines Netzwerks wird gezeigt, dass die nichtlineare Dynamik in Bose-Einstein Kondensaten in der Tat genutzt werden kann, um atomare neuronale Netzwerke zu implementieren.

ABSTRACT

In recent years, machine learning techniques have found countless applications in industry, research, as well as everyday life. The realization of such systems relies heavily on the ever-growing capabilities of digital computing systems. However, physical systems, and especially their dynamics, have already proved their potential to be alternative implementation platforms. Especially optical systems have been a focal point of this research with proposals of free-space and integrated photonics optical neural networks.

In this thesis, we aim to aid in the search of new physical implementation platforms of artificial neural networks. Contributing to the field of optical neural networks, we develop a description of temperature-dependent intensity noise suppression in quantum dot superluminescent diodes. Those light sources emit spatially directed radiation, which is high-powered, spectrally broadband and typically subject to thermal fluctuations, described by the central degree of second-order coherence $g^{(2)}(0) = 2.0$. In 2011, the group of Prof. Dr. W. Elsäßer at the Technical University Darmstadt observed that intensity fluctuations are reduced to $g^{(2)}(0) = 1.33$ when tuning the running temperature of the diode to $T = 190\,\mathrm{K}$. We show, that this effect can be described via a photon statistics manipulation of the emission due to interaction with the pumped diode material. Therefore, intensity noise suppression in such diodes can be explained as a temperature driven saturation effect.

Furthermore, we go beyond purely optical implementations of artificial neural networks looking at atomic systems as the main focus of potential new machine learning platforms. Specifically, we make use of the movement of thermal atoms in optically shaped potential landscapes as well as harnessing the nonlinear dynamics of coherent matter waves in Bose-Einstein condensates. Considering thermal atoms inside a box potential, we demonstrate that a neuron algorithm can be implemented using optical dipole potentials as input and the measurement of the particle number remaining at the end of the procedure as output. This thermal cloud neuron shows similar features compared to conventional implementations and is able to solve benchmark problems.

Moving even more towards a purely atomic implementation of an artificial neural network, this thesis describes an implementation based on intrinsic nonlinearities of coherent matter waves. There, we present an in-detail description of the four-wave mixing process of plane waves in a homogeneous condensate. Analytical as well as numerical investigations of the dynamics of this process are performed, revealing Josephson-like oscillations within the four-wave mixing states. Interpreting the complex amplitudes of three momentum components of the four-wave mixing process as input, and the fourth one as output, we demonstrate the implementation of a complex-valued neuron. A single realization of such a four-wave mixing neuron is able to solve the benchmark XOR problem. By parallelizing such neurons and setting up communicating layers of a network, we show that nonlinear dynamics of Bose-Einstein condensates can be used to implement an atomic neural network.

CONTENTS

1

INTRODUCTION

In the last couple of decades, machine learning techniques have seen a rapid rise to prominence, finding ways to significance in nearly all aspects of current everyday life [1]. By learning to recognize patterns in data using enormous training datasets, these techniques have proven to be a disruptive technology in fields such as health care [2–4], finance [5, 6], manufacturing [7, 8], energy generation [9–11], autonomous driving [12–14] and many more. An impact of comparable magnitude by a technology on industry and research has not been seen since the introduction of computers [15]. Especially deep learning techniques have been used successfully to improve the performance of machine learning systems, going beyond classical pattern recognition and moving more towards the realm of artificial intelligence.

Yet, the increase in complexity, and therefore capability, of machine learning systems is starting to point out limitations and possible problems. Especially the energy consumption of modern computing systems, on which machine learning techniques are traditionally implemented, has become a point of concern. To power the artificial neural networks AlphaGo and AlphaZero, which can beat human experts at strategy games, roughly 200 W are needed [16]. Putting this into perspective, the human brain consumes approximately 20 W of power while performing many more parallel tasks than just playing a strategy game. Using current technology, the power consumption of an artificial neural network the size of the human brain would be around 7.9 MW [17]. Therefore, there is a growing need for alternative implementation platforms of machine learning systems beyond digital computing systems based on the classical von Neumann architecture.

OPTICAL APPROACHES TO DEEP LEARNING

One platform of high interest are optical implementations of neural networks [18]. With their potential of massively parallelized data processing using free-space optics as well as integrated photonic chips, while consuming relatively small amounts of power, optical neural networks have gathered significant attention in recent years.

In this thesis, we aim to add to the field of optical neural network implementations by characterizing the temperature-dependent intensity noise characteristics of a specialized light source, the quantum dot superluminescent diode. Such diodes are already being implemented in realizations of machine learning systems [19–23]. However, suppressing intensity noise while conserving the broadband emission properties of the diode, yielding silent white light, has the potential to increase the performance of such networks even further.

The basis of our discussion is given by experimental measurements of *M. Blazek* and *W. Elsäßer* from 2011 [24]. They observed a reduction of intensity noise in such diodes in terms of the central degree of second-order coherence $g^{(2)}(0)$ from thermal noise $g^{(2)}(0) = 2.0$ at room temperature to about $g^{(2)}(0) = 1.33$ when tuning the temperature to $T = 190\,\mathrm{K}$. Here,

we present a description of this effect via photon statistics manipulation of the emission due to the interaction with the pumped diode material. In doing so, we are able to interpret the observations as a temperature driven saturation effect inside the diode material. A better understanding of such effects might aid the determination of optimal operating conditions of these diodes, as well as possibly designing new diodes with optimized emission characteristics. Both those advances would also prove to be beneficial in the implementation of optimized diodes in artificial neural networks.

ATOMIC APPROACHES TO DEEP LEARNING

In addition to investigating optical systems and their possible applications in machine learning, we focus on another approach to set up artificial neural networks using atomic systems. Up to this point, the role of atomic systems in the realm of machine learning has usually been restricted to certain parts of implementations, for example by introducing nonlinearity to optical neural networks [25]. In this thesis, we investigate two systems, the thermal cloud neural network and the four-wave mixing neural network, in which the main focus of the implementation of the artificial neural networks lies on the atomic system itself.

In a first approach to an atomic artificial neural network, we investigate a classical thermal gas in optically shaped potential landscapes. In the sense of the thermal cloud neural network, a neuron is realized by implementing an input via an optical dipole potential. The externally trapped thermal atoms equilibrate in the new landscape. By turning off the external trap, the particle number still trapped by the input laser can be interpreted as the output of the neuron. We investigate the nonlinear activation function of such a neuron and show that a network of thermal cloud neurons is able to solve benchmark problems of machine learning.

Going beyond classical atomic gases, we use the highly nonlinear dynamics of coherent matter waves in Bose-Einstein condensates to set up an artificial neural network. During the advent of Bose-Einstein condensation in dilute atomic gases, nonlinear effects in the condensate have been widely used to determine the occurrence of a phase transition [26, 27]. In this thesis, we focus on the process of four-wave mixing in coherent matter waves, which is well-known from nonlinear optics [28]. There, if phase-matching conditions are fulfilled in a third-order nonlinear medium, three frequencies of an optical field can interact in a way such that an initially absent fourth frequency emerges. An equivalent process in coherent matter waves was predicted theoretically [29, 30] and observed experimentally [31, 32]. There, momentum components of the matter waves took over the role of optical frequencies from the initial scenario.

Considering the idealized case of a homogeneous periodic Bose-Einstein condensate, we show that the four-wave mixing process using plane waves exhibits Josephson-like oscillations [33–35]. Based on these results, we demonstrate the implementation of a complex-valued neuron, where we identify the complex amplitudes of three momentum components of the four-wave mixing process as input and the fourth component as output. Using an in-detail investigation of the nonlinear activation function of this implementation, we show that steepest descent learning can be used to train such four-wave mixing neurons by solving the benchmark XOR problem.

Expanding on this implementation, we subsequently investigate the parallelization ability of the four-wave mixing neuron and search for possible ways to implement communication between multiple layers. Especially using the extended internal structure of the condensate atoms turns out to be beneficial in setting up a neural network structure. As a proof of principle, we train a four-wave mixing neural network to distinguish logical six-bit sequences using the

backpropagation algorithm. The success of this procedure shows that also larger network structures can be constructed using the four-wave mixing neuron.

STRUCTURE OF THESIS

This thesis is structured in five parts. In Part I, an overview over artificial neural networks is presented. After discussing the inspirational sources of machine learning, biological neurons and the human brain, the basic concept of an artificial neuron is presented in Chapter 2. This discussion is expanded by introducing artificial neural networks in Chapter 3. Starting from simple single-layer models, training processes of deep neural networks using the backpropagation algorithm are introduced. As performance benchmarks for neural network implementations, we discuss the solution of the XOR problem, as well as the recognition of handwritten digits based on the MNIST dataset. State-of-the-art physical implementations of artificial neural networks are discussed in Chapter 4.

In Part II, an example case of a neural network implementation is given using thermal atoms in optically shaped potential landscapes. There, the requirements of setting up such an implementation can be investigated in detail. After discussing the stochastic simulation of thermal atoms in Chapter 5, the algorithm for implementing an artificial neuron via shaping of optical dipole potentials is described in Chapter 6, yielding the thermal cloud neuron. Subsequently, in Chapter 7, this neuron concept is embedded into a network structure, constituting the thermal cloud neural network and showing the learning capabilities of this implementation.

Going beyond thermal atoms, a neural network implementation based on coherent matter waves in Bose-Einstein condensates, given by the four-wave mixing neural network, is presented in Part III. The properties, different forms of realizations and descriptions of the dynamics of such condensates given by the Gross-Pitaevskii equation are discussed in Chapter 8. Focussing on a particular nonlinear effect in Bose-Einstein condensates, the four-wave mixing process using a superposition of plane waves is described in detail in Chapter 9. Using a dimensionless description of the dynamics of such a system and appropriate coordinates reveals Josephson-like oscillations, which can be described analytically as well as numerically. This detailed investigation of the four-wave mixing process is exploited in Chapter 10 to develop an algorithm to implement a complex-valued neuron, the four-wave mixing neuron. Numerically investigating the nonlinear activation function of this neuron enables subsequent training using steepest descent learning and solving the XOR problem with just a single complex-valued neuron. By showing ways to parallelize four-wave mixing neurons and to set up communicating layers of a network, the four-wave mixing neural network is introduced in Chapter 11. Using a sample training task, the learning capabilities of this system are demonstrated, proving that nonlinear dynamics in Bose-Einstein condensates can indeed be used to set up trainable feedforward neural networks.

In Part IV, we present an investigation of temperature-dependent intensity noise suppression in quantum dot superluminescent diodes, yielding silent white light. Using a stochastic model of the diode emission, its first- and second-order temporal correlation properties are investigated in Chapter 12. In addition to optical power spectra and second-order temporal correlation functions, this reveals an exponential probability distribution of the emitted intensity. Those emission statistics can be manipulated via the interaction with the pumped diode material. In Chapter 13, we show that describing the diode material as pumped three-level atoms, a modulation of the central degree of second-order coherence $g^{(2)}(0)$ in dependence of the system parameters can be observed. Using experimental measurements regarding the temperature dependence of the emitted mean intensity of the diode, we are able to recon-

struct the observed intensity noise reduction to around $g^{(2)}(0) = 1.33$ at $T = 190\,\mathrm{K}$. Our model enables an interpretation of the observed effect as a temperature driven saturation effect in the diode material.

Finally, this thesis concludes with a summary of the results and an outlook to further research opportunities and directions in Part V. Central results of this thesis are submitted for publication or have already been published in scientific journals:

- K. N. Hansmann, R. Walser [36]
  Forming complex neurons by four-wave mixing in a Bose-Einstein condensate
  Physical Review A **109**, 013302 (2024)
  Based on Chapters 9 and 10.

- K. N. Hansmann, R. Walser [37]
  Stochastic Simulation of Emission Spectra and Classical Photon Statistics of Quantum Dot Superluminescent Diodes
  Journal of Modern Physics **12**, 22-34 (2021)
  Based on Chapter 12.

- K. N. Hansmann, F. Dommermuth, W. Elsäßer, R. Walser [38]
  Silent White Light
  Submitted to Physical Review Letters (2023)
  Based on Chapter 13.

Part I

ARTIFICIAL NEURAL NETWORKS

2

ARTIFICIAL NEURONS

Artificial neural networks (ANNs) are systems, which were developed to simulate the calculation processes inside the human nervous system. The computational units of such a network, called artificial neurons (ANs), are therefore modelled after human neurons. The aim of this implementation is the development of artificial intelligence (AI) which performs calculations in the same manner as the brain. Due to the enormous complexity of the human brain this is a difficult task.

Early developments of ANNs came shortly after the proposal of the modern computer by *Alan Turing* in 1936 [39]. The first model of a neuron was introduced by *McCulloch* and *Pitts* in 1943 [40]. They created algorithms called threshold logic to mimic the response characteristics of human neurons. The next advancement was achieved by *Frank Rosenblatt* in 1958 when he introduced the concept of the perceptron [41]. This network is build around a McCulloch-Pitts neuron and was the first model to implement supervised learning in ANNs. However, the need of big data sets, high computational demands and limited abilities of the perceptron led to concerns about the usability of ANNs [42]. As the processing capabilities of computers developed through the decades, more and more achievements were made regarding ANNs. In 1974 *John Werbos* introduced the backpropagation algorithm for the training of multilayer ANNs [43]. This led to the possibility of developing larger and more complex network structures, which have come to be known as deep learning [44].

This first part of the presented thesis introduces the basic concepts of ANs (Chapter 2) and ANNs (Chapter 3). Subsequently, state-of-the-art physical implementations of ANNs are reviewed in Chapter 4. In this chapter, the biological inspiration of ANs is discussed in Section 2.1, before introducing the theoretical description of ANs in Section 2.2.

2.1 BIOLOGICAL NEURONS AND THE HUMAN BRAIN

2.1.1 Human Information Processing

Humans and all other living organisms constantly interact with their environment. To be able to respond properly to their surroundings, the flood of information received from the environment has to be processed. This happens in the nervous system of a human and can be cut down into three stages: processing of sensory inputs, sensory integration and processing

Figure 2.1: Schematic depiction of information processing in the human nervous system. Graphic adapted from [45].

of motoric output signals (see Fig. 2.1) [46]. All these tasks in the nervous system are performed by specialized cells called neurons.

The main part of signal processing is performed in the human brain. The brain hosts around 86 billion neurons, where some neurons have up to 100,000 connections to their neighbours [46]. This in total presents a calculation capacity surpassing those of state-of-the-art supercomputers. In modelling how learning happens or could happen in the brain, the field of learning algorithms and ANNs came to fruition [44].

2.1.2 Neural Structure and Functionality

The basic computation units of the human nervous system are specialized cells called neurons [47]. They appear in a vast variety of shapes and forms. A schematic visualization of an exemplary neuron can be seen in Fig. 2.2.

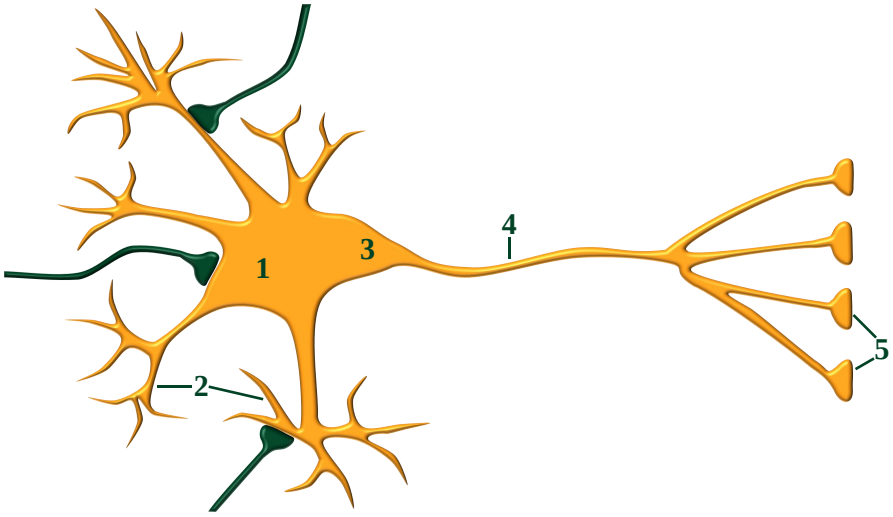

Figure 2.2: Schematic depiction of an exemplary neuron with soma (1), dendrites (2), axon hillock (3), axon (4) and axon terminals (5). Graphic adapted from [47].

A neuron consists of a cell body, or soma, and a number of cellular extensions, which vary in their functionality. The soma contains the nucleus of the neuron and is the location of protein synthesis. Information is passed on between neurons on connection sites called synapses. On the receiving neuron, these can either be located on the soma itself or on tree-like extensions, called dendrites. If an electric signal arrives at the synapse, calcium ions enter the small gap between the neurons, diffuse to the receiving side and activate receptors there. Synapses can either increase (excitatory) or decrease (inhibitory) the activity in the target neuron. The magnitude of influence a synapse has on the receiving neuron depends on the location of the synapse on the neuron and strength of the synaptic connection.

All incoming signals arrive at the soma and reach a region called axon hillock. If the incoming potential surpasses a neuron-specific activation potential, an electric signal is produced in the axon hillock. This response has an ”all-or-none” characteristic. If the threshold is surpassed, the neuron produces an outgoing signal at full strength. Otherwise, there is no response at all. Therefore, neurons show a highly nonlinear response to incoming signals. The outgoing signal propagates along the axon towards the axon terminals, where it is passed on to following neurons [47].

Connections inside the brain can be altered depending on their importance. For example, if the activity of one synapse is correlated to the activity of another one, changes can appear that strengthen this synaptic connection. Furthermore, new synapses can be formed if the activity of two neurons is correlated. Changes in the synaptic regions can also alter the strength of the connection produced by the synapse. All these processes are a part of the neural plasticity of the brain [46], which leads to a change in positions and strengths of connections between neurons, and therefore learning in the brain.

2.2 ARTIFICIAL NEURONS

On the basis of biological neurons, a model for ANs as the basic constituents of ANNs has been introduced by *McCulloch* and *Pitts* in 1943 [40] and has been further developed by *Rosenblatt* in 1958 [41]. These neurons posses three basic properties:

- Input signals x_j reach a neuron k via synapses with scalar weights w_{kj},
- a linear combiner sums up all weighted input signals,
- an activation function produces an output signal y_k.

A schematic representation of the working principle of an AN can be seen in Fig. 2.3.

All inputs x_j presented to a neuron k are weighted with the respective weights w_{kj} and produce the induced local field v_k at the summing junction as

$$v_k = \sum_{j=0}^{n} w_{kj} x_j. \tag{2.1}$$

The term w_{k0}, where $x_0 = 1$ is always fixed, describes an externally applied bias, increasing or decreasing the net input of the neuron. The output is produced by applying a nonlinear activation function (NAF) φ to the induced local field

$$y_k = \varphi(v_k). \tag{2.2}$$

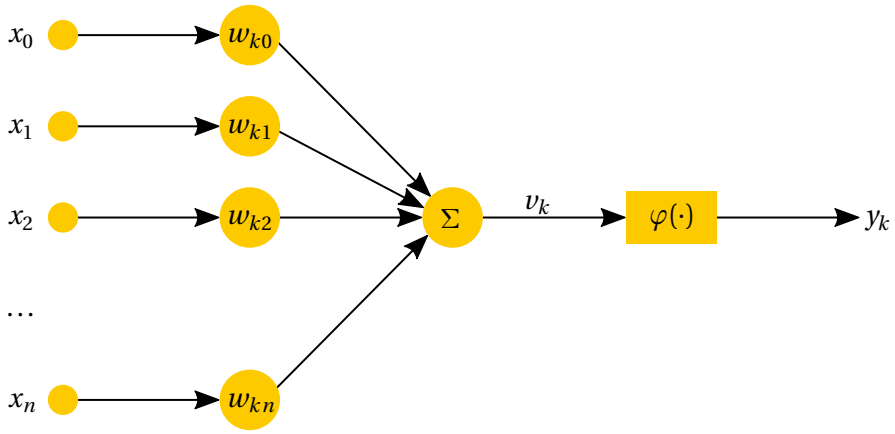

Figure 2.3: Model of an artificial neuron. A total of n inputs x_j are weighted with respective weights w_{kj}. A fixed input $x_0 = 1$ in combination with w_{k0} represents a neuron-specific bias. The weighted inputs are summed up at a summing junction to produce the local induced field v_k. This signal is processed by a nonlinear activation function $\varphi(v_k)$ producing the output y_k. Graphic adapted from [45].

The NAF $\varphi(v_k)$ thus defines the output y_k of a neuron. Depending on the application of the neuron and the problem at hand, these NAFs take on vastly different forms. Some typical examples are:

- Threshold/Heaviside function

$$\varphi_{\mathrm{H}}(v_k) = \begin{cases} 1, & \text{for } v_k \geq 0, \\ 0, & \text{for } v_k < 0. \end{cases} \tag{2.3}$$

- Piecewise-linear function

$$\varphi_{\mathrm{pl}}(v_k) = \begin{cases} 1, & \text{for } v_k \geq 1/2, \\ v_k + 1/2, & \text{for } 1/2 > v_k > -1/2, \\ 0, & \text{for } -1/2 \geq v_k. \end{cases} \tag{2.4}$$

- Sigmoidal function/Fermi distribution

$$\varphi_{\mathrm{s}}(v_k) = \frac{1}{1+\exp(-\alpha v_k)}. \tag{2.5}$$

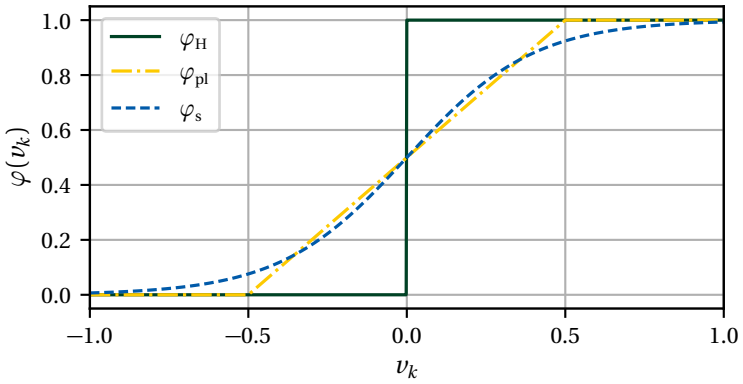

Figure 2.4: Nonlinear activation function response φ to induced local field v_k. Exemplary types of nonlinearities are threshold φ_{H} (green, solid), piecewise-linear φ_{pl} (yellow, dash-dotted) and sigmoidal φ_{s} (blue, dashed; $\alpha = 5$) functions.

3

ARTIFICIAL NEURAL NETWORKS

TYPES OF ARTIFICIAL NEURAL NETWORKS All neural networks have neurons as their basic constituents. However, networks might differ in their architecture, which typically influences the algorithm used in training the network. In a layered network the neurons are organized in the form of layers. Computations in one layer take place at the same time. A network is called feedforward, if the input layer projects onto the output layer but not vice versa. Accordingly, there are no feedback connections in which outputs of neurons are fed back to themselves. When such feedback connections are present, one has to deal with the topic of recurrent neural networks [48–50].

As neural networks typically posses a significant number of layers, a depth can be assigned to the architecture, giving rise to the term deep learning [44]. Such deep neural networks can be further subcategorized by the way in which the layers are connected. If every neuron in a layer is connected to every single neuron in the adjacent layers, the network is said to be fully-connected. If there are only partial connections between the layers, one enters the realm of convolutional neural networks [51].

HISTORY AND PERFORMANCE CAPABILITIES OF ARTIFICIAL NEURAL NETWORKS Three main periods can be identified in the history of ANNs: cybernetics in the 1940s to 1960s [40, 41, 52, 53], connectionism in the 1980s and 1990s [49, 54, 55] and deep learning from 2006 to present day [56–59]. Over recent years, faster computers with larger memories and the introduction of hidden layers have led to a doubling of ANN sizes roughly every 2.4 years [44].

As larger networks are able to achieve higher accuracy on more complex tasks, also the size of available datasets has grown significantly. A rough rule of thumb states, that a supervised deep learning algorithm will achieve acceptable performance with around 5,000 labeled examples and will match or surpass human performance when trained with at least 10 million labeled examples [44]. Such large datasets, available for ANN training, are, for example, MNIST [60], CIFAR-10 [61], ImageNet [62, 63], Sports-1M [64] or IBM's dataset constructed from the Canadian Hansard [65].

Nowadays, ANNs show excellent performance in tasks like image recognition [66], speech recognition [67, 68], pedestrian detection [69] and traffic sign classification [70]. Many leading companies, including Google, Microsoft, IBM, NVIDIA and others, rely on deep learning and a lot of commercial programs and software libraries like Torch [71], TensorFlow [72] and ChatGPT are readily available.

In this thesis, we aim to introduce new ways to implement artificial neurons and set up neural networks. Therefore, we limit the following discussion of architecture and learning algorithms to feedforward fully-connected neural networks, as this will be the way in which our implementations will be realized.

In this chapter, the concepts of input processing and learning in single-layer (Section 3.1) and multilayer (Section 3.2) feedforward neural networks are discussed. In Section 3.3, we introduce two benchmark problems, the XOR problem and the recognition of hand-written digits based on the MNIST dataset, against which the implementations of ANNs will be tested throughout this thesis.

3.1 SINGLE-LAYER FEEDFORWARD NETWORKS AND ERROR-CORRECTION LEARNING

In a single-layer feedforward network, the input of source nodes projects directly onto the output layer of calculation neurons. Source nodes do not count as a layer, as no computations are performed. An exemplary block diagram of such a simple network can be seen in Fig. 3.1. One possible definition of learning in the context of ANNs is that it "is a process by which the free parameters of a neural network are adapted through a process of stimulation by the environment in which the network is embedded. The type of learning is determined by the manner in which the parameter change takes place" [73]. One popular algorithm for the training of single-layer feedforward networks is so-called error-correction learning [45].

The signal-flow graph for error-correction learning of a single neuron can be seen in Fig. 3.2. Consider a single-layer feedforward network with n input nodes and q output nodes. The network is stimulated by an input vector $\boldsymbol{x}^{(i)} = \left(x_1^{(i)}, \ldots, x_n^{(i)}\right)$, where i denotes the instant in time at which the excitation is applied to the system. The external behaviour of the system is described by the training dataset

$$\mathcal{T} : \left\{\boldsymbol{x}^{(i)}, \boldsymbol{d}^{(i)}; i = 1, \ldots, m\right\}, \tag{3.1}$$

where $\boldsymbol{d}^{(i)} = \left(d_1^{(i)}, \ldots, d_q^{(i)}\right)$ is the desired response associated with $\boldsymbol{x}^{(i)}$ and m is the size of the dataset. In response to the stimulus, the network produces an output $\boldsymbol{y}^{(i)} = \left(y_1^{(i)}, \ldots, y_q^{(i)}\right)$ (2.2).

Starting from an arbitrary setting of the synaptic weights $W = (w_{kj})$, $k = 1, \ldots, q$, $j = 1, \ldots, n$, the goal of the learning procedure is to adjust the weights to minimize the difference between the desired and actual outputs, described by means of a cost function $\mathcal{E}(W)$. A typical cost function is the squared error averaged over the training sample set [45]

$$\mathcal{E} = \frac{1}{m} \sum_{i=1}^{m} \mathcal{E}^{(i)}. \tag{3.2}$$

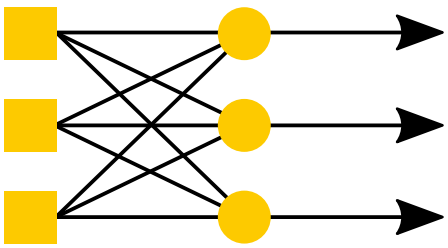

Figure 3.1: Exemplary block diagram of a single-layer feedforward network. The squares symbolize source nodes where no calculations are performed, while circles represent artificial neurons. Signals flow from left to right through the network. Graphic adapted from [45].

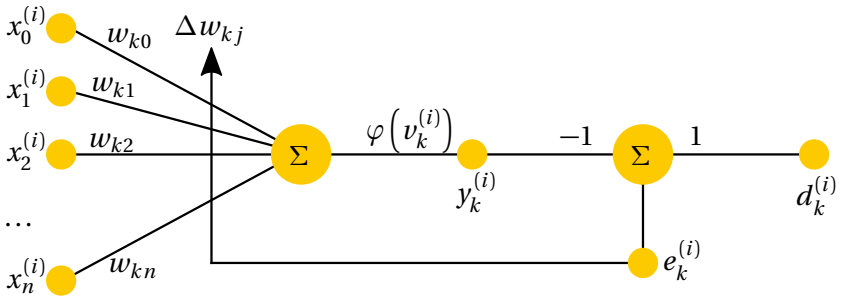

Figure 3.2: Signal-flow graph for error-correction learning in a single-layer feedforward network with one output neuron. At an instant i, the network output $y_k^{(i)}$ is compared to a desired response $d_k^{(i)}$, which together produce an error signal $e_k^{(i)}$. This can be used to update the neuron weights, minimizing a cost function $\mathcal{E}(W)$. Graphic adapted from [45].

There, $\mathcal{E}^{(i)}$ is the total instantaneous error measure

$$\mathcal{E}^{(i)} = \frac{1}{2}\sum_{k=1}^{q}\left|e_k^{(i)}\right|^2, \qquad e_k^{(i)} = d_k^{(i)} - y_k^{(i)}, \tag{3.3}$$

with $e_k^{(i)}$ the error signal produced by neuron k.

This yields an unconstrained optimization problem [74], meaning to find an optimal solution W^* such that

$$\mathcal{E}(W^*) \leq \mathcal{E}(W) \tag{3.4}$$

for all weight matrices W. The necessary condition for optimality is

$$\nabla\mathcal{E}(W^*) = 0, \tag{3.5}$$

where ∇ is the gradient operator in weight space.

In a steepest descent method [74], adjustments applied to the weight matrix W are in the direction of steepest descent, that is in the direction of the negative gradient. An update rule for an iterative procedure can therefore be formulated as

$$\Delta W(n) = W(n+1) - W(n) = -\eta\nabla\mathcal{E}(W), \tag{3.6}$$

where n symbolizes one iteration and η is a positive learning rate. Depending on the way supervised learning is actually performed, two different methods can be identified: batch learning and on-line learning.

3.1.1 Batch Learning

In batch learning [45], adjustments to the synaptic weights are performed after all m examples in the training sample set $\mathcal{T}$ have been presented to the network (see Fig. 3.3). One run through the sample set $\mathcal{T}$ is called an epoch. Accordingly, adjustments are made on an epoch-by-epoch basis. Inserting (3.2) into (3.6) and using the chain rule, (2.1) and (2.2) yields

$$\frac{\partial\mathcal{E}}{\partial w_{kj}} = \frac{1}{m}\sum_{i=1}^{m}\frac{\partial\mathcal{E}^{(i)}}{\partial e_k^{(i)}}\frac{\partial e_k^{(i)}}{\partial y_k^{(i)}}\frac{\partial y_k^{(i)}}{\partial v_k^{(i)}}\frac{\partial v_k^{(i)}}{\partial w_{kj}} = -\frac{1}{m}\sum_{i=1}^{m}e_k^{(i)}\varphi'\left(v_k^{(i)}\right)x_j^{(i)}. \tag{3.7}$$

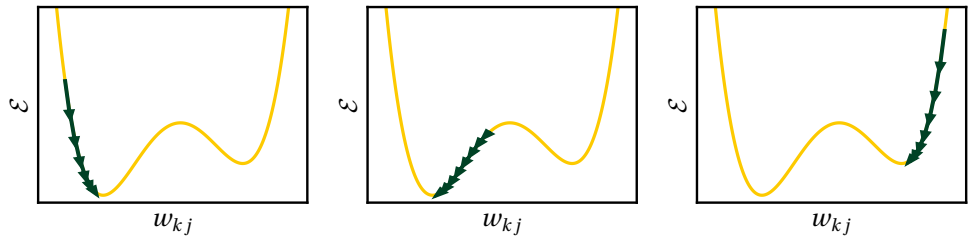

Figure 3.3: Visualization of steepest descent method, cost function $\mathcal{E}$ versus synaptic weight w_{kj}. The cost function is iteratively minimized by changing the synaptic weight along the negative of the gradient $\partial_{w_{kj}}\mathcal{E}$. Depending on the starting point and learning rate, global or local minima might be reached.

Consequently, the update rule for batch learning is given by

$$\Delta w_{kj} = \frac{\eta}{m}\sum_{i=1}^{m}\delta_{kj}^{(i)}x_j^{(i)}, \tag{3.8}$$

with the local gradient

$$\delta_{kj}^{(i)} = e_k^{(i)}\varphi'\left(v_k^{(i)}\right). \tag{3.9}$$

Batch learning yields an accurate estimation of the gradient vector, thereby guaranteeing convergence to a local minimum. Furthermore, a high degree of parallelization is possible during the training process. However, batch learning is rather demanding in terms of storage requirements.

3.1.2 On-Line Learning

In on-line learning [45], adjustments to the weights are performed on an example-by-example basis. The cost function to be minimized is therefore the total instantaneous error measure $\mathcal{E}^{(i)}$ (3.3). An epoch consists of m training samples. At an instant i, a pair $\left\{\boldsymbol{x}^{(i)}, \boldsymbol{d}^{(i)}\right\}$ is presented to the network and weight adjustments are performed using the method of steepest descent according to

$$\Delta w_{kj}^{(i)} = \eta\delta_{kj}^{(i)}x_j^{(i)}. \tag{3.10}$$

Subsequently, the next sample is presented to the network. This process is repeated until all m samples of the epoch have been presented.

Given the training samples are presented to the neuron in a random manner, the use of the on-line learning makes the search in multi-dimensional weight space stochastic in nature [45]. This stochasticity has the desirable effect of making it less likely for the learning process to get stuck in a local minimum. While using less storage and avoiding local minima, on-line learning limits the possibility of parallelization in the learning process. Yet, as it is easy to implement and it provides effective solutions to large-scale and difficult pattern-classification problems, we will make use of on-line learning throughout this thesis.

3.2 MULTILAYER FEEDFORWARD NETWORKS AND BACKPROPAGATION ALGORITHM

Feedforward neural networks with additional layers between the source nodes and the output nodes are called multilayer feedforward networks. The layers which are neither input nor output layers are referred to as hidden layers, as they have no direct contact with the environment (see Fig. 3.4). A network with m source nodes, h hidden nodes and q nodes in the output layer is called a m-h-q-network. Training a multilayer feedforward network presents more of a challenge than training a single-layer network, as there is no desired output signal for the neurons in the hidden layers. Because of that, the backpropagation algorithm as introduced by *John Werbos* [43] is used to train such networks.

We describe the implementation of the backpropagation algorithm in an on-line manner. If neuron k is an output node, the local gradient (3.9) can be simply calculated, as an explicit desired response in the form of $d_k^{(i)}$ exists. However, if neuron l is a hidden node, there is no explicit desired response. Yet, the local gradient for hidden neuron l can still be calculated by determining the portion of the error signal, for which neuron l is responsible. In this case, using (2.2), the local gradient is given by [45],

$$\delta_{lj}^{(i)} = -\frac{\partial \mathcal{E}^{(i)}}{\partial y_l^{(i)}} \frac{\partial y_l^{(i)}}{\partial v_l^{(i)}} = -\frac{\partial \mathcal{E}^{(i)}}{\partial y_l^{(i)}} \varphi'\left(v_l^{(i)}\right). \tag{3.11}$$

As the total instantaneous error measure of the network can be calculated according to (3.3), the local gradient $\delta_{lj}^{(i)}$ can be determined using the chain rule and (2.2) as

$$\frac{\partial \mathcal{E}^{(i)}}{\partial y_l^{(i)}} = \sum_{k=1}^{q} \frac{\partial \mathcal{E}^{(i)}}{\partial e_k^{(i)}} \frac{\partial e_k^{(i)}}{\partial v_k^{(i)}} \frac{\partial v_k^{(i)}}{\partial y_l^{(i)}} = -\sum_{k=1}^{q} e_k^{(i)} \varphi'\left(v_k^{(i)}\right) w_{kl}. \tag{3.12}$$

In consequence, the local gradient for a hidden neuron is given by

$$\delta_{lj}^{(i)} = \varphi'\left(v_l^{(i)}\right) \sum_{k=1}^{q} \delta_{kj}^{(i)} w_{kl}, \tag{3.13}$$

enabling the training of neurons in the hidden layer according to (3.10).

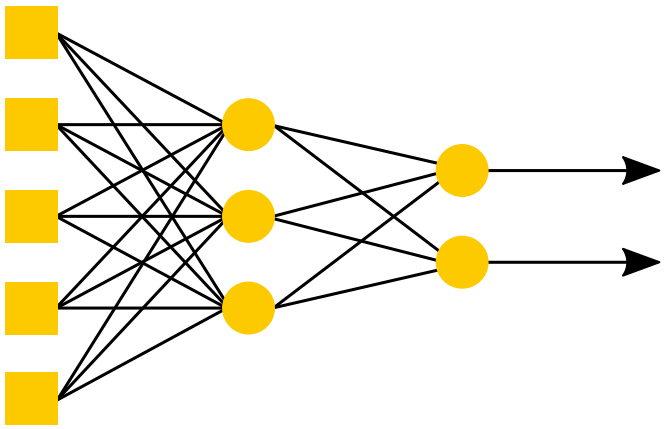

Figure 3.4: Exemplary block diagram of a 5-3-2-multilayer feedforward network. The squares symbolize source nodes where no calculations are performed, while circles represent artificial neurons. The middle layer is called a hidden layer, as the neurons have no direct contact with the environment. Graphic adapted from [45].

3.3 BENCHMARK PROBLEMS

The network architecture as well as the explicit form of the NAF of a multilayer feedforward neural network are typically adapted to the purpose the network is supposed to fulfil. To assess the performance of a neural network, benchmark problems are needed, which can be presented to a network and used to compare its capability to other implementations.

The simplest example of such a benchmark test used in the investigation of neural networks is the Exclusive OR (XOR) problem [75]. The XOR problem represents a logical operation consisting of two binary inputs and one binary output. However, this problem is not solvable using just a single neuron [42]. Therefore, a multilayer architecture is needed to solve the problem, making it a suitable benchmark test to investigate the abilities of the backpropagation algorithm.

A more advanced problem with real-life applications is the recognition of handwritten digits based on the *Modified National Institute of Standards and Technology* (MNIST) dataset [60]. In contrast to the XOR problem, the classification of the MNIST dataset presents a much bigger challenge as the input vectors presented to the network, the variety of input samples and the needed network complexity are much larger. Yet, the recognition of handwritten digits possesses real-life applications and is a good test whether the network can be applied in such settings.

3.3.1 Exclusive OR-Problem

The XOR problem is a logical operation which is true, if and only if its arguments are different. Expressing this using binary inputs x_1 and x_2, the truth table of the XOR problem can be seen in Table 3.1. Accordingly, the XOR operation yields an output $y = 0$ for the input combinations $(0,0)$ and $(1,1)$ and output $y = 1$ for $(1,0)$ and $(0,1)$ [75].

This problem is not linearly separable (see Fig. 3.5). Therefore, a single neuron is not sufficient to solve this problem [42]. Yet, a 2-2-1 feedforward network is able to produce decision boundaries, separating all outputs 0 and 1, thus solving the XOR problem [76]. Assuming a sigmoidal NAF φ_s according to (2.5) with $\alpha = 5$ for all ANs in the network architecture shown in Fig. 3.5, a neural network can be trained to solve the XOR problem using the backpropagation algorithm.

In the on-line training procedure, 1,000 randomized training samples are used in each of the 200 epochs. The learning rate is steadily decreased throughout the training procedure. Starting from $\eta = 0.1$ for the first ten epochs, it drops to $\eta = 0.05$ for epochs 10 to 40, to $\eta = 0.02$ for epochs 40 to 100 and to $\eta = 0.01$ for the final 100 epochs. After each epoch, the network performance in solving the XOR problem is evaluated by calculating the averaged squared error $\mathcal{E}$ according to (3.2) for all $m = 4$ possible input-output pairs of the XOR problem. Starting from random weights and biases, $\mathcal{E}$ decreases rapidly after each epoch (see Fig. 3.6).

Table 3.1: Truth table of the XOR problem.

x_1	0	0	1	1
x_2	0	1	0	1
y	0	1	1	0

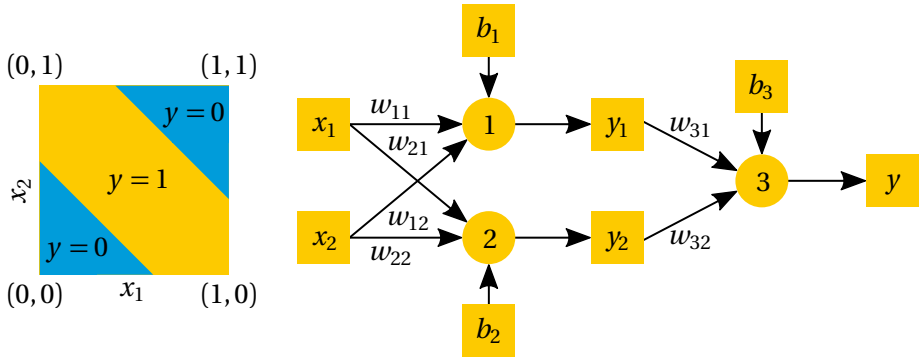

Figure 3.5: (Left) Decision boundaries of XOR problem. Two boundaries are needed to separate the possible input-output-pairs implying that the problem is not linearly separable.
(Right) Network architecture used to solve the XOR problem using artificial neurons with sigmoidal activation function. Terms $b_k = w_{k0}x_0$, where $x_0 = 1$, symbolize neuron-specific biases. The 2-2-1 network is able to produce decision boundaries needed to solve the XOR problem. Graphics adapted from [45].

A test sample is categorized as being identified correctly in the case that the network output is within a tolerance of ±0.05 of the desired value 0 or 1. At the end of the 200$^{\text{th}}$ epoch, the network is able to identify every single test sample correctly.

The ability to solve the XOR problem starting from random initial weights is a proof-of-principle for the backpropagation algorithm. Yet, this algorithm is not limited to such simple problems. It can also be applied to more complicated tasks with higher dimensional input and output.

3.3.2 Recognition of Handwritten Digits - MNIST Dataset

One possible application for ANNs is the automated recognition of images. A popular benchmark test in this area is the recognition of handwritten digits based on the MNIST dataset [60]. This set consists of 60,000 training and 10,000 test samples. Each sample is a scan of a handwritten digit from 0 to 9, which were produced by employees of the American Census Bureau as well as american high school students. The images are saved on a 28×28-pixel grid with

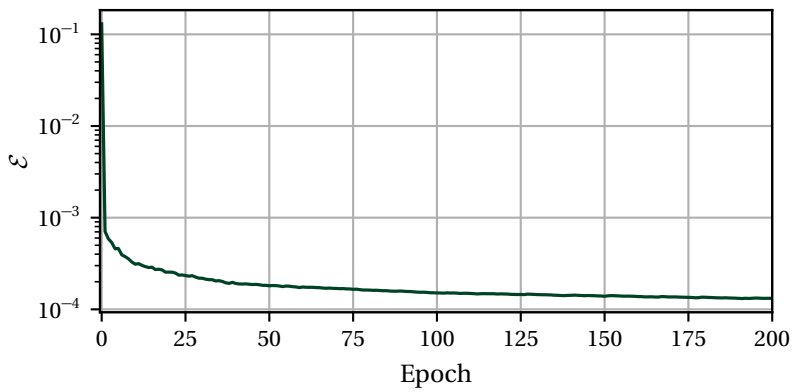

Figure 3.6: Averaged squared error $\mathcal{E}$ versus epochs for solving the XOR problem using artificial neurons with sigmoidal activation function and $\alpha = 5$ in the architecture shown in Fig. 3.5.

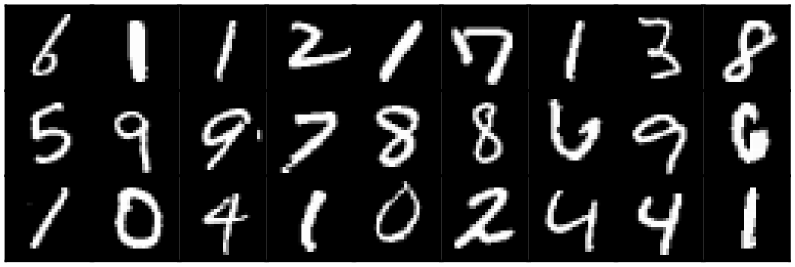

Figure 3.7: Examples of hand-written digits from the MNIST dataset [60].

the center of mass of the pixels positioned at the middle of the grid. Each pixel has a grayscale value ranging from 0 to 255, which is scaled to the interval $[0,1]$ for implementation. Deconvolving these images, the input vectors for an ANN are 784-dimensional [77]. An excerpt of these samples can be seen in Fig. 3.7.

A large variety of neural networks and other machine learning techniques, ranging from linear classifiers, support vector machines and fully-connected feedforward neural networks to convolutional neural networks, has been used to solve this problem [78]. For the purpose of this thesis, we constrain further investigations to fully-connected feedforward networks, as the newly developed implementation schemes for ANNs are going to have this architecture.

In line with [78], a 784-300-10 neural network is subjected to on-line training on the MNIST dataset using the backpropagation algorithm. There, each of the ten output neurons corresponds to a digit from 0 to 9. If, for example, the fifth output neuron produces an output of 1, while all other neurons yield 0, the network has identified a sample as the digit 4. The training is performed over 100 epochs, where in each epoch, 6,000 random training samples from the MNIST dataset are presented to the network. The learning rate is fixed at $\eta = 0.05$. Every neuron in the network possesses a sigmoidal nonlinear activation function of the form (2.5) with $\alpha = 1$.

After each epoch, the performance of the network is evaluated by calculating the averaged squared error $\mathcal{E}$ according to (3.2). As the number of available test input-output pairs is much larger than in the case of the XOR problem, $m = 1{,}000$ test samples are randomly chosen after

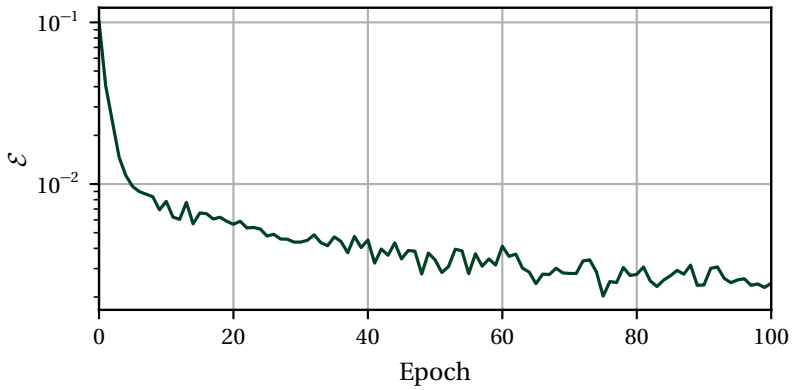

Figure 3.8: Averaged squared error $\mathcal{E}$ versus epochs for the recognition of handwritten digits from the MNIST dataset using a 784-300-10 ANN. $\mathcal{E}$ is reduced significantly throughout the course of the training yielding $\mathcal{E} = 2.4 \cdot 10^{-3}$ at the end of the procedure.

each epoch the determine $\mathcal{E}$. As can be seen in Fig. 3.8, the initial averaged squared error for random weights is quite large. Throughout the course of the training, $\mathcal{E}$ is steadily decreased, yielding $\mathcal{E} = 2.4 \cdot 10^{-3}$ at the end of the procedure.

The network is defined to identify a sample with a digit, if the corresponding output neuron yields a value of 0.8 or higher. If no neuron reaches this threshold, the sample is categorized as being not identified by the network. At the end of the 100^{th} epoch, the network identifies 937 out of the 1,000 test samples. A comparison between the network and the desired responses can be seen in Table 3.2. Apart from the samples, which were not identifiable for the network, there are very few samples, which were identified wrongly. In total, the network shows an average error rate of 7.13%, which is in the range of literature values [78] despite a comparatively short training period.

The success of the network in identifying handwritten digits shows, that quite simple network architectures are able to fulfil such tasks. Therefore, the fully-connected architectures used throughout this thesis to realize and test new implementations of artificial neurons and neural networks have the potential to be applied to real-world problems.

Table 3.2: Sample digit versus digits identified by a 784-300-10 network after 100 training epochs in percent. A test sample is assigned to a digit, if the corresponding neuron produces an output larger than 0.8. If no neuron reaches this threshold, the sample is not identified by the network and is categorized under "-". In total, the network has an error rate of 7.13%

	0	1	2	3	4	5	6	7	8	9	-
0	93.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.7
1	0.0	96.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.4
2	0.0	0.0	95.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.3
3	0.0	0.0	0.0	92.7	0.0	0.0	0.0	0.9	0.0	0.0	6.4
4	0.0	0.0	0.0	0.0	90.0	0.0	0.0	0.0	0.0	2.2	7.8
5	0.0	0.0	0.0	0.0	0.0	97.0	0.0	0.0	0.0	0.0	3.0
6	1.3	0.0	0.0	0.0	0.0	0.0	95.0	0.0	0.0	0.0	3.7
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	91.2	0.0	0.8	8.0
8	0.0	0.0	0.0	0.0	0.0	1.2	0.0	0.0	90.4	0.0	8.4
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.8	0.0	86.8	10.5

4

PHYSICAL IMPLEMENTATIONS OF ARTIFICIAL NEURAL NETWORKS

The concepts of architecture and learning procedures in ANNs require hardware to be implemented on. Traditionally, the platform of choice were electronics-based computers as they offered programmability, mass-fabrication, robust memory solutions and high precision. Especially since 2012, when *Krizhevsky et al.* efficiently used graphics-processing units (GPUs) to implement backpropagation training [66], enormous strides were made in this area. To this day, implementations of ANNs on GPUs appreciate huge popularity due to their outstanding parallelization capabilities [79, 80].

However, growing demands on ANNs in terms of dataset sizes, network complexity and scalability start to put limitations onto GPU-based implementations. Especially the energy consumption of ever-growing processor arrays is a key aspect in this discussion [81]. Therefore, there is an active search for special-purpose hardware to implement ANNs on.

In this chapter, we review established realizations of ANNs based on physical systems. In Section 4.1 we focus on optical implementations, typically referred to as optical neural networks. Additionally we discuss further implementations like wave physics approaches, 2D materials and mechanical systems in Section 4.2.

4.1 OPTICAL NEURAL NETWORKS

One possible approach to circumvent the shortcomings of electronic implementations of ANNs is the usage of optical systems [18, 82]. The upside of such realizations is the ability of optical system to perform parallel vector-matrix multiplications at very high speed [83, 84] while showing small energy consumption [85]. The approaches in optical neural networks are manifold. In this overview, we will focus on all-optical realizations and photonic circuit implementations.

4.1.1 All-Optical Realizations

As early as 1985, free-space optical systems were used to realize Hopfield neural networks [86]. Over the years, many different techniques such as microlens arrays [87, 88], holography [89], fiber interconnections [90] and diffractive networks [91] have been used to implement neural networks in free-space optics.

While optical implementations of neural networks excel at the linear parts of the neuron activity, implementing the crucial nonlinearity of an AN is a big challenge in optical neural networks. Nowadays, the approaches to tackle this problem include using interactions of

free-space optical setups with thermal atoms [25, 92], Kerr-nonlinearities [93] and saturated absorption in semiconductor mirrors [94].

4.1.2 Photonic Circuits

Photonic systems already occupy an important role in telecommunication. With photonic waveguides in a size range of around 500 nm and information bandwidth densities in the range of terabits per second available on fabricated photonic chips, they are also very promising platforms for the implementation of ANNs. When considering large processor sizes and low-precision operations, photonic linear computing systems outperform electronic implementations in terms of energy requirements, speed and computational density [95].

Consequently, a number of approaches have been considered to implement ANNs based on photonic systems [82]. These include, for example, the usage of Mach-Zehnder interferometers on silicon photonic chips [96], micro-ring weight banks to implement recurrent silicon photonic neural networks [97] and the exploitation of division multiplexing techniques on photonic circuits and phase-change materials [98].

4.2 FURTHER IMPLEMENTATIONS

In addition to optical approaches, other routes have been taken to provide hardware platforms for ANNs [99, 100]. For example, in 2022, *Wright et al.* introduced a general approach to implement backpropagation training in controllable physical systems [101]. They directly train the hardware's physical transformations to perform the desired computations. They demonstrated the validity of their approach by applying it to optical, as well as mechanical and electronic systems. In a somewhat similar approach, *Hughes et al.* identify a mapping between the dynamics of wave physics and the computation in recurrent neural networks [102].

A vast variety of further approaches to physically implement ANNs exist today, ranging from coupled spin-torque nano-oscillators [103] to the usage of 2D materials [104] and many more. In this thesis, we aim to contribute to this growing field via the introduction of two approaches, the thermal cloud neural network (Part II) and the four-wave mixing neural network (Part III). Additionally, we investigate the emission properties of quantum dot superluminescent diodes (Part IV), which possess beneficial characteristics to be implemented successfully in ANNs.

Part II

THERMAL CLOUD NEURAL NETWORK

5

STOCHASTIC SIMULATION OF THERMAL ATOMS

In this part, we present an implementation of an AN and, subsequently, of an ANN based on the motion of thermal atoms in an optically-shaped potential landscape, which in the following we refer to as the thermal cloud neural network. As we aim to realize ANs using quantum optical methods and systems of ultracold atoms, a sensible first approach to this problem is to investigate possibilities to implement an algorithm based on the classical motion of thermal atoms in optical potentials. Starting from describing the motion of thermal atoms in Chapter 5, we develop an implementation of the thermal cloud neuron in Chapter 6. Subsequently, this neuron is embedded in a network structure to solve benchmark problems from the realm of ANNs in Chapter 7.

THERMAL ATOMS Consider a cloud of neutral atoms in thermal equilibrium at a temperature T. In real gases, such atoms interact mutually via some interaction potential, which typically is isotropic and short-ranged for neutral atoms in their electronic ground state. For gases at sufficiently low atom densities, i.e. so-called diluted gases, these interactions can be considered as pairwise [105]. Under these conditions the motion of such thermal atoms in potentials, either being applied externally or as an effective potential due to the interactions, can be described classically via the concept of Brownian motion [106, 107].

BROWNIAN MOTION In 1827, while observing small pollen grains suspended in water, *Robert Brown* discovered a random movement of the grains in the absence of external forces [108]. An explanation of Brownian motion was independently developed in the early twentieth century by *Albert Einstein* [109] and *Marian von Smoluchowski* [110]. They recognized, that the motion of the pollen grains was caused by collisions between the grains and the water molecules. However, the motion of the grains turned out to be of such high complexity, that a probabilistic description in term of statistically independent impacts was necessary.

In this chapter, we first discuss the motion of Brownian particles, described by Kramers' equation, in Section 5.1. A simulation method for the Brownian motion of thermal atoms based on the Störmer-Verlet scheme is derived in Section 5.2. As a first application of this simulation scheme, we investigate the equilibration time of a cloud of thermal atoms in a periodic confinement in Section 5.3.

5.1 KRAMERS' EQUATION

5.1.1 Stochastic Equations of Motion for Brownian Particle

The starting point for the description of the dynamics of a thermal atomic gas as Brownian particles is given by the Langevin equation [111, 112]

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -\zeta\frac{\mathrm{d}x}{\mathrm{d}t} + \sqrt{D}X(t). \tag{5.1}$$

This equation determines the (one-dimensional) movement of a particle with mass m and position x due to friction with coefficient ζ and a stochastically fluctuating force $\sqrt{D}X$, thereby imprinting the stochastic nature onto $x(t)$. D is called diffusion coefficient and X is a white noise process, constructed such that its mean value and two-time correlations vanish,

$$\langle X(t)\rangle = 0, \qquad \langle X(t)X(t')\rangle = \delta(t-t'), \tag{5.2}$$

where $\delta(t-t')$ is a delta distribution which is only non-zero for $t-t'=0$. Originally, Langevin called this fluctuating term complementary force. In 1940, *Hendrik A. Kramers* added a term to this equation describing the motion of the particle due to an external force field K [113],

$$m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -\zeta\frac{\mathrm{d}x}{\mathrm{d}t} + \sqrt{D}X(t) + K. \tag{5.3}$$

This force field is supposed to be inferable from a potential

$$K = -\partial_x V(x). \tag{5.4}$$

Furthermore, the fluctuating force is also assumed to be inferable from a potential B as [114]

$$\sqrt{D}X(t) = -\partial_x B, \qquad B = -\sqrt{D}X(t)x. \tag{5.5}$$

Equations of Motion

The Lagrangian equation of motion equivalent to (5.3), including friction according to Rayleigh's dissipation function, is given by [115]

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} + \frac{\partial D_\mathrm{R}}{\partial \dot{x}} = 0, \qquad L = \frac{m}{2}\dot{x}^2 - (V(x)+B), \tag{5.6}$$

with the dissipation function

$$D_\mathrm{R} = \frac{k\dot{x}^2}{2}, \qquad k = \frac{\zeta}{m}. \tag{5.7}$$

As D_R only depends on the particle velocity $\dot{x}$, (5.6) is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \tilde{L}}{\partial \dot{x}} - \frac{\partial \tilde{L}}{\partial x} = 0, \tag{5.8}$$

with the modified Lagrangian $\tilde{L} = Le^{kt}$.

In order to determine the Hamiltonian of the system via a Legendre transform [115]

$$H(x,\Pi,t)=\Pi\dot{x}(x,\Pi,t)-\mathcal{L}(x,\Pi,t), \tag{5.9}$$

where $\mathcal{L}(x,\Pi,t)$ is the generalized Lagrangian, the generalized momentum Π has to be determined as

$$\Pi=\frac{\partial\tilde{L}}{\partial\dot{x}}=m\dot{x}e^{kt}. \tag{5.10}$$

Therefore, the Hamiltonian of the system is given by

$$H(x,\Pi,t)=\frac{\Pi^2}{2m}e^{-kt}+(V(x)+B)e^{kt}. \tag{5.11}$$

The Hamiltonian equations of motions of the particle in terms of position x and generalized momentum Π can be determined as

$$\begin{aligned}\dot{x}&=\frac{\partial H}{\partial\Pi}=\frac{\Pi}{m}e^{-kt},\\ \dot{\Pi}&=-\frac{\partial H}{\partial x}=-\left(V'(x)-\sqrt{D}X(t)\right)e^{kt}.\end{aligned} \tag{5.12}$$

Rewriting these equations in terms of the mechanical momentum

$$p=m\dot{x}=\Pi e^{-kt} \tag{5.13}$$

yields the dynamics of the system in terms of stochastic differential equations as [116]

$$\begin{aligned}\mathrm{d}x&=\frac{p}{m}\mathrm{d}t,\\ \mathrm{d}p&=-\left(V'(x)+kp\right)\mathrm{d}t+\sqrt{D}\mathrm{d}W(t).\end{aligned} \tag{5.14}$$

There, the stochastically fluctuating term $X(t)\mathrm{d}t$ is identified as the Wiener noise increment $\mathrm{d}W(t)$ [116], which satisfies

$$\langle\mathrm{d}W(t)\rangle=0, \qquad \langle\mathrm{d}W(t)\mathrm{d}W(t')\rangle=\mathrm{d}t\,\delta(t-t'). \tag{5.15}$$

Ito Stochastic Differential Equations

The set of equations (5.14) are called Ito stochastic differential equations. An arbitrary stochastic quantity $s(t)$ is said to obey an Ito stochastic differential equation

$$\mathrm{d}s(t)=a(s(t),t)\mathrm{d}t+b(s(t),t)\mathrm{d}W(t), \tag{5.16}$$

if for all t and t_0, the solution $s(t)$ is given by

$$s(t)=s(t_0)+\int_{t_0}^{t}\mathrm{d}t'\,a(s(t'),t')+\int_{t_0}^{t}\mathrm{d}W(t')\,b(s(t'),t'). \tag{5.17}$$

In this equation, the last term is identified as an Ito stochastic integral, which can be calculated as [116]

$$\int_{t_0}^{t}\mathrm{d}W(t')\,b(s(t'),t')=\underset{n\to\infty}{\text{ms-lim}}\left\{\sum_{i=1}^{n}b(s(t_{i-1}),t_{i-1})[W(t_i)-W(t_{i-1})]\right\}. \tag{5.18}$$

There, the mean square limit ms-lim denotes, that a sequence of random variables S_n converges to S, if

$$\lim_{n\to\infty}\left\langle (S_n - S)^2 \right\rangle = 0. \tag{5.19}$$

In this instance, one can write [116]

$$\underset{n\to\infty}{\text{ms-lim}} S_n = S. \tag{5.20}$$

5.1.2 Kramers' Fokker-Planck Equation

The described dynamics (5.14) can be written as a multi-variable system in terms of the phase-space coordinate

$$\boldsymbol{z} = \begin{pmatrix} x \\ p \end{pmatrix}, \tag{5.21}$$

yielding

$$\mathrm{d}\boldsymbol{z} = \boldsymbol{A}\mathrm{d}t + b\,\mathrm{d}\boldsymbol{W}(t) = \begin{pmatrix} p/m \\ -(V'(x)+kp) \end{pmatrix}\mathrm{d}t + \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{D} \end{pmatrix}\mathrm{d}\boldsymbol{W}(t), \tag{5.22}$$

where $\mathrm{d}\boldsymbol{W}(t)) = (\mathrm{d}W_x(t), \mathrm{d}W_p(t))^{\mathrm{T}}$, $\boldsymbol{A}$ is called the drift vector and $B = b\,b^{\mathrm{T}}$ is the diffusion matrix . Such a stochastic differential equation for a multi-variable system is connected to a Fokker-Planck equation, describing the dynamics of the probability distribution function $f(x,p,t)$, via the relation [116]

$$\partial_t f(x,p,t) = -\sum_i \partial_i \left[A_i f(x,p,t)\right] + \frac{1}{2}\sum_{i,j}\partial_i\partial_j\left[B_{ij} f(x,p,t)\right], \tag{5.23}$$

with $i,j \in [x,p]$. Applying this to (5.22) yields Kramers' equation

$$\partial_t f(x,p,t) = -\partial_x\left[\frac{p}{m} f(x,p,t)\right] + \partial_p\left[\left(V'(x)+kp\right) f(x,p,t)\right] + \frac{1}{2} D\,\partial_p^2 f(x,p,t). \tag{5.24}$$

There, the diffusion matrix B can be identified as

$$B = \begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix}. \tag{5.25}$$

5.1.3 Maxwell-Boltzmann Statistics for Brownian Particles

As Kramers' equation describes the motion of Brownian particles, and in extension thermal atoms, its stationary solution $f_{\mathrm{s}}(x,p)$ has to satisfy the Maxwell-Boltzmann statistics [117]

$$f_{\mathrm{s}}(x,p) = \mathcal{N}\exp\left(-\frac{H}{k_{\mathrm{B}}T}\right), \qquad H = \frac{p^2}{2m} + V(x), \tag{5.26}$$

where the normalization constant $\mathcal{N}$ can be chosen such that $\int \mathrm{d}x\,\mathrm{d}p\; f_{\mathrm{s}}(x,p) = 1$. To ensure this, the diffusion coefficient D has to be further determined. The stationary solution of (5.24) can be calculated using the detailed balance condition [116]

$$\varepsilon_i A_i(\varepsilon\boldsymbol{z}) f_{\mathrm{s}}(\boldsymbol{z}) = \left[-A_i(\boldsymbol{z}) + \sum_j \partial_j B_{ij}(\boldsymbol{z})\right] f_{\mathrm{s}}(\boldsymbol{z}), \tag{5.27}$$

with the detailed balance transformation

$$\boldsymbol{\varepsilon}\begin{pmatrix} x \\ p \end{pmatrix} = \begin{pmatrix} x \\ -p \end{pmatrix}, \qquad \boldsymbol{\varepsilon} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{5.28}$$

Evaluating the detailed balance condition yields

$$\partial_p f_\mathrm{s}(x,p) = -2\frac{\zeta p}{mD} f_\mathrm{s}(x,p). \tag{5.29}$$

One possible solution for this condition is given by

$$f_\mathrm{s}(x,p) = \exp\left(-\frac{\zeta}{mD}p^2\right) g(x). \tag{5.30}$$

As this has to be a solution of (5.24) as well, the term $g(x)$ has to satisfy

$$\partial_x g(x) = -\frac{2\zeta V'(x)}{D} g(x). \tag{5.31}$$

The full stationary solution is therefore given by

$$f_\mathrm{s}(x,p) = \mathcal{N}\exp\left[-\frac{2\zeta}{D}\left(\frac{p^2}{2m} + V(x)\right)\right]. \tag{5.32}$$

Comparing this to the Maxwell-Boltzmann statistics (5.26), setting

$$D = 2\zeta k_\mathrm{B} T \tag{5.33}$$

ensures thermal statistics for the atoms. This is known as the Einstein relation [109].

5.2 STÖRMER-VERLET-LIKE SIMULATION SCHEME

To simulate the Brownian motion of thermal atoms, a simulation scheme for the stochastic differential equations (5.14) has to be derived. For this, we develop a Störmer-Verlet-like simulation scheme.

Störmer-Verlet Scheme

A general solution to Hamiltonian equations of motion for the generalized phase-space coordinates

$$\boldsymbol{\chi} = \begin{pmatrix} x \\ \Pi \end{pmatrix} \tag{5.34}$$

is given by [118]

$$\boldsymbol{\chi}(t) = \exp\left[t\sum_{j=1}^{2}\{\chi_{0,j}, H(\boldsymbol{\chi}_0)\}\partial_{\chi_{0,j}}\right]\boldsymbol{\chi}_0, \tag{5.35}$$

with the Poisson bracket

$$\{\chi_{0,j}, H(\boldsymbol{\chi}_0)\} = \frac{\partial \chi_{0,j}}{\partial x_0}\frac{\partial H}{\partial \Pi_0} - \frac{\partial H}{\partial x_0}\frac{\partial \chi_{0,j}}{\partial \Pi_0}. \tag{5.36}$$

Introducing the kinetic and potential part of the dynamics as

$$T = \{x_0, H(\boldsymbol{\chi}_0)\}\partial_{x_0}, \qquad U = \{\Pi_0, H(\boldsymbol{\chi}_0)\}\partial_{\Pi_0}, \tag{5.37}$$

respectively, the time evolution of $\boldsymbol{\chi}(t)$ can be rewritten as

$$\boldsymbol{\chi}(t) = \exp[t(T+U)]\boldsymbol{\chi}_0. \tag{5.38}$$

Now, there exists a set of numbers $(c_1, \ldots, c_n)$ and $(d_1, \ldots, d_n)$ such that

$$\exp[t(T+U)] = \prod_{i=1}^{n} \exp[c_i t T]\exp[d_i t U] + \mathcal{O}(t^{n+1}). \tag{5.39}$$

The integer n is called the order of the integrator [119]. For $n = 2$, this procedure yields

$$\boldsymbol{\chi}(t) = \exp[c_1 t T]\exp[c_2 t T]\exp[d_1 t U]\exp[d_2 t U]. \tag{5.40}$$

The coefficients c_1, c_2, d_1 and d_2 can be calculated by expanding the exact solution and the approximation (5.40) into Taylor series up to an order of $n = 2$ and subsequently verifying (5.39). This results in the set of equations

$$1 = c_1 + c_2, \quad 1 = d_1 + d_2, \quad 1 = (c_1 + c_2)^2, \quad 1 = (d_1 + d_2)^2, \quad 1 = (c_1 + c_2)(d_1 + d_2). \tag{5.41}$$

One exemplary solution to these equations is given by the set of coefficients

$$c_1 = \frac{1}{2}, \qquad c_2 = \frac{1}{2}, \qquad d_1 = 1, \qquad d_2 = 0. \tag{5.42}$$

Applying this set to (5.40) yields the Störmer-Verlet scheme [120]

$$\boldsymbol{\chi}(t) = S_2(t)\boldsymbol{\chi}_0 = \exp\left[\frac{t}{2}T\right]\exp[tU]\exp\left[\frac{t}{2}T\right]\boldsymbol{\chi}_0. \tag{5.43}$$

Störmer-Verlet-Like Scheme for Brownian Motion

This method can be applied to the Hamiltonian of Brownian motion (5.11) by assuming instantaneous values for the stochastically fluctuating force $X(t)$. Hence, evaluating the Poisson brackets in (5.37), the kinetic and potential parts of the Störmer-Verlet scheme are given by

$$T = \frac{\Pi_0}{m} e^{-kt}\partial_{x_0}, \qquad U = -(V'(x_0) - \sqrt{D}X(t))e^{kt}\partial_{\Pi_0}. \tag{5.44}$$

Thus, using the explicit forms of T and U, while taking into account only first-order terms in the expansion of $S_2(t)$, the time derivative of the phase-space coordinate $\dot{\boldsymbol{z}}(t)$ can be calculated as

$$\dot{\boldsymbol{z}}(t) = \begin{pmatrix} p_0/2m \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -(V'(x_0) + k p_0) + \sqrt{D}X(t) \end{pmatrix} + \begin{pmatrix} p_0/2m \\ 0 \end{pmatrix}. \tag{5.45}$$

Discretizing the set of differential equations for time step Δt yields a Störmer-Verlet-like simulation scheme for Brownian motion of thermal atoms

$$\begin{aligned} x_{i+1/2} &= x_i + \frac{\Delta t}{2}\frac{p_i}{m}, \\ p_{i+1} &= p_i - \Delta t\left[V'(x_{i+1/2}) + \frac{\zeta}{m}p_i\right] + \sqrt{2\zeta k_\mathrm{B} T}\Delta W(t), \\ x_{i+1} &= x_{i+1/2} + \frac{\Delta t}{2}\frac{p_{i+1}}{m}. \end{aligned} \tag{5.46}$$

There, $\Delta W(t)$ is a Gaussian random process with $\langle \Delta W \rangle = 0$ and $\langle \Delta W^2 \rangle = \Delta t$. The simulation scheme for the stochastic differential equations can be interpreted either as an Euler- or a Milstein-scheme, as both are equivalent in this case. For more information on time-discrete approximation of stochastic differential equations see Appendix A.

5.3 EQUILIBRATION TIME IN A PERIODIC CONFINEMENT

5.3.1 Thermal Atoms in Periodic Confinement

As an application of the stochastic simulation of a thermal gas, the characteristic timescales of the dynamics of such a system are investigated. The system under investigation is a one-dimensional thermal gas in a confinement of length L with periodic boundary conditions, that is $(x,p) = (x+L,p)$. Inside the confinement no potential gradient is present, $V'(x) = 0$. Therefore, the stochastic differential equations describing the dynamics of this system are given by

$$\begin{pmatrix} \mathrm{d}x \\ \mathrm{d}p \end{pmatrix} = \begin{pmatrix} 0 & 1/m \\ 0 & \zeta/m \end{pmatrix} \begin{pmatrix} x \\ p \end{pmatrix} \mathrm{d}t + \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{D} \end{pmatrix} \begin{pmatrix} \mathrm{d}W_x(t) \\ \mathrm{d}W_p(t) \end{pmatrix}. \tag{5.47}$$

We define this system to be in equilibrium, if the positions of the atoms are uniformly distributed along the confinement. To verify this condition, the variance

$$\mathrm{Var}(x) = \left\langle (x - \langle x \rangle)^2 \right\rangle \tag{5.48}$$

of the thermal gas, where $\langle \cdot \rangle$ denotes the average over all particles, is compared to the value for a uniform position distribution described by

$$f_\mathrm{u}(x) = \begin{cases} 1/L, & \text{for } x \in [0,L], \\ 0, & \text{otherwise.} \end{cases} \tag{5.49}$$

In the uniform case, the moments of the distribution can be calculated analytically as

$$\left\langle x^n \right\rangle = \int_{-\infty}^{\infty} \mathrm{d}x \; x^n p_\mathrm{u}(x). \tag{5.50}$$

Accordingly, the variance for uniformly distributed particles in the confinement is given by

$$\mathrm{Var}_\mathrm{u}(x) = \frac{1}{L} \int_0^L \mathrm{d}x \left(x - \frac{L}{2} \right)^2 = \frac{L^2}{12}. \tag{5.51}$$

5.3.2 Simulation Results for Rubidium-87

The investigation of the timescales is performed for Rubidium-87 (^{87}Rb) atoms with an atomic mass of $m = 1.443 \cdot 10^{-25}\,\mathrm{kg}$ [121]. The atoms are assumed to be cooled to their Doppler temperature $T_\mathrm{D} = 145.57\,\mu\mathrm{K}$ [122] and a total of $N = 10{,}000$ atoms are trapped in a confinement of length $L = 10\,\mu\mathrm{m}$ [123]. For an estimation of the diffusion coefficient $D = 2\zeta k_\mathrm{B} T$, the friction coefficient $\zeta = 6\pi\eta a$ [111, 112], with η the viscosity and a the particle diameter, has to be determined. For a dilute ideal gas, the shear viscosity can be calculated as [124]

$$\eta = \frac{2}{3\pi^{3/2}} \frac{\sqrt{m k_\mathrm{B} T}}{a}. \tag{5.52}$$

With a particle diameter for ^{87}Rb of $a \approx 600\,\mathrm{pm}$ [125], the friction coefficient can be estimated to $\zeta = 5.65 \cdot 10^{-17}\,\mathrm{kg/s}$. The atomic parameters of ^{87}Rb are summarized in Appendix B.

The thermal de Broglie wavelength of these particles can be determined as [126]

$$\lambda_{\mathrm{dB}} = \sqrt{\frac{2\pi\hbar^2}{m k_{\mathrm{B}} T}} \approx 7.66 \cdot 10^{-12}\,\mathrm{m}. \tag{5.53}$$

Compared to the mean free path of the particles $l = L/N = 10^{-10}\,\mathrm{m}$, the extent of λ_{dB} is rather small. Therefore, the system is expected to behave like a cloud of classical particles. Initially, the particle positions are drawn randomly from a Gaussian distribution with mean $\langle x_0 \rangle = L/2$ and variance $\mathrm{Var}(x_0) = L^2/64$. To be conform with the Maxwell-Boltzmann distribution for thermal atoms (5.26), the initial momentum distribution is also Gaussian with a vanishing mean $\langle p_0 \rangle = 0$ and variance $\mathrm{Var}(p_0) = m k_{\mathrm{B}} T$.

The particle position distributions at initialization and after $t = 400\,\mathrm{ms}$, as well as the time evolution of the variance $\mathrm{Var}(x)$ can be seen in Fig. 5.1. From the initial Gaussian distribution, the variance of the position distribution rises and corresponds to the one of the uniform distribution within one percent after $t_{\mathrm{e}} = 344\,\mathrm{ms}$.

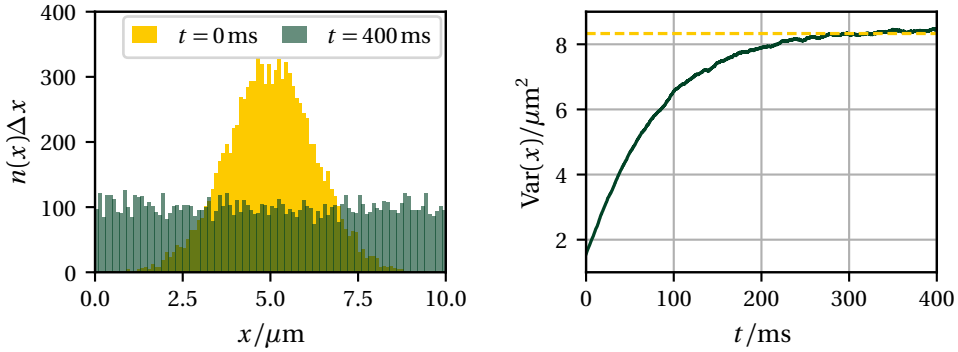

Figure 5.1: (a) Density of particle positions $n(x)$ for position discretization $\Delta x = 0.1\,\mu\mathrm{m}$ at $t = 0\,\mathrm{ms}$ and $t = 400\,\mathrm{ms}$. The initial Gaussian distribution with $\langle x_0 \rangle = 5\mu\mathrm{m}$ and $\mathrm{Var}(x_0) = 1.56\,\mu\mathrm{m}^2$ evolves into a uniform distribution along the confinement. (b) Variance $\mathrm{Var}(x)$ versus time t. The variance of the thermal gas (green, solid) corresponds to the one of the uniform distribution $\mathrm{Var}_{\mathrm{u}}(x) = 8.33\,\mu\mathrm{m}^2$ (yellow, dashed) within one percent after $t_{\mathrm{e}} = 344\,\mathrm{ms}$.

5.3.3 Comparison with Free Particles

To investigate the influence of the periodic boundary on the equilibration time of the system, the evolution of the variance is compared to the case of free particles, the dynamics of which are described by (5.47). In this case, a general solution for the particle positions is given by

$$x(t) = x_0 + \frac{1}{\zeta}\left(e^{-\zeta t/m} - 1\right) p_0 + \int \mathrm{d}W_p(t') \frac{\sqrt{D}}{\zeta}\left(e^{-\zeta(t-t')/m} - 1\right). \tag{5.54}$$

The statistical properties of this solution can be investigated using Ito calculus. There, two main results are of interest here. Firstly, terms of the form

$$\int \mathrm{d}W(t)\, G(t') = 0 \tag{5.55}$$

vanish for nonanticipating functions $G(t)$ [127]. A function is nonanticipating if its dynamics are independent of the behaviour of the Wiener process $\mathrm{d}W(t)$ in the future of t. This condition is fulfilled for physical processes due to the principle of causality [116]. Secondly, doubly-stochastic integrals can be evaluated for nonanticipating functions $G(t)$ via the relation [127]

$$\int_{t_0}^{t} \big(\mathrm{d}W(t')\big)^2\, G(t') = \int_{t_0}^{t} \mathrm{d}t'\, G(t'). \tag{5.56}$$

Therefore, the variance for free particles can be calculated as

$$\begin{aligned}\mathrm{Var_f}(x) =& \mathrm{Var}(x_0) + \frac{1}{\zeta^2}\big(e^{-\tau}-1\big)^2\,\mathrm{Var}(p_0) + \frac{2}{\zeta}\big(e^{-\tau}-1\big)\,\mathrm{Cov}(x_0,p_0)\\ &+ \frac{D\,m}{\zeta^3}\left[\frac{1}{2}\big(1-e^{-2\tau}\big) - 2\big(1-e^{-\tau}\big) + \tau\right],\end{aligned} \tag{5.57}$$

with the covariance $\mathrm{Cov}(x_0,p_0) = \langle (x_0 - \langle x_0\rangle)(p_0 - \langle p_0\rangle)\rangle$ and $\tau = \zeta\, t/m$. Hence, for large τ and the given parameters for ^{87}Rb, the free particles reach the variance for a uniform distribution in the confinement of length L after $t_\mathrm{f} \approx 95\,\mathrm{ms}$. Consequently, the variance of a set of free particle rises faster than for particles confined in periodic boundaries. This implies, that the presence of a periodic confinement indeed influences the dynamics of a thermal atomic gas.

6

THERMAL CLOUD NEURON

On the basis of the movement of thermal atoms in an optically produced potential landscape, we introduce the implementation of an AN, the thermal cloud neuron. This realization relies on algorithmic changes of optical dipole potentials and the subsequent equilibration of the thermal atoms in these new environments. By interpreting optical potential depths as input and remaining particle numbers as output of an AN, the NAF of the thermal cloud neuron can be determined. Through a careful choice of shapes and depths of the optical dipole potentials, we are able to implement an AN with a sigmoidal-like NAF.

After introducing the concept of optical dipole potentials with a focus on basic atom-light-interaction and the realization of box potentials in Section 6.1, the algorithmic procedure to prepare the inputs of the thermal cloud neuron is presented in Section 6.2. Subsequently, the NAF of the thermal cloud neuron is investigated in Section 6.3.

6.1 OPTICAL DIPOLE POTENTIALS

6.1.1 Atom-Light Interaction

The situation of interest is a cold dilute gas of neutral atoms moving much slower than speed of light and interacting with one or several lasers in the visible or near-infrared range. The intensity of the laser is chosen to be sufficiently low such that no ionization or higher-order effects occur. Therefore, non-relativistic quantum mechanics can be used to describe the atom and its center-of-mass motion.

Dipole Interaction between Two-Level Atom and Quantized Electromagnetic Field

Restricting the description to a single atom, the Hamiltonian of the system is [106]

$$\hat{H} = \hat{H}_\mathrm{a} + \hat{H}_\mathrm{r} + \hat{H}_\mathrm{i}, \tag{6.1}$$

where $\hat{H}_\mathrm{a}$, $\hat{H}_\mathrm{r}$ and $\hat{H}_\mathrm{i}$ are the Hamiltonians of the atom, the radiation field and the interaction between them, respectively. For the atom, only two electronic states, the ground $|g\rangle$ and excited state $|e\rangle$, are assumed to be relevant. Taking into account the external and internal degrees of freedom of the atom, its Hamiltonian can be written as

$$\hat{H}_\mathrm{a} = \frac{\hat{\boldsymbol{p}}^2}{2m} + \hbar\omega_g |g\rangle\langle g| + \hbar\omega_e |e\rangle\langle e|, \tag{6.2}$$

where $\hat{\boldsymbol{p}}$ is the center-of-mass momentum of the atom and $\hbar\omega_i$ is the energy of state $|i\rangle$.

The quantized electromagnetic field is expanded on the complete set of plane wave modes with frequencies ω_j, yielding

$$\hat{H}_\mathrm{r} = \sum_j \hbar\omega_j \hat{a}_j^\dagger \hat{a}_j, \tag{6.3}$$

where $\hat{a}_j$ is the annihilation operator of the j-th mode. In this section, modifications of the incident radiation due to the interaction with the atom are not taken into account.

To describe the interaction between atom and radiation field, it is assumed that the extent of the atom is much smaller than the wavelength of the electromagnetic field, which typically holds for radiation in the visible or near-infrared spectrum. This assumption is called the dipole approximation [128]. Accordingly, the interaction Hamiltonian is given by

$$\hat{H}_\mathrm{i} = -\hat{\boldsymbol{d}} \cdot \hat{\boldsymbol{E}}(\boldsymbol{r}), \tag{6.4}$$

where $\hat{\boldsymbol{d}}$ is the atomic dipole moment and $\hat{\boldsymbol{E}}(\boldsymbol{r})$ is the transverse electromagnetic field operator evaluated at the center-of-mass position $\boldsymbol{r}$ of the atom. For the two-level-atom, the dipole moment operator can be written as

$$\hat{\boldsymbol{d}} = \boldsymbol{d}_{eg} |e\rangle\langle g| + \boldsymbol{d}_{ge} |g\rangle\langle e|, \tag{6.5}$$

with the dipole matrix elements $\boldsymbol{d}_{ij} = \langle i|\hat{\boldsymbol{d}}|j\rangle$.

In the electromagnetic field of a laser, very few modes contain the vast majority of all photons, while all other modes are basically unoccupied. Hence, the field $\hat{\boldsymbol{E}}$ can be approximated as a combination of a classical monochromatic external field

$$\boldsymbol{E}_\mathrm{c}(\boldsymbol{r}, t) = \boldsymbol{\mathcal{E}}(\boldsymbol{r}) e^{-\mathrm{i}\omega t} + \boldsymbol{\mathcal{E}}^*(\boldsymbol{r}) e^{\mathrm{i}\omega t}, \tag{6.6}$$

with complex amplitude $\boldsymbol{\mathcal{E}}$ and frequency ω, and the vacuum field $\hat{\boldsymbol{E}}_\mathrm{v}$.

Light Forces

The classical light field exerts a force on the atom. This mechanical effect of light on an atom is described by the force operator $\hat{\boldsymbol{F}}$ [107]. It can be calculated using Heisenberg's equation of motion for the atomic momentum operator, yielding

$$\hat{\boldsymbol{F}} = \partial_t \hat{\boldsymbol{p}} = \frac{\mathrm{i}}{\hbar}\left[\hat{H}, \hat{\boldsymbol{p}}\right] = \frac{\mathrm{i}}{\hbar}\left[\hat{H}_i, \hat{\boldsymbol{p}}\right]. \tag{6.7}$$

The quantity describing the actual motion of the atom is the expectation value $\boldsymbol{F} = \mathrm{Tr}\{\hat{\boldsymbol{F}}\hat{\rho}\}$, where $\hat{\rho}$ is the density operator of the total system consisting of the atom as well as the radiation field. Partial traces over the radiation field and the atomic degrees of freedom can be performed subsequently $\boldsymbol{F} = \mathrm{Tr}_a\{\mathrm{Tr}_r\{\hat{\boldsymbol{F}}\hat{\rho}\}\} = \mathrm{Tr}_a\{\hat{\boldsymbol{F}}\hat{\varrho}\}$, where $\hat{\varrho}$ is the reduced density operator of the atomic system.

Assuming the atomic wave function is strongly localized around its center of mass and that its extent is small compared to the spatial variation of the electric field, the partial trace over the atomic degrees of freedom can be executed. This yields the expectation value of the force operator in terms of the Rabi frequency $\Omega = \boldsymbol{d}_{eg}\boldsymbol{\mathcal{E}}/\hbar = |\Omega| e^{i\phi}$ [106, 129]

$$\begin{aligned}\boldsymbol{F} \approx& \varrho_{eg}\hbar\nabla\Omega^* + \varrho_{ge}\hbar\nabla\Omega \\ =& 2\hbar|\Omega|^2 \mathrm{Im}\left(\frac{\varrho_{eg}}{\Omega^*}\right)\nabla\phi + \hbar\,\mathrm{Re}\left(\frac{\varrho_{eg}}{\Omega}\right)\nabla|\Omega|^2,\end{aligned} \tag{6.8}$$

where $\varrho_{ij} = \langle i|\hat{\varrho}|j\rangle$ are elements of the atomic density matrix. The first term describes the radiation pressure, which acts along the wave vector $\boldsymbol{k}$ of the light field. The second term is the optical dipole force, which points along the gradient of the Rabi frequency, and therefore along ∇I, where $I = 2\varepsilon_0 c|\boldsymbol{\mathcal{E}}|^2$ is the intensity of the light field.

Atomic Equations of Motion

To further evaluate these expressions, the density matrix element ϱ_{eg} has to be calculated. The dynamics of the atomic density operator are given by the master equation [130]. In the case that the light field is not extremely detuned $|\Delta| = |\omega - \omega_{eg}| \ll \omega_{eg}$ from the atomic transition $\omega_{eg} = \omega_e - \omega_g$ and not exceedingly strong, $\Omega \ll \omega$, rapidly oscillating terms can be neglected. This approach is known as the rotation wave approximation [131]. Therefore, the equations of motion for the elements of the atomic density matrix can be written in a suitable co-rotating frame as [128]

$$\begin{aligned}
\partial_t \varrho_{ee} &= -\Gamma\rho_{ee} + \mathrm{i}\Omega^*\varrho_{ge} - \mathrm{i}\Omega\varrho_{eg},\\
\partial_t \varrho_{gg} &= \Gamma\rho_{gg} - \mathrm{i}\Omega^*\varrho_{ge} + \mathrm{i}\Omega\varrho_{eg},\\
\partial_t \varrho_{eg} &= -\left(\frac{\Gamma}{2} - \mathrm{i}\Delta\right)\varrho_{eg} - \mathrm{i}\Omega^*(\varrho_{ee} - \varrho_{gg}),\\
\partial_t \varrho_{ge} &= -\left(\frac{\Gamma}{2} + \mathrm{i}\Delta\right)\varrho_{ge} + \mathrm{i}\Omega(\varrho_{ee} - \varrho_{gg}).
\end{aligned} \tag{6.9}$$

The term Γ describes the decay of the excited state due to spontaneous emission.

As the internal dynamics of the atom are much faster than those of the electric field, the stationary solution $\partial_t \tilde{\varrho}_{eg} = 0$ of the internal dynamics can be used. This yields

$$\tilde{\varrho}_{eg} = \frac{\mathrm{i}\Gamma - 2\Delta}{4\Omega}\frac{s(\Delta)}{1+s(\Delta)}, \qquad s = \frac{2|\Omega|^2}{\Delta^2 + (\Gamma/2)^2}. \tag{6.10}$$

Strongly Detuned Regime - Optical Dipole Force

In the investigation of optical dipole forces, we consider cases where the light field is strongly detuned from the atomic transition, $|\Delta| \gg \Omega, \Gamma$. In this case, also Doppler corrections of the frequencies due to the motion of the particles can be neglected. The dissipative and dipole forces from (6.8) can be written as

$$\boldsymbol{F}_{\mathrm{dis}} = \frac{\hbar|\Omega|^2\Gamma}{\Delta^2}\nabla\phi, \qquad \boldsymbol{F}_{\mathrm{dip}} = -\frac{\hbar}{\Delta}\nabla|\Omega|^2. \tag{6.11}$$

As $\boldsymbol{F}_{\mathrm{dis}}$ scales with Δ^{-2} and $\boldsymbol{F}_{\mathrm{dip}}$ scales with Δ^{-1}, the dipole force dominates for large detunings. Furthermore, the dipole force is conservative and can therefore be inferred as the negative gradient of the optical dipole potential [132] given by

$$V_{\mathrm{dip}}(\boldsymbol{r}) = \frac{\mu}{\Delta} I(\boldsymbol{r}), \tag{6.12}$$

where $\mu = 3\pi c^2\Gamma/2\omega_{eg}^3$. The sign of the detuning determines the main behaviour of the potential. In the blue detuned case ($\Delta > 0$) the potential is repulsive, while in the red detuned case ($\Delta < 0$) it is attractive.

6.1.2 Box Potentials

The ability to trap atoms in optical dipole potentials according to (6.12) has made these systems a popular platform to study many-body physics [133, 134]. In early experiments in this field, optical dipole traps predominantly possessed harmonic shapes with the opportunity to create low-dimensional traps, double wells and optical lattices [27, 135–137].

In recent years, optical box traps have become increasingly popular [123]. As such box traps typically have flat bottoms and sharp walls, they are able to produce uniform particle distributions inside the trap [138–140]. This homogeneity of the gas, being quantum or classical, has allowed the study of a wide range of effects [141–145].

To produce box potentials, sculpted blue-detuned laser beams are used to generate a repulsive potential, constructing the box walls and confining the atoms in the process (see Fig. 6.1). Initially, such box traps are typically three-dimensional. To produce lower-dimensional boxes, tight confinements along a direction are used to freeze out the atom motion in an additional dimension. Attractive box potentials, produced by red-detuned laser beams, can also be used to confine atoms. However, this method is experimentally more demanding than using repulsive walls [123, 146].

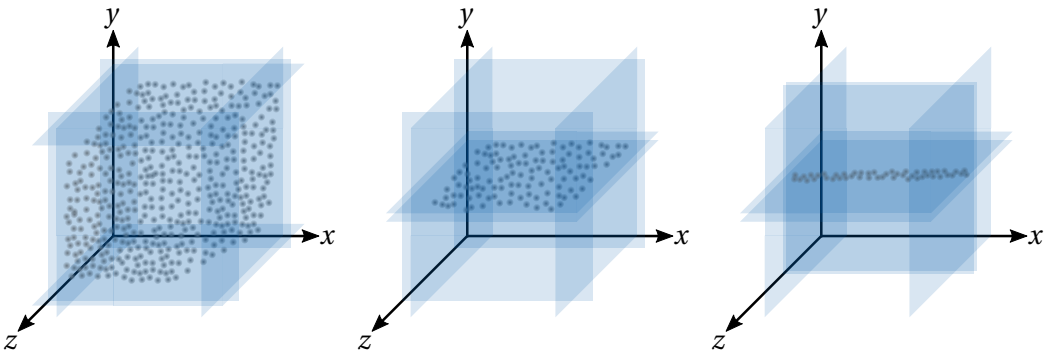

Figure 6.1: Box potentials constructed out of blue-detuned light sheets, confining particles in 3D, 2D and 1D (left to right).

Popular methods to produce specifically sculpted laser beams for box traps are spatial light modulators (SLMs) and digital micromirror devices (DMDs). Liquid-crystal-based SLMs are typically rectangular arrays of up to 10^6 pixel elements with individually controllable indices of refraction [147]. The intensity pattern in the vicinity of the Fourier plane of a SLM is controlled by spatially modulating the phase delay of a single laser beam. This method can be used to create multiple beams needed to realize a box trap from only one laser [138].

On the other hand, DMDs are typically rectangular arrays of around 10^6 mirrors, which can individually be turned 'on' or 'off' by changing their tilt angle [148]. In doing so, the amplitude of a laser beam can directly be spatially modulated and arbitrary intensity patterns can be imaged onto atomic clouds [149]. Therefore, DMDs are very convenient devices to produce arbitrarily shaped box traps, such as squares in 2D and cubes in 3D. These shapes can also be dynamically changed during the experiments due to the high refreshing rates of DMDs [150].

It has to be noted, that state-of-the-art box traps are not perfect. The sharpness of the box walls are limited by the optical wavelength of the used laser beams. This is typically not negligible compared to the box dimensions of interest [123]. Yet, as experimental methods are always evolving and new technologies become available, we assume perfect box potentials in the following discussion of the implementation of the thermal cloud neuron.

6.2 INPUT PREPARATION

To use thermal atoms to implement an AN, such a gas is considered to propagate in an optically produced potential landscape. There, the input of a neuron is prepared by shaping this potential. This is possible as a classical light field generates a dipole potential according to (6.12). The implementation of the neuron input is described for atoms being confined in one dimension as the concept is easily transferable to higher dimensions.

The synaptic weights of a neuron $\boldsymbol{w}$ are assumed to be stored externally. When presented with an input vector $\boldsymbol{x}$, the induced local field v of the thermal cloud neuron is calculated as

$$v = \sum_{i=0}^{n} w_i x_i. \tag{6.13}$$

As weights w_i, as well as inputs x_i can be both, negative and positive, the value range of the induced local field is $v \in [-\infty, \infty]$.

The input potential for the thermal cloud neuron is produced by a laser beam generating a Gaussian intensity distribution $I(x)$ along the x-direction where the atoms are confined

$$I(x) = I \exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right). \tag{6.14}$$

The width, described by the standard deviation σ, and the central position x_0 are always identical, independent of the neuron input. The resulting optical dipole potential is given by

$$V(x) = V_0 \exp\left(-\frac{(x-x_0)^2}{2\sigma^2}\right), \qquad V_0 = \frac{\mu I}{\Delta}. \tag{6.15}$$

As v is a real number, the amplitude of the optical dipole potential V_0 is set to be determined by the induced local field of the neuron v according to

$$V_0 = v\, k_\mathrm{B} T. \tag{6.16}$$

While the central intensity I is always positive, the detuning Δ can be negative as well. Therefore, the whole value range of v can be implemented using this procedure.

In an experimental realization, two lasers are needed for the input preparation, one being blue- and one being red-detuned to the atomic transition. The input of the thermal cloud neuron is then implemented by tuning the peak intensity of a single laser, a red-detuned one for positive v and a blue-detuned one for negative v.

6.3 NONLINEAR ACTIVATION FUNCTION

6.3.1 Implementation Algorithm

A nonlinear response to the input v is produced by changing the potential landscape seen by the atoms using an algorithmic procedure (see Fig. 6.2). First, the atomic gas is trapped in an one-dimensional box potential of width $L = 10\,\mu$m. This trapping potential is realized via periodic boundary conditions $(x, p) = (x + L, p)$, resulting in an unitary position distribution of the particles, which is characteristic for box traps.

Subsequently, the input to the neuron v is implemented via a laser, producing an additional potential with amplitude V_0 along the x-direction according to (6.15) and (6.16). Depending on the sign of v, a red- or a blue-detuned laser has to be used.

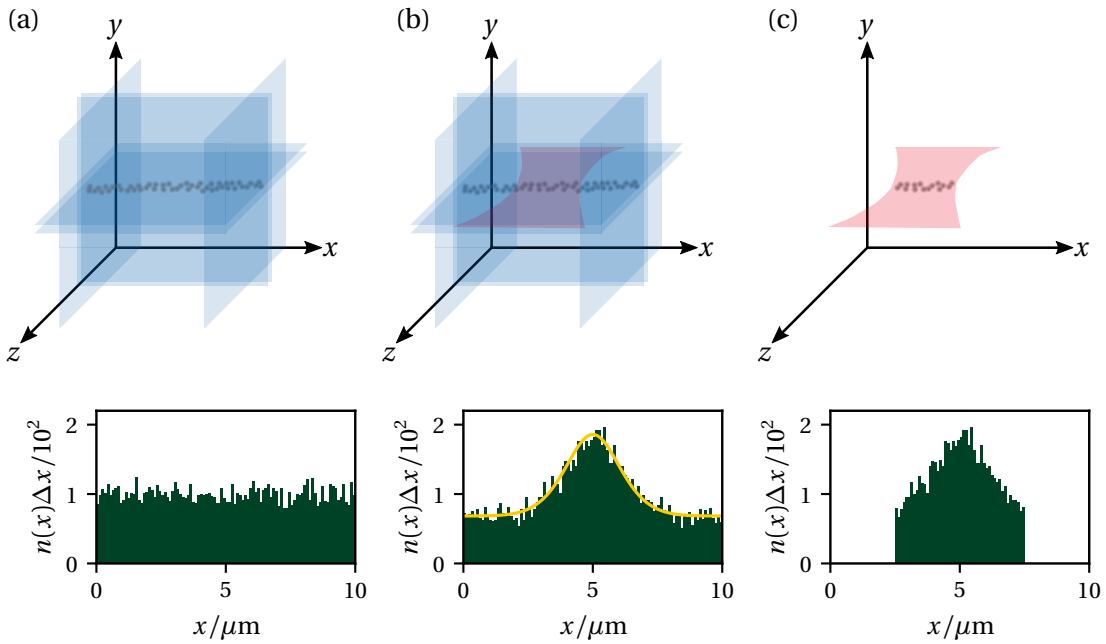

Figure 6.2: Procedure to implement the nonlinear activation function of the thermal cloud neuron. (a) The particles are initially trapped in an one-dimensional box potential with a width of $L = 10\,\mu$m, yielding an unitary position distribution along the x-direction. (b) The input of the neuron is realized via an optical potential (6.15) with central position $x_0 = 5\,\mu$m, standard deviation $\sigma = 1.25\,\mu$m and amplitude $V_0 = v\,k_B\,T$. The sign of v determines whether a red- or blue-detuned laser has to be used. After equilibration in the new potential, the position distribution is given by (6.17) (yellow). (c) The number of particles in the region $[x_0 - 2\sigma, x_0 + 2\sigma]$ is detected and identified as output of the neuron.
Numerical simulations are performed using the procedure described in (5.46) for $N = 10{,}000$ ^{87}Rb atoms cooled to the Doppler temperature (atomic parameters see Appendix B). The input is set to $v = 1$. The histograms show the densities of particle number positions $n(x)$ for a discretization $\Delta x = 0.1\,\mu$m.

According to the Maxwell-Boltzmann distribution (5.26), the position distribution in equilibrium is determined by the external potential $V(x)$. Therefore, under the influence of the input potential, the particles will distribute according to

$$n(x) = N\mathcal{N}\exp\left(-v\,e^{-(x-x_0)^2/(2\sigma^2)}\right), \tag{6.17}$$

where $\mathcal{N} = 1/\int \mathrm{d}x\ \exp\left(-v\,e^{-(x-x_0)^2/(2\sigma^2)}\right)$ (see Fig. 6.2 (b)). For an attractive potential, the particles will accumulate at the central position of the laser x_0, while they will evacuate this region for a repulsive potential. The amount of particles trapped by an attractive potential depends on the depth of the potential as well as the temperature of the gas.

6.3.2 Output Calculation

After equilibration in the new potential landscape, a detection of the number of particles in the region $[x_0-2\sigma, x_0+2\sigma]$ is performed. This can, for example, be done via an absorption measurement [151–153]. The result, normalized with regard to the total particle number N, is identified as the output of the neuron $y \in [0,1]$. Hence, y can be calculated in dependence of the input v as

$$y = \mathcal{N} \int_{x_0-2\sigma}^{x_0+2\sigma} \mathrm{d}x \, \exp\left(-v\, e^{-(x-x_0)^2/(2\sigma^2)}\right). \tag{6.18}$$

As two equilibration processes are performed during this implementation, the timescale of the neuron process using $N = 10{,}000$ ^{87}Rb atoms cooled to the Doppler temperature (atomic parameters see Appendix B) can be estimated to $t_\mathrm{N} \approx 688\,\mathrm{ms}$ (see Section 5.3).

The NAF $y(v)$, as well as its derivative $\partial_v y(v)$ with respect to the input v can be seen in Fig. 6.3. Similar to standard implementations of ANs, the activation function has a sigmoidal-like shape. The availability of an input-output-relation $y(v)$, as well as the partial derivative $\partial_v y(v)$ enables the implementation and training of a neural network.

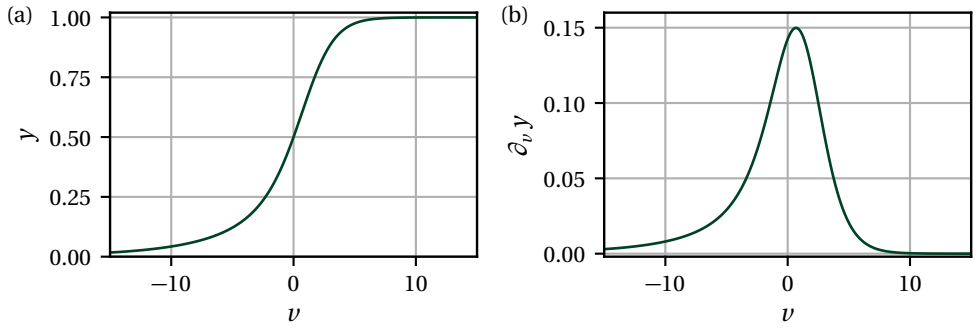

Figure 6.3: (a) Output of thermal cloud neuron y versus input v as described by (6.18) for $x_0 = 5\,\mu$m and $\sigma = 1.25\,\mu$m. (b) The derivative of y with respect to v is calculated numerically.

7

THERMAL CLOUD NEURAL NETWORK

To construct a feedforward neural network out of thermal cloud neurons, the output of a neuron has to be translated into an input for the adjacent layer. The output of a thermal cloud neuron is given by the measurement of the relative particle number remaining at the end of the neuron activity. This, in turn, determines the amplitude of the input laser of the next layer, described by the peak intensity and detuning. By repeatedly converting measured particle numbers to laser parameters, deep networks can be set up using the thermal cloud neuron. With this, exemplary network architectures can be realized and trained to perform different benchmark tasks as introduced in Section 3.3.

The implementation of the thermal cloud neural network is described in Section 7.1. Following that, this network is subjected to training on the XOR problem in Section 7.2 and on the recognition of handwritten digits based on the MNIST dataset in Section 7.3. Subsequently, the performance capabilities of the thermal cloud neural network are discussed in Section 7.4.

7.1 NETWORK IMPLEMENTATION

7.1.1 Parallelization and Layer Setup

A large number of thermal cloud neuron implementations can be run in parallel simply by choosing different central locations for the input lasers as well as by creating neighboring box traps for the thermal atoms. As these sites will be well separated, the individual neuron implementations will not influence each other. Additionally, adjacent layers of the network will be well separated, as the thermal cloud neuron algorithm terminates with a measurement, allowing for sequential realization of the layers.

7.1.2 Inter-Layer Communication

As described in the previous chapter, the output of a thermal cloud neuron j is given by the measurement of a relative particle number y_j. The value range of such an output is therefore given by $y_j \in [0,1]$. Assuming a total of m neurons in a layer of the network, the input for a neuron k in the following layer is calculated externally as

$$v_k = \sum_{j=0}^{m} w_{kj}\, y_j. \tag{7.1}$$

Accordingly, the amplitude V_0 of the input optical dipole potential (6.15) for neuron k can be determined according to (6.16) with $v = v_k$. Due to the value range of the synaptic weights w_{kj}, the input amplitude can again be positive as well as negative.

This conversion from measured particle number portions to laser parameters enables forward propagation of signals through the network, thus allowing for the possibility to build large and deep neural networks. In thermal cloud neural networks training is performed according to the backpropagation algorithm described in Section 3.2, as the synaptic weights of the network are stored externally.

7.1.3 Encoding of Binary Outputs

To be able to use the NAF of the thermal cloud neuron (see Fig. 6.3) to its full extent, an encoding scheme for inputs and outputs is required. Keeping the value range of the NAF in mind, we introduce the following encoding scheme for binary outputs, which are encountered for the benchmark examples introduced in Section 3.3: a binary output 0 is identified with a thermal cloud neuron output $y = 0.2$, while 1 is realized by $y = 0.8$. The input encodings are defined for the XOR problem and the recognition of handwritten digits separately.

7.2 XOR PROBLEM

The first investigated problem is the XOR problem as introduced in Section 3.3.1. To process this problem using a thermal cloud neural network, a fully-connected feedforward 2-2-1-network is used (cf. Fig. 3.5).

Taking into account the characteristics of the NAF of the thermal cloud neurons, inputs to the network are implemented as $x_i = 0.1$ for inputs 0 and as $x_i = 1$ for inputs 1 of the XOR problem. The complete encoding scheme can be seen in Table 7.1.

Table 7.1: Input-output mapping for the XOR problem using the thermal cloud neural network.

Input 1	Input 2	x_1	x_2	Output	y
0	0	0.1	0.1	0	0.2
0	1	0.1	1.0	1	0.8
1	0	1.0	0.1	1	0.8
1	1	1.0	1.0	0	0.2

The backpropagation algorithm described in Section 3.2 can be used to train the network to solve the XOR problem. In agreement with the procedure in Section 3.3.1, the on-line training is performed over 200 epochs with 1,000 randomized training samples per epoch. The learning rate of the training is set to $\eta = 0.1$ for the first epoch, $\eta = 0.01$ for the following 49 epochs and to $\eta = 0.005$ for every epoch from this point onwards. After each epoch, the performance of the network in solving the XOR problem is tested by calculating the averaged squared error $\mathcal{E}$ according to (3.2) for all $m = 4$ possible input-output pairs of the XOR problem.

Starting from random initial values for the weights and biases of the network, $\mathcal{E}$ decreases rapidly over the course of the training procedure (see Fig. 7.1). A sample is classified as being identified correctly, if the network response is within ± 0.1 of the desired value. At the end of the training procedure, the thermal cloud neural network is able to identify every test sample correctly.

While not reaching the same performance as the network implementation presented in Section 3.3.1,the thermal cloud neural network in the presented architecture can be trained to solve the XOR problem. Therefore, there is proof-of-principle that thermal cloud neurons can be used to set up trainable feedforward neural networks.

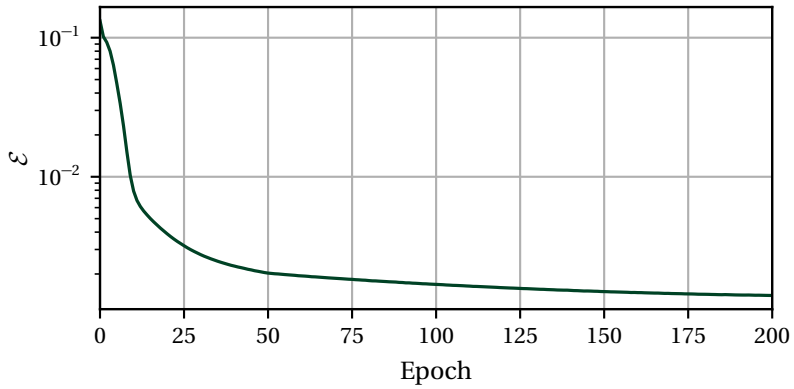

Figure 7.1: Averaged squared error $\mathcal{E}$ versus epochs for solving the XOR problem using thermal cloud neurons. Starting from random values for weights and biases, $\mathcal{E}$ decreases rapidly.

7.3 MNIST DATASET

In line with Section 3.3.2, thermal cloud neurons are also used to set up a network for the recognition of handwritten digits based on the MNIST dataset. The network architecture is chosen to be a fully-connected 784-300-10-network. The grayscale values of the 784-dimensional input vector are normalized to the range $[0,1]$. The inputs to the thermal cloud neural network are chosen accordingly, lying in the range $x_i \in [0,1]$.

Again, the backpropagation algorithm introduced in Section 3.2 is used to train the network in an on-line manner. Starting from random weights and biases, the training procedure is performed over 100 epochs with a fixed learning rate of $\eta = 0.05$. In every epoch, the network is subjected to 6,000 random training samples. The performance capability of the network is investigated by determining the averaged squared error $\mathcal{E}$ according to (3.2) using $m = 1{,}000$ randomly chosen test samples after each epoch (see Fig. 7.2). Throughout the course of the training, $\mathcal{E}$ is decreased, reaching $\mathcal{E} = 2.3 \cdot 10^{-3}$ at the end of the 100^{th} epoch.

We define the network to identify a sample with a digit, if the output of the corresponding neuron is within a margin of ± 0.2 of the desired output value of $y = 0.8$. If none of the neurons is within this range, the sample is categorized as being not identifiable by the network. After

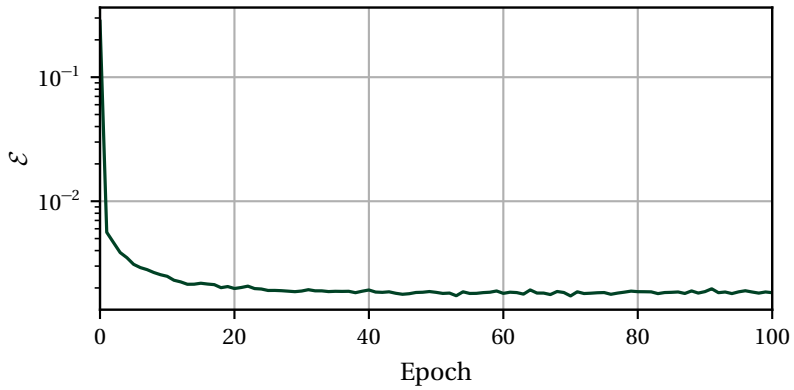

Figure 7.2: Averaged squared error $\mathcal{E}$ versus epochs for the recognition of handwritten digits based on the MNIST dataset using a 784-300-10 thermal cloud neural network. At the end of the procedure, $\mathcal{E}$ is reduced to $\mathcal{E} = 2.3 \cdot 10^{-3}$.

Table 7.2: Sample digit versus digits identified by a 784-300-10 thermal cloud neural network after 100 training epochs in percent. A test sample is assigned to a digit, if the corresponding neuron produces an output $y = 0.8 \pm 0.2$. If no neuron is within this range, the sample is not identified by the network and is categorized under "-". In total, the network has an error rate of 16.45%

	0	1	2	3	4	5	6	7	8	9	-
0	93.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.7
1	0.0	97.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.4
2	0.0	0.0	81.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	18.9
3	0.0	0.0	0.0	75.9	0.0	1.1	0.0	0.0	0.0	0.0	23.0
4	0.0	1.0	0.0	0.0	83.5	0.0	1.0	0.0	0.0	0.0	14.6
5	0.0	0.0	0.0	0.0	0.0	75.0	0.0	0.0	0.0	0.0	25.0
6	0.0	0.2	0.0	0.0	0.0	0.0	77.2	0.0	0.0	0.0	22.8
7	0.0	0.4	0.0	0.0	0.0	0.0	0.0	85.6	0.0	0.0	14.4
8	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	85.3	0.0	12.7
9	0.0	0.3	0.0	0.0	2.0	0.0	0.0	0.0	0.0	81.0	17.0

the training procedure, the network identifies 847 out of the 1,000 test samples. A comparison between the network responses and the actual digits of the test samples can be seen in Table 7.2.

While the network is not able to identify all test samples, very few samples are identified incorrectly. In total, the network has an average error rate of 16.45%. This is significantly larger than the exemplary implementation using ANs and a sigmoidal NAF shown in Section 3.3.2 despite comparable values of $\mathcal{E}$ at the end of the training procedure. This hints at the fact, that the NAF of the thermal cloud neuron might not be very well suited for this problem. Yet, the network is still able to identify the majority of the test samples correctly, showing that there is the possibility for real-world application of the thermal cloud neural network.

7.4 DISCUSSION

The thermal cloud neural network demonstrates the possibility to use cold atoms in an optical setup to implement a neural network. Using an algorithmic procedure, a nonlinear response to a weighted input can be produced. Additionally, multiple neuron realizations can be run in parallel as well as in adjacent layers, enabling the setup of deep neural networks.

However, there are two main disadvantages in the presented implementation. Firstly, the thermal cloud neural network is not fully integrated. After each layer, the information encoded in the thermal atoms has to be extracted via a measurement to produce the inputs for the next layer. Therefore, one has to interact actively with the network during the processing of a signal, constantly converting between electronic and optical signals.

Secondly, due to the nature of thermal ^{87}Rb atoms cooled to their Doppler temperature, which we suggest to use in the implementation, the whole process is rather slow. As shown before, one neuron implementation takes a couple of hundreds of milliseconds. Even under the assumption that neurons in one layer can be run perfectly in parallel without any delay and that the timespan for the measurement of the particle numbers as output and implementation of the new laser parameters as input can be neglected, a run through a two-layer thermal cloud neural network takes more than one second. In comparison to state-of-the-art physical

implementations of artificial neural networks (see Chapter 4), this is quite slow, possibly preventing the applicability in the real world.

Yet, the thermal cloud neural network is a good platform to understand the intricacies in the implementation of an artificial neural network. One has to have very good control over the NAF and has to be able to encode a problem for the network in such a way that makes full use of the NAF input and output range.

Part III

FOUR-WAVE MIXING NEURAL NETWORK

8

COHERENT MATTER WAVES

In this part, the implementation of an AN and an ANN based on the four-wave mixing (FWM) process in coherent matter waves is presented. In contrast to thermal atoms, Bose-Einstein condensed gases show an intrinsic nonlinearity. We aim to harness this characteristic to set up a nonlinear activation function of an AN, subsequently culminating in the description of the FWM neuron. Bose-Einstein condensates (BECs), as well as their dynamics in terms of the Gross-Pitaevskii equation (GPE), are discussed in Chapter 8. In Chapter 9, the FWM process with plane waves in coherent matter waves is investigated in detail, yielding a Josephson-like description of the dynamics. On the basis of the population and phase oscillations in FWM, the implementation of the FWM neuron is introduced in Chapter 10, also showing its capability to solve the XOR problem. Finally, parallelization opportunities and possible ways to set up communicating layers using the FWM neuron are discussed in Chapter 11, manifesting the FWM neural network.

In this chapter the basic properties of BECs are discussed. After a review of BECs and their state-of-the-art realizations and applications in Section 8.1, the dynamics of coherent matter waves in the context of a mean-field theory are analyzed in Section 8.2, yielding the Gross-Pitaevskii equation.

8.1 BOSE-EINSTEIN CONDENSATES

8.1.1 Bose-Einstein Condensation

The theoretical prediction of Bose-Einstein condensation dates back nearly 100 years to the work of *Satyendranath Bose* [154] and *Albert Einstein* [155]. They showed that in a gas of non-interacting, massive bosons a finite fraction of the total number of particles would occupy the lowest-energy single-particle state below a critical temperature.

Superfluid liquid Helium-4 is a prototype BEC as pointed out by *Fritz London* in 1938 [156]. Yet, due to the strong interaction between helium atoms, the number of atoms in the zero-momentum state is reduced even at absolute zero temperature. Therefore, the occupancy of the zero-momentum state is difficult to measure directly.

The fact that interactions reduce the occupation of the lowest-energy single-particle state dramatically led to the search for weakly interacting Bose gases to achieve a higher condensate fraction. As such a possible candidate, spin-polarized hydrogen was identified [126]. A gas of hydrogen atoms in a magnetic field would be stable against the formation of molecules and would remain in a gaseous phase to arbitrarily low temperatures. Indeed, Bose-Einstein condensation of atomic hydrogen was achieved in 1998 [157].

However, at this point in time, atomic hydrogen was not the most promising candidate for Bose-Einstein condensation anymore. Due to the dramatic advances made in laser cooling of

alkali atoms, such systems became prime candidates to achieve Bose-Einstein condensation[1]. In 1995, the groups of *Eric A. Cornell*, *Wolfgang Ketterle* and *Carl E. Wieman* reported the observation of Bose-Einstein condensation in trapped, dilute, weakly interacting atomic gases of Rubidium-87 and Sodium-23, earning them the Nobel prize in 2001[2].

8.1.2 Bose-Einstein Condensation in Dilute Atomic Gases

Bose-Einstein condensation in an ideal gas sets in when the temperature falls below a critical temperature T_c, such that the thermal de Broglie wavelength

$$\lambda_{\mathrm{dB}} = \sqrt{\frac{2\pi\hbar^2}{m k_{\mathrm{B}} T}} \tag{8.1}$$

is comparable to the mean interparticle spacing

$$d = n^{-1/3}, \tag{8.2}$$

where n is the particle density [126]. Thus,

$$\lambda_{\mathrm{dB}}(T_c) = d, \qquad T_c \simeq 3.3125 \frac{\hbar^2 n^{2/3}}{m k_{\mathrm{B}}}. \tag{8.3}$$

A general definition of a BEC, also including interactions, is given by *Penrose* and *Onsager* [160]. This states, that a BEC occurs for the case that N_c, the largest eigenvalue of the single particle density matrix

$$\rho(x, x') = \langle \psi^\dagger(x)\psi(x') \rangle, \tag{8.4}$$

is of the order of the total particle number in the system. Equivalently, this criterion corresponds to the existence of off-diagonal long range order,

$$\rho(x, x') \to n_c \neq 0, \qquad |x - x'| \to 0, \tag{8.5}$$

in the system [161].

A typical starting point for experimental realization of a BEC is an atomic gas at room temperature. The atoms are trapped and cooled by laser cooling methods to about $10\,\mu$K [132]. Subsequently, the cooled down cloud is trapped magnetically via the Zeeman interaction of the electron spin with an inhomogeneous magnetic field. In the magnetic trap, the cloud of atoms is cooled down further by evaporative cooling. There, the highest-energy atoms evaporate from the trap by converting their spin to an untrapped state via external radiation. With the escape of those atoms, the average energy of the remaining atoms is reduced. Using these methods, atomic clouds can be cooled to the nano-Kelvin range while achieving densities of about $10^{15}\,\mathrm{cm}^{-3}$. For dilute gases in a three-dimensional harmonic trap, those values cause the phase space density

$$\rho = n \lambda_{\mathrm{dB}}^3 \tag{8.6}$$

1 The Nobel prize in physics 1997 was awarded to *Steven Chu*, *Claude Cohen-Tannoudji* and *William D. Phillips* "for the development of methods to cool and trap atoms with laser light" [158]

2 The Nobel prize in physics 2001 was awarded to *Eric A. Cornell*, *Wolfgang Ketterle* and *Carl E. Wieman* "for the achievement of Bose-Einstein condensation in dilute gases of alkali atoms, and for early fundamental studies of the properties of the condensates" [159]

to exceed the critical value for the transition towards a BEC of [126]

$$\rho_c \simeq 2.612. \tag{8.7}$$

After the first realization of a BEC in cold atomic gases, research focused on equilibrium properties, elementary and nonlinear excitations (e.g. vortices) and effects of thermal fluctuations [162]. Subsequent work moved towards enhanced control of the interaction strengths via Feshbach resonances [163], exploring the dimensionality of the BECs [135] and setting up quantum simulators of solid states e.g. via optical lattice potentials [134].

Nowadays, there is a plethora of active experimental research regarding Bose-Einstein condensation in atomic systems [162]. Those range from BECs in guided or controlled geometries, e.g. atom chips [164], ring-traps [165] and box-like potentials [138], to multicomponent and spinor BECs [166] and nonlinear effects including dark and bright solitons [167]. Further fields of high interest are the superfluid-Mott insulator transition [168], observations of the Josephson effect [169, 170], the use of atoms in optical lattices to create strongly correlated systems as quantum simulators of condensed matter systems [171] and atom interferometry [172]. Ongoing theoretical investigations concern the justification and applicability of the GPE [173], nonequilibrium thermodynamics, effects of disorder [174, 175], and many more.

8.1.3 Further Approaches to Bose-Einstein Condensation

While Bose-Einstein condensation in atomic gases is well-studied and still provides a thriving research field, the universal phenomenon exhibits a multitude of diverse manifestations. Already in the early 1990s, excitons in semiconductors were thought of as a good candidate for Bose-Einstein condensation. Excitons are quasiparticles formed by an electron and an electron hole attracted via the Coulomb force inside a semiconductor. Due to their property of being a boson, Bose-Einstein condensation is within the realm of possibility. Recently, two main methods to achieve condensation in excitonic systems were explored. Firstly, magnetic fields and electronic gating were used to control the densities of electrons and holes in a coupled quantum well system. The pairing of electrons and holes in such systems yields thermodynamically stable excitons, enabling condensation [176]. The alternative approach is to place excitons in a high-Q optical cavity[3], creating strong coupling between the cavity and the exciton state. Eigenstates of such systems are known as polaritons. Through coupling, the effective mass of the exciton is reduced, implying a larger de Broglie wavelength and therefore larger quantum effects at typical densities and temperatures [178, 179].

The usage of high-Q cavities can also be exploited to achieve Bose-Einstein condensation in photons [180]. While the cavity provides an increased lifetime of the photons, repeated absorption and re-emission into dye molecules inside the cavity enables thermalization of the photons.

Similarly, Bose-Einstein condensation of magnons created as a nonequilibrium population in a ferromagnet can be realized [181]. There, magnons are quasiparticles corresponding to quantized spin waves describing the collective motion of spins.

Additionally, the question of Bose-Einstein condensation in even more exotic systems is of high interest. This includes BECs of neutrons in neutron stars [182], kaon condensation across the whole universe [183], axion condensation as a candidate for cold dark matter [184, 185] and graviton BECs as a new approach to quantum gravity [162, 186].

3 The Q factor of a resonant cavity is given by $Q = \omega_0 E/P$, where ω_0 is the resonant frequency, E is the energy stored in the cavity and $P = -\partial E/\partial t$ is the dissipated power. The average lifetime of a resonant photon inside the cavity is proportional to the cavity's Q factor [177].

Hence, Bose-Einstein condensation is a vast field in modern research. As such, a theoretical description of the equilibrium and dynamical properties of such systems is needed. In this thesis, we focus on a mean-field description of the condensate yielding the GPE.

8.2 GROSS-PITAEVSKII EQUATION

The system under consideration is a bosonic fluid of neutral atoms of mass m without spin confined in space by an external potential $U(\boldsymbol{r})$ at $T = 0\,\mathrm{K}$. This can be described by a Hamiltonian of a pairwise interacting many-body system

$$\hat{H} = \sum_i \hat{H}_0(\boldsymbol{r}_i, \boldsymbol{p}_i) + \frac{1}{2}\sum_{i,j} \hat{V}(\boldsymbol{r}_i - \boldsymbol{r}_j), \qquad \hat{H}_0 = -\frac{\hbar^2}{2m}\nabla^2 + U(\boldsymbol{r}). \tag{8.8}$$

As this is a system of identical particles, it is convenient to rewrite the Hamiltonian in second quantization [105, 187]

$$\begin{aligned}\hat{H} = &\int \mathrm{d}^3 r\, \mathrm{d}^3 r'\, \hat{\psi}^\dagger(\boldsymbol{r}')\langle \boldsymbol{r}'|\hat{H}_0|\boldsymbol{r}\rangle\hat{\psi}(\boldsymbol{r}) \\ &+ \frac{1}{2}\int \mathrm{d}^3 r_1'\, \mathrm{d}^3 r_2'\, \mathrm{d}^3 r_1\, \mathrm{d}^3 r_2\, \hat{\psi}^\dagger(\boldsymbol{r}_1')\hat{\psi}^\dagger(\boldsymbol{r}_2')\langle \boldsymbol{r}_1'\boldsymbol{r}_2'|\hat{V}|\boldsymbol{r}_1\boldsymbol{r}_2\rangle\hat{\psi}(\boldsymbol{r}_2)\hat{\psi}(\boldsymbol{r}_1).\end{aligned} \tag{8.9}$$

The field operators $\hat{\psi}(\boldsymbol{r})$ and $\hat{\psi}^\dagger(\boldsymbol{r})$ destroy or create bosonic particles at position $\boldsymbol{r}$ in Fock space

$$|\boldsymbol{r}\rangle = \hat{\psi}^\dagger(\boldsymbol{r})|0\rangle. \tag{8.10}$$

They satisfy the bosonic commutator relations

$$\left[\hat{\psi}(\boldsymbol{r}), \hat{\psi}^\dagger(\boldsymbol{r}')\right] = \delta(\boldsymbol{r} - \boldsymbol{r}'), \qquad \left[\hat{\psi}(\boldsymbol{r}), \hat{\psi}(\boldsymbol{r}')\right] = 0. \tag{8.11}$$

8.2.1 Effective Interactions

In dilute gases interaction strengths are small for typical atomic separations. Only if two atoms approach each other, the many-body wave function shows rapid spatial variations. To avoid the calculation of short-range correlations between atoms, it is convenient to introduce an effective interaction. Using this approximation, one enters the realm of a mean-field theory.

Considering two particles of equals mass without internal degrees of freedom, the wave function for the relative motion of the particles is given as the sum of an incoming plane wave and a scattered wave $\psi_{\mathrm{sc}}(\boldsymbol{r})$ [126]

$$\psi(\boldsymbol{r}) = e^{\mathrm{i}\boldsymbol{k}\boldsymbol{r}} + \psi_{\mathrm{sc}}(\boldsymbol{r}). \tag{8.12}$$

In momentum representation, this is equivalent to

$$\psi(\boldsymbol{k}) = (2\pi)^3\delta(\boldsymbol{k} - \boldsymbol{k}') + \psi_{\mathrm{sc}}(\boldsymbol{k}). \tag{8.13}$$

This wave function satisfies the Schrödinger equation in momentum space

$$\frac{\hbar^2}{m}\left(k^2 - k'^2\right)\psi_{\mathrm{sc}}(\boldsymbol{k}') = V(\boldsymbol{k} - \boldsymbol{k}') + \frac{1}{\mathcal{V}}\sum_{\boldsymbol{k}''} V(\boldsymbol{k}' - \boldsymbol{k}'')\psi_{\mathrm{sc}}(\boldsymbol{k}''), \tag{8.14}$$

where $\mathcal{V}$ is a volume. Thus, the scattered wave is given by

$$\psi_{\mathrm{sc}}(\boldsymbol{k}') = \left[\frac{\hbar^2}{m}(k^2 - k'^2) + \mathrm{i}\delta\right]^{-1} T\left(\boldsymbol{k} - \boldsymbol{k}'; \frac{\hbar^2 k^2}{m}\right). \tag{8.15}$$

The infinitesimal imaginary part $\mathrm{i}\delta$, $\delta > 0$, is introduced to ensure that only outgoing waves are present in the scattered wave [187]. The scattering matrix T satisfies the Lippmann-Schwinger equation [126]

$$T(\boldsymbol{k} - \boldsymbol{k}'; E) = V(\boldsymbol{k} - \boldsymbol{k}') + \frac{1}{\mathcal{V}} \sum_{\boldsymbol{k}''} V(\boldsymbol{k}' - \boldsymbol{k}'') \left(E - \frac{\hbar^2 k''^2}{m} + \mathrm{i}\delta\right)^{-1} T(\boldsymbol{k} - \boldsymbol{k}''; E). \tag{8.16}$$

At large distances and for zero energy ($E = k = 0$), the scattered wave in coordinate space is given by

$$\psi_{\mathrm{sc}}(r) = -\frac{m\,T(0;0)}{4\pi\hbar^2} \frac{1}{r}. \tag{8.17}$$

Here, one can identify the s-wave scattering length

$$a_{\mathrm{s}} = \frac{m}{4\pi\hbar^2} T(0;0). \tag{8.18}$$

In Born approximation only the first term on the right hand side of the Lippmann-Schwinger equation (8.16) is taken into account. Hence, an effective interaction potential is given by

$$V(0) = T(0;0) = \frac{4\pi\hbar^2 a_{\mathrm{s}}}{m} := g. \tag{8.19}$$

In coordinate space, this result corresponds to a contact interaction

$$V(\boldsymbol{r} - \boldsymbol{r}') = g\,\delta(\boldsymbol{r} - \boldsymbol{r}') \tag{8.20}$$

where $\boldsymbol{r}$ and $\boldsymbol{r}'$ are the positions of the two particles. Already in 1936, *Enrico Fermi* introduced this concept of a pseudopotential to describe the s-wave scattering of a free neutron by a nucleus [188].

8.2.2 Many-Body Schrödinger Equation

Using the contact potential (8.20), the Hamiltonian (8.9) simplifies to

$$\hat{H} = \int \mathrm{d}^3 r\, \hat{\psi}^\dagger(\boldsymbol{r}) \left[-\frac{\hbar^2}{2m}\nabla^2 + U(\boldsymbol{r})\right] \hat{\psi}(\boldsymbol{r}) + \frac{g}{2} \int \mathrm{d}^3 r\, \hat{\psi}^\dagger(\boldsymbol{r}) \hat{\psi}^\dagger(\boldsymbol{r}) \hat{\psi}(\boldsymbol{r}) \hat{\psi}(\boldsymbol{r}). \tag{8.21}$$

A general many-body state may be written as

$$|\psi\rangle = \sum_{N=0}^{\infty} \int \mathrm{d}r^{3N}\, |r_1 \dots r_N\rangle \langle r_1 \dots r_N | \psi\rangle_N, \tag{8.22}$$

with the particle number N and the order of the assignment $\psi(r_1, \dots, r_N) = \langle r_1 \dots r_N | \psi\rangle_N$. The equations of motion of this state can be recovered using the Lagrangian density [189, 190]

$$\mathcal{L} = \langle \psi | \mathrm{i}\hbar\partial_t - \hat{H} | \psi \rangle. \tag{8.23}$$

From this, the canonical momentum can be determined as

$$\pi = \frac{\delta \mathcal{L}}{\delta \dot{\psi}} = \mathrm{i}\hbar\psi^*. \tag{8.24}$$

According to the Hamiltonian equations of motion,

$$\dot{\pi} = \frac{\delta \mathcal{L}}{\delta \psi}. \tag{8.25}$$

The Euler-Lagrange equations yield the Schrödinger equation in Fock space

$$\mathrm{i}\hbar\partial_t|\psi\rangle = \hat{H}|\psi\rangle. \tag{8.26}$$

8.2.3 Gross-Pitaevskii Equation

In a BEC, the atoms in the system occupy the same quantum states macroscopically. Therefore, the state of the system can be approximated by a coherent state [191]

$$\hat{\psi}(\boldsymbol{r})|\psi\rangle_\mathrm{c} = \psi(\boldsymbol{r})|\psi\rangle_\mathrm{c}, \tag{8.27}$$

where the eigenvalue $\psi(\boldsymbol{r})$ is called the order parameter of the coherent state of bosons. Combining (8.21) and (8.27), the Lagrangian functional for the condensate is given by

$$L = \int \mathrm{d}^3 r \left(\mathrm{i}\hbar\psi^*\partial_t\psi - \mathcal{E}(\boldsymbol{r})\right) \tag{8.28}$$

with the energy density

$$\mathcal{E}(\boldsymbol{r}) = \frac{\hbar^2}{2m}|\nabla\psi|^2 + U(\boldsymbol{r})n(\boldsymbol{r}) + \frac{g}{2}n^2(\boldsymbol{r}). \tag{8.29}$$

Here, $n(\boldsymbol{r}) = |\psi(\boldsymbol{r})|^2$ denotes the particle density.

Evaluating the Euler-Lagrange field equation using functional differentiation

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta L}{\delta \dot{\psi}} = \frac{\delta L}{\delta \psi} \tag{8.30}$$

yields the Gross-Pitaevskii equation [192, 193]

$$\mathrm{i}\hbar\partial_t\psi(\boldsymbol{r},t) = \left[-\frac{\hbar^2\nabla^2}{2m} + U(\boldsymbol{r}) + g\,n(\boldsymbol{r})\right]\psi(\boldsymbol{r},t). \tag{8.31}$$

The total particle number of the system is conserved $N = \int \mathrm{d}^3 r\; n(\boldsymbol{r})$.

9

FOUR-WAVE MIXING WITH PLANE WAVES

In this chapter, we present an in-detail discussion of the FWM process with plane waves based on [36]. So far, theoretical investigations of FWM in coherent matter waves typically considered slowly-varying envelope approximations of three-dimensional wave packets due to the experimental possibilities at the time [30, 194]. With the development and subsequent refinement of optical box traps (cf. Section 6.1.2), arbitrary-dimensional homogeneous BECs became experimentally accessible [138, 139]. Therefore, an in-depth description of FWM with plane waves in homogeneous BECs is of theoretical as well as experimental interest.

After introducing the FWM process in optical systems as well as in coherent matter waves in Section 9.1, we discuss the ideal FWM process in homogeneous BECs in Section 9.2. This investigation yields a Josephson-like description of the population and phase dynamics of the coherent matter waves participating in the process. In the discussion of ideal FWM, two setups are considered: two-dimensional FWM and one-dimensional FWM with multiple internal states of the atoms. In Section 9.3, the influence of population in additional momentum states on the FWM dynamics is investigated using numerical simulations of the GPE on a discrete periodic grid.

9.1 FOUR-WAVE MIXING

9.1.1 Four-Wave Mixing in Nonlinear Optics

Nonlinear Optics

The process of FWM is well-known in nonlinear optics [28]. Nonlinear optics studies the modification of optical properties of a material in the presence of light and the phenomena that occur as a consequence. In conventional (linear) optics, the polarization $P^{(1)}(t)$ induced in a medium depends linearly on the applied electric field strength $E(t)$ according to

$$P^{(1)}(t) = \varepsilon_0 \int_{-\infty}^{t} \mathrm{d}t'\, \chi^{(1)}(t-t')E(t'). \tag{9.1}$$

There, the proportionality constant $\chi^{(1)}$ is the linear susceptibility of the medium. The integration range from $-\infty$ to t reflects the fact, that the polarization depends only on past and not on future values of the electric field strength. This expression can be Fourier transformed to the frequency domain, yielding

$$P^{(1)}(\omega) = \varepsilon_0 \chi^{(1)}(\omega)E(\omega). \tag{9.2}$$

An extension to include nonlinear responses of the medium to presented electric field strengths is possible. For an isotropic medium, the polarization $P(\omega)$ can be expressed using a power series in $E(\omega)$, as

$$P(\omega) = \varepsilon_0 \left(\chi^{(1)}(\omega) E(\omega) + \chi^{(2)}(\omega) E^2(\omega) + \chi^{(3)}(\omega) E^3(\omega) + \dots \right). \tag{9.3}$$

$\chi^{(2)}(\omega)$ and $\chi^{(3)}(\omega)$ are called second- and third-order nonlinear susceptibility, respectively.

In typical media, the value of the susceptibilities decreases by orders of magnitudes with increasing susceptibility order. Well-known second-order nonlinear effects are second-harmonic, sum- and difference-frequency generation. Third-order processes include third-harmonic generation, self-focusing and optical mixing processes.

Degenerate Optical Four-Wave Mixing

In this thesis, we are interested in a particular third-order optical mixing process, namely FWM. In optical systems, this process is often realized as degenerate FWM where all four interacting waves have the same frequency. In this process, a lossless nonlinear medium characterized by $\chi^{(3)}$ is illuminated by two strong counter-propagating pump waves $\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ and a signal wave $\boldsymbol{E}_3$. As a result of the nonlinear interaction between those waves, a fourth wave $\boldsymbol{E}_4$ is generated. The geometry of degenerate FWM can be seen in Fig. 9.1.

An in-detail calculation of the degenerate FWM process can e.g. be found in [28]. Here, we present a shortened version to highlight to generation of the fourth wave. The four waves can all be assumed to be plane waves represented as

$$\boldsymbol{E}_i(\boldsymbol{r}, t) = \int_{-\infty}^{\infty} \frac{\mathrm{d}^3 k}{(2\pi)^{3/2}} \frac{\mathrm{d}\omega}{(2\pi)^{1/2}} \, \delta(\boldsymbol{k} - \boldsymbol{k_i}) \delta(\omega - \omega_i) \boldsymbol{E}(\boldsymbol{k}, \omega) e^{\mathrm{i}(\boldsymbol{k}\boldsymbol{r} - \omega t)} = \boldsymbol{E}_i e^{\mathrm{i}(\boldsymbol{k}_i \boldsymbol{r} - \omega_i t)}, \tag{9.4}$$

where $i = 1, \dots, 4$ and $\boldsymbol{E}_i$ are slowly-varying amplitudes. In addition to a large number of other terms, the nonlinear polarization $\boldsymbol{P}_{\mathrm{nl}}$ produced within the medium will have a term of the form

$$\boldsymbol{P}_{\mathrm{nl}} = 6\chi^{(3)} \boldsymbol{E}_1 \boldsymbol{E}_2 \boldsymbol{E}_3^* e^{\mathrm{i}(\boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{k}_3)\boldsymbol{r}} \tag{9.5}$$

in the degenerate case ($\omega_1 = \omega_2 = \omega_3 = \omega_4$). As the two pump waves are counter-propagating, the relation between their wave vectors is given by

$$\boldsymbol{k}_1 + \boldsymbol{k}_2 = 0, \tag{9.6}$$

yielding

$$\boldsymbol{P}_{\mathrm{nl}} = 6\chi^{(3)} \boldsymbol{E}_1 \boldsymbol{E}_2 \boldsymbol{E}_3^* e^{-\mathrm{i}\boldsymbol{k}_3 \boldsymbol{r}}. \tag{9.7}$$

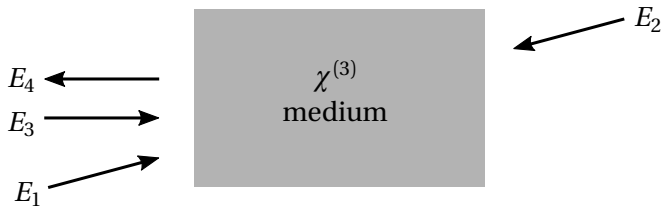

Figure 9.1: Geometry of degenerate FWM in a $\chi^{(3)}$ nonlinear medium. Graphic adapted from [28].

This contribution acts as a phase-matched source term for a conjugate wave $\boldsymbol{E}_4$ with wave vector

$$\boldsymbol{k}_4 = -\boldsymbol{k}_3. \tag{9.8}$$

The amplitude of the generated fourth wave will be proportional to $\boldsymbol{E}_1 \boldsymbol{E}_2 \boldsymbol{E}_3^*$.

There is also an alternative explanation of the degenerate FWM process. The incoming signal wave $\boldsymbol{E}_3$ together with e.g. the pump wave $\boldsymbol{E}_1$ forms a spatially varying intensity distribution. This interference pattern causes a change in the refractive index due to the nonlinear response of the medium. The second pump wave $\boldsymbol{E}_2$ is scattered on the resulting diffraction grating, producing the outgoing conjugate wave $\boldsymbol{E}_4$.

The FWM process exists beyond the degenerate case. Then, to satisfy the phase matching conditions in the nonlinear polarization

$$\boldsymbol{P}_{\mathrm{nl}} = 6\chi^{(3)} \boldsymbol{E}_1 \boldsymbol{E}_2 \boldsymbol{E}_3^* e^{\mathrm{i}[(\boldsymbol{k}_1+\boldsymbol{k}_2-\boldsymbol{k}_3)\boldsymbol{r}-(\omega_1+\omega_2-\omega_3)t]}, \tag{9.9}$$

two relations have to hold:

$$\boldsymbol{k}_1 + \boldsymbol{k}_2 = \boldsymbol{k}_3 + \boldsymbol{k}_4, \qquad \omega_1 + \omega_2 = \omega_3 + \omega_4. \tag{9.10}$$

Therefore, the dispersion relation $\omega(\boldsymbol{k})$ of the medium has to be taken into account. For the special case of linear dispersion

$$\omega(\boldsymbol{k}) = c|\boldsymbol{k}|, \tag{9.11}$$

holding for electromagnetic waves in vacuum, a one-dimensional realization of FWM can be found where all wave vectors are parallel.

9.1.2 Four-Wave Mixing in Coherent Matter Waves

Following the advent of the BEC in cold atomic gases (cf. Section 8.1), theoretical investigations as well as experiments demonstrated a process equivalent to FWM in optical systems. There, momentum components of the BEC took over the role of optical frequencies.

In 1998, *Trippenbach et al.* [29] first investigated the feasability of realization of FWM in BECs using numerical simulations of the GPE with three interacting wave packets. Already then, they pointed out a key difference to the optical case: whereas the dispersion relation is linear in k for the case of light, it is quadratic in k for massive particles. As the creation of a fourth wave packet through FWM is limited to cases where momentum and energy conservation are simultaneously satisfied, not all geometries (for example a one-dimensional implementation) from the optical case are readily available using BECs.

Shortly after the theoretical proposal of the process, FWM in matter waves was realized by the group of *William D. Phillips* in 1999 [31]. Using Bragg diffraction, they simultaneously created three overlapping wave packets with different momenta in a BEC of sodium atoms. Using a half-collision geometry, they observed a fourth wave packet after separation of the momentum components produced by the nonlinear interaction inside the BEC. Further experiments by the group of *Wolfgang Ketterle* showed that FWM can e.g. also be used to create pair-correlated atomic beams [32].

Subsequent research on FWM in coherent matter waves was quite diverse. While early theoretical investigations continued to focus on numerical calculations, specifically considering wave packets [30, 195], also exact quantum theories of collinear FWM in a multi-component

BEC were developed [196]. Such setups were also utilized in the realm of the GPE, showing that collinear FWM in BECs with multiple spin states enables the realization of one-dimensional geometries [197, 198]. Additionally, work has been dedicated to FWM in periodic optical potentials, showing that the specific dispersion relation can also be used to implement one-dimensional FWM [199].

Due to the experimental availability of homogeneous BECs in arbitrary dimensions [138, 139], a theoretical investigation of FWM with plane waves should now be of particular interest. Therefore, we present such an investigation, showing that the population and phase dynamics of the momentum components can be described via Josephson-like oscillations.

9.2 IDEAL FOUR-WAVE MIXING

9.2.1 General Description

Plane Wave Ansatz

The dynamics of a weakly interacting BEC described by the order parameter $\psi(\boldsymbol{r},t)$ are given by the GPE (8.31). The Lagrangian functional of the GPE is given by [189, 190]

$$L=\int \mathrm{d}^3 r\,(\mathrm{i}\hbar\psi^*\partial_t\psi-\mathcal{E}(\boldsymbol{r})), \tag{9.12}$$

with the expression for the energy density

$$\mathcal{E}(\boldsymbol{r})=\frac{\hbar^2}{2m}|\nabla\psi|^2+Un+\frac{g}{2}n^2. \tag{9.13}$$

In the following, we consider the case of a homogeneous BEC with $U=0$ and periodic boundary conditions. Then, a wave function $|\psi\rangle=|\psi_\alpha\rangle+|\psi_\beta\rangle$ is a coherent superposition of plane waves $|\boldsymbol{k}_j\rangle$ with complex amplitudes α_j and β_j. It consists of a FWM state

$$\left|\psi_\alpha\right\rangle=\sum_{j=1}^{4}\sqrt{N}\alpha_j\left|\boldsymbol{k}_j\right\rangle \tag{9.14}$$

and a residual wave

$$\left|\psi_\beta\right\rangle=\sum_{j>4}\sqrt{N}\beta_j\left|\boldsymbol{k}_j\right\rangle, \tag{9.15}$$

which is orthogonal $\langle\psi_\alpha|\psi_\beta\rangle=0$ to the FWM state.

The complex amplitudes α_j, in terms of absolute value and phase, are given by

$$\alpha_j=\sqrt{n_j}e^{-i\varphi_j}. \tag{9.16}$$

Thus, $n_j=|\alpha_j|^2$ is the probability to be in the momentum state $|\boldsymbol{k}_j\rangle$. The mode functions

$$\left\langle\boldsymbol{r}\middle|\boldsymbol{k}_j\right\rangle=\frac{1}{\sqrt{V}}e^{\mathrm{i}\boldsymbol{k}_j\boldsymbol{r}} \tag{9.17}$$

are normalized in a cuboid with lengths (L_1,L_2,L_3) and a volume $V=L_1L_2L_3$. For periodic boundary conditions, the wave-numbers $k_j=2\pi\kappa_j/L_j$ are quantized with $\kappa_j\in\mathbb{Z}$ and the plane wave states are orthonormal $\langle\boldsymbol{k}_i|\boldsymbol{k}_j\rangle=\delta_{ij}$.

Four-Wave Mixing Conditions

The conditions for FWM are momentum and energy conservation [31]

$$\boldsymbol{p}_1 + \boldsymbol{p}_2 = \boldsymbol{p}_3 + \boldsymbol{p}_4, \qquad E_1 + E_2 = E_3 + E_4. \tag{9.18}$$

In the plane wave ansatz, these conditions can be written as

$$\boldsymbol{k}_1 + \boldsymbol{k}_2 = \boldsymbol{k}_3 + \boldsymbol{k}_4, \qquad \omega_1 + \omega_2 = \omega_3 + \omega_4, \tag{9.19}$$

which are identical to the conditions derived in the optical case (9.10), however now under the consideration of the dispersion relation of free massive particles

$$\omega_j = \omega(\boldsymbol{k}_j) = \frac{\hbar |\boldsymbol{k}_j|^2}{2m}. \tag{9.20}$$

Hence, the frequency scales quadratically with the wave number in contrast to the linear relation in optics (9.11).

To satisfy the conditions (9.19), three wave vectors, e.g. $\boldsymbol{k}_1$, $\boldsymbol{k}_2$ and $\boldsymbol{k}_3$, can be chosen arbitrarily, while the fourth vector follows as the linear superposition $\boldsymbol{k}_4 = \boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{k}_3$. For visualization, the wave vectors can be chosen to lie on a two-dimensional plane in three-dimensional space as the dynamics of the system will not be influenced by the orientation of the vectors, as long as the conditions (9.19) are satisfied. Possible constellations of the wave vectors of a FWM setup in 2D can be seen in Fig. 9.2.

Due to the quadratic nature of the dispersion relation (9.20), the conditions (9.19) cannot be satisfied by wave vectors from one dimension, as energy and momentum conservation cannot be fulfilled simultaneously. However, in Section 9.2.3 we show that one-dimensional FWM can be realized by using the internal energy structure of neutral atoms.

Experimentally, momentum states fulfilling the FWM conditions can be prepared using atomic beamsplitters based on Bragg diffraction [200–202]. The developed initialization sequence to avoid population outside of the FWM states can be seen in Appendix C.

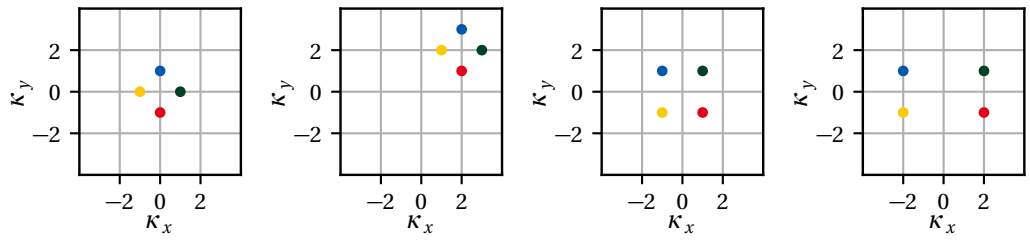

Figure 9.2: Possible configurations of wave vectors $\boldsymbol{k}_1$ (green), $\boldsymbol{k}_2$ (yellow), $\boldsymbol{k}_3$ (blue) and $\boldsymbol{k}_4$ (red), where $k_j = 2\pi\kappa_j / L_j$, satisfying the FWM conditions (9.19). Due to the FWM momentum condition, the wave vectors participating in FWM can be visualized on a two-dimensional plane in three-dimensional space.

Euler-Lagrange Dynamics

In the ideal FWM scenario, the residual wave is absent, $\beta_l = 0$. Consequently, $\sum_{j=1}^{4} n_j = 1$. The physical Lagrangian functional of the system can be determined by inserting the ansatz (9.14) into (9.12), while ensuring (9.19). This yields

$$\begin{aligned} L =& \mathrm{i}\hbar \sum_{j=1}^{4} \alpha_j^* \partial_t \alpha_j - N\hbar \sum_{j=1}^{4} \omega_j n_j + \frac{gN^2}{2V} \sum_{j=1}^{4} n_j^2 - 2\frac{gN^2}{V}(\alpha_1^* \alpha_2^* \alpha_3 \alpha_4 + \text{c.c.}) \\ &+ \frac{2gN^2}{V}(n_1 n_2 + n_1 n_3 + n_1 n_4 + n_2 n_3 + n_2 n_4 + n_3 n_4). \end{aligned} \tag{9.21}$$

Accordingly, the Euler-Lagrange equations, describing the dynamics of the FWM amplitudes α_j, are given by [30, 194, 195]

$$\begin{aligned} \mathrm{i}\hbar \partial_t \alpha_1 =& \hbar\omega_1 \alpha_1 + \frac{gN}{V}\left[(n_1 + 2(n_2 + n_3 + n_4))\alpha_1 + 2\alpha_2^* \alpha_3 \alpha_4\right], \\ \mathrm{i}\hbar \partial_t \alpha_2 =& \hbar\omega_2 \alpha_2 + \frac{gN}{V}\left[(n_2 + 2(n_1 + n_3 + n_4))\alpha_2 + 2\alpha_1^* \alpha_3 \alpha_4\right], \\ \mathrm{i}\hbar \partial_t \alpha_3 =& \hbar\omega_3 \alpha_3 + \frac{gN}{V}\left[(n_3 + 2(n_1 + n_2 + n_4))\alpha_3 + 2\alpha_4^* \alpha_1 \alpha_2\right], \\ \mathrm{i}\hbar \partial_t \alpha_4 =& \hbar\omega_4 \alpha_4 + \frac{gN}{V}\left[(n_4 + 2(n_1 + n_2 + n_3))\alpha_4 + 2\alpha_3^* \alpha_1 \alpha_2\right]. \end{aligned} \tag{9.22}$$

Dimensionless Description

It is convenient to switch to a dimensionless description of the problem. For this, a dimensionless time is introduced as $\tau = \gamma t$, with the frequency

$$\gamma = \frac{gN}{\hbar V}. \tag{9.23}$$

The self-frequencies are also scaled with this constant, yielding $\bar{\omega}_j = \omega_j/\gamma$. Furthermore, the Lagrangian function is shifted by a constant and scaled according to $\mathcal{L} = 1 + VL/gN^2$. Thus, the mathematical Lagrangian functional reads

$$\mathcal{L} = \sum_{j=1}^{4} \mathrm{i}\alpha_j^* \dot{\alpha}_j - \mathcal{E}, \tag{9.24}$$

with the dimensionless energy of the system

$$\mathcal{E} = \sum_{j=1}^{4} \varepsilon_j + 2(\alpha_1^* \alpha_2^* \alpha_3 \alpha_4 + \text{c.c.}). \tag{9.25}$$

There, $\dot{\alpha}_j$ denotes the partial derivative with respect to dimensionless time $\partial_\tau \alpha_j$ and the mean-field shifted single particle energies are defined as

$$\varepsilon_j = \bar{\omega}_j n_j - \frac{n_j^2}{2}, \qquad \mu_j = \frac{\partial \varepsilon_j}{\partial n_j} = \bar{\omega}_j - n_j. \tag{9.26}$$

Accordingly, a compact form of the equations of motion for the complex amplitudes is given by

$$\begin{aligned} \mathrm{i}\dot{\alpha}_1 &= \mu_1\alpha_1 + 2\alpha_2^*\alpha_3\alpha_4, \\ \mathrm{i}\dot{\alpha}_2 &= \mu_2\alpha_2 + 2\alpha_1^*\alpha_3\alpha_4, \\ \mathrm{i}\dot{\alpha}_3 &= \mu_3\alpha_3 + 2\alpha_4^*\alpha_1\alpha_2, \\ \mathrm{i}\dot{\alpha}_4 &= \mu_4\alpha_4 + 2\alpha_3^*\alpha_1\alpha_2. \end{aligned} \tag{9.27}$$

Clearly, these equations are highly symmetrical, which can be explored using the polar decomposition of the complex amplitudes (9.16). As can be seen in Fig. 9.3 (a), the relative particle numbers n_j perform periodic oscillations. Due to the source terms in the equations of motion (9.27), the initially unpopulated fourth momentum state sees an increase in population during the oscillation. Furthermore, the simulations reveal, that the total particle number, as well as the energy of the system, are conserved.

The interaction term in (9.25) coherently couples the subspaces $\{|\boldsymbol{k}_1\rangle, |\boldsymbol{k}_2\rangle\} \longleftrightarrow \{|\boldsymbol{k}_3\rangle, |\boldsymbol{k}_4\rangle\}$ through the relative phase difference

$$\phi = \varphi_1 + \varphi_2 - \varphi_3 - \varphi_4 \tag{9.28}$$

and the population imbalance

$$m = n_1 + n_2 - n_3 - n_4. \tag{9.29}$$

Therefore, the oscillatory dynamics of the system can be investigated in m-ϕ-phase-space. Depending on the choice of the quantities

$$m_{12} = n_1 - n_2, \qquad m_{34} = n_3 - n_4, \tag{9.30}$$

periodic as well as aperiodic orbits of the mathematical pendulum with a separatrix in between occur. For $m_{12} = m_{34} = 0$ (see Fig. 9.3 (c)) only closed trajectories can be observed. The covered area in phase space depends on $m_0 = m(t = 0)$. Choosing instead $m_{12} = 0.4$ and $m_{34} = 0.02$, also transient trajectories appear. Furthermore, the possible value range becomes limited, depending on the choice of m_{12} and m_{34}.

To obtain an analytical description of the FWM process in addition to numerical solutions, a coordinate transformation is performed. This will reveal the possibility to describe the FWM dynamics as Josephson oscillations.

9.2.2 Josephson Description of Four-Wave Mixing Amplitudes

Coordinate Transformation

With the help of Lagrangian field theory, the time-independent Hamiltonian energy (9.12) and the FWM state ansatz (9.14), we obtain a discrete nonlinear set of four Hamiltonian equations with a number of symmetries. This constrains the dynamics to a two-dimensional phase-space, analogous to the mathematical pendulum. Due to the phase-invariant structure of the self-energy $g\,n^2$, typical Josephson oscillations emerge. This analogy enables the search for an analytical description of the FWM oscillations. We introduce new coordinates by

$$\begin{aligned} \alpha_1 &= \sqrt{n_1}e^{-\mathrm{i}(\Phi+\phi/4+\varphi)}, \quad \alpha_2 = \sqrt{n_2}e^{-\mathrm{i}(\Phi+\phi/4-\varphi)}, \\ \alpha_3 &= \sqrt{n_3}e^{-\mathrm{i}(\Phi-\phi/4+\theta)}, \quad \alpha_4 = \sqrt{n_4}e^{-\mathrm{i}(\Phi-\phi/4-\theta)}. \end{aligned} \tag{9.31}$$

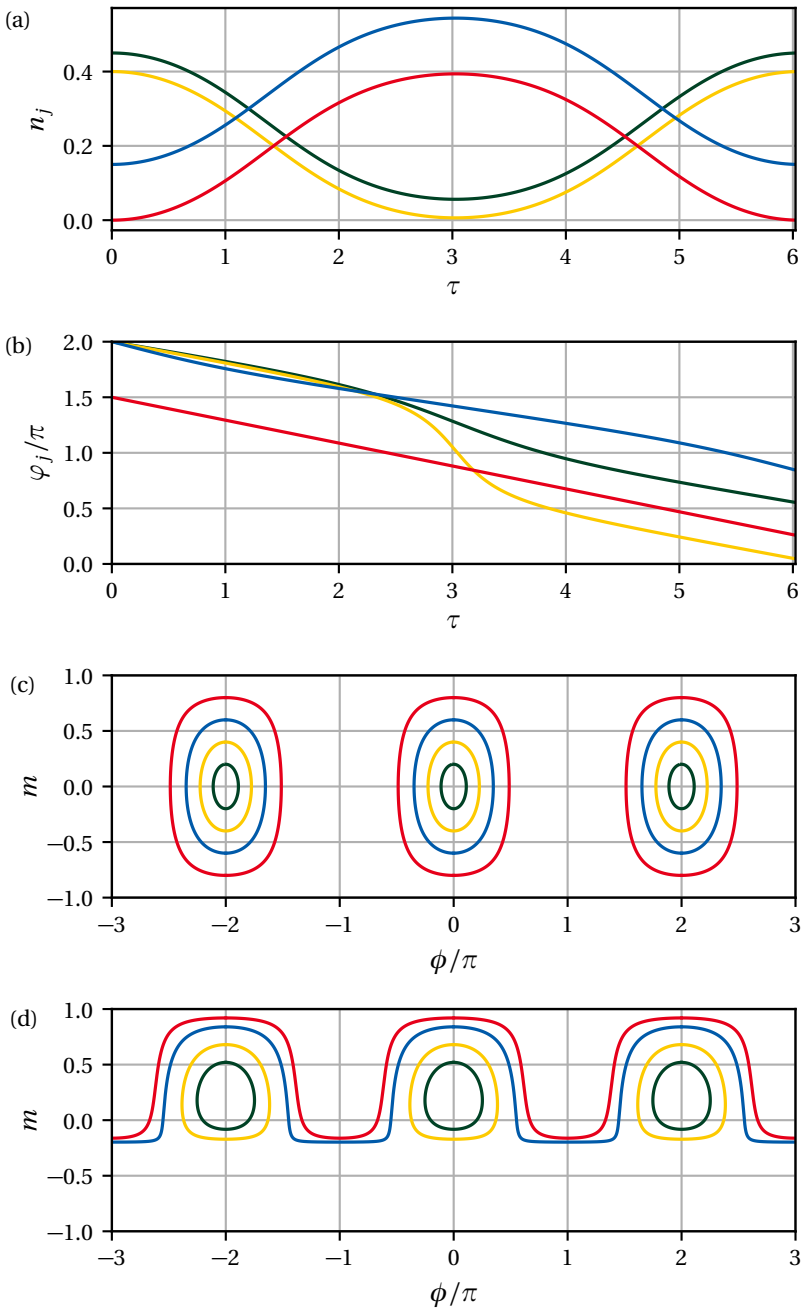

Figure 9.3: Dynamics of FWM amplitudes in terms of (a) relative particle numbers n_j and (b) phases φ_j versus dimensionless time τ for momentum components $|\boldsymbol{k}_1\rangle$ (green), $|\boldsymbol{k}_2\rangle$ (yellow), $|\boldsymbol{k}_3\rangle$ (blue) and $|\boldsymbol{k}_4\rangle$ (red) and $\bar{\omega}_j = 1$. (c) & (d) FWM dynamics in m-ϕ-phase-space. For $m_{12} = m_{34} = 0$ (c) only closed trajectories occur, while aperiodic solutions are present for $m_{12} = 0.4$ and $m_{34} = 0.02$ (d). (c) green: $m_0 = 0.2$, yellow: $m_0 = 0.4$, blue: $m_0 = 0.6$, red: $m_0 = 0.8$; (d) green: $m_0 = 0.52$, yellow: $m_0 = 0.68$, blue: $m_0 = 0.84$, red: $m_0 = 0.92$. For all trajectories, $\phi_0 = 0$ was set.

From the global phase invariance of (9.24), one finds that the total occupation in the FWM state $\sum_{j=1}^{4} n_j$ is conserved. This can be used to construct a generating function as

$$R = \frac{\mathrm{i}}{2} e^{2\mathrm{i}\Phi}\Big(\alpha_1^2 e^{2\mathrm{i}(\phi/4+\varphi)} + \alpha_2^2 e^{2\mathrm{i}(\phi/4-\varphi)} + \alpha_3^2 e^{2\mathrm{i}(-\phi/4+\theta)} + \alpha_4^2 e^{2\mathrm{i}(-\phi/4-\theta)}\Big). \tag{9.32}$$

According to the rules of Hamiltonian mechanics [203], this generating function relates old coordinates $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ to new coordinates $(\Phi, \phi, \varphi, \theta)$. In turn, one can obtain the old momenta

$$\pi_j = \frac{\partial R}{\partial \alpha_j} = \mathrm{i}\alpha_j^*, \tag{9.33}$$

as well as the new momenta

$$P_\Phi = -\frac{\partial R}{\partial \Phi} = n_1 + n_2 + n_3 + n_4, \tag{9.34}$$

$$P_\phi = -\frac{\partial R}{\partial \phi} = \frac{n_1 + n_2 - n_3 - n_4}{4} \equiv \frac{m}{4}, \tag{9.35}$$

$$P_\varphi = -\frac{\partial R}{\partial \varphi} = n_1 - n_2 \equiv m_{12}, \tag{9.36}$$

$$P_\theta = -\frac{\partial R}{\partial \theta} = n_3 - n_4 \equiv m_{34}. \tag{9.37}$$

In terms of the new coordinates the dimensionless Lagrangian $\mathcal{L}$ reads

$$\mathcal{L} = \dot{\Phi} + \frac{m}{4}\dot{\phi} + m_{12}\dot{\varphi} + m_{34}\dot{\theta} - H(m, \phi), \tag{9.38}$$

with a generic Josephson Hamiltonian energy

$$H(m, \phi) = \frac{\eta}{4}\cos\phi - \frac{m^2}{8} + \mathcal{C}, \tag{9.39}$$

$$\eta = \sqrt{\left[(1+m)^2 - 4m_{12}^2\right]\left[(1-m)^2 - 4m_{34}^2\right]}. \tag{9.40}$$

Here, we have denoted an energy offset

$$\mathcal{C} = \sum_{j=1}^{4} \frac{\bar{\omega}_j}{4} + \frac{\bar{\omega}_{12} m_{12} + \bar{\omega}_{34} m_{34}}{2} + \frac{m_{12}^2 + m_{34}^2}{4} - \frac{7}{8}. \tag{9.41}$$

and transition energies $\bar{\omega}_{12} = \bar{\omega}_1 - \bar{\omega}_2$ and $\bar{\omega}_{34} = \bar{\omega}_3 - \bar{\omega}_4$. As $\mathcal{L}$ does not depend on Φ, φ or θ, these phases are cyclic [204]. Therefore, the conjugate momenta, total particle number N and population differences m_{12} (9.36) and m_{34} (9.37) are conserved. Consequently, the equations of motion for Φ, φ and θ can be solved in quadrature.

As can be seen by calculating the Legendre transform of (9.38), $H(m, \phi)$ (9.39) is the Hamiltonian of the system and m and ϕ are the canonical variables of the system. Accordingly, the dynamics of the system are given by the two coupled Josephson-like differential equations [33–35]

$$\dot{m} = -4\partial_\phi H = \eta \sin\phi, \qquad \dot{\phi} = 4\partial_m H = -m + \cos\phi\, \partial_m \eta. \tag{9.42}$$

Comparison to Josephson effect in superconductors

The Josephson effect was predicted theoretically by *Brian Josephson* in 1962 [205]. He considered a thin insulating layer separating two superconductors, predicting the existence of a tunnel current $I(t)$ carried by Cooper pairs through the junction. The two basic Josephson equations for such a junction are given by [206]

$$I(t) = I_c \sin\varphi, \qquad \dot{\varphi} = \frac{2eV}{\hbar}, \tag{9.43}$$

where I_c is the maximum critical current, $\varphi = \varphi_1 - \varphi_2$ is the phase difference between the two superconducting electrodes and V is the voltage across the junction.

In the absence of an external voltage $V = 0$, a direct current across the insulator can be observed due to tunnelling. This current is constant as it is proportional to the sine of the Josephson phase φ, which itself is constant. This is called the DC Josephson effect.

For a finite voltage $V \neq 0$, the Josephson phase varies in time according to

$$\varphi(t) = \varphi_0 + \frac{2eV}{\hbar} t \tag{9.44}$$

yielding an alternating current through the junction

$$I(t) = I_c \sin\big(\varphi(t)\big). \tag{9.45}$$

This is called the AC Josephson effect. Also, an inverse effect can be observed [207]. In the presence of microwave radiation with frequency ω, current steps at constant voltages

$$V_n = \frac{n\hbar}{2e}\omega, \tag{9.46}$$

where $n = \pm 1, \pm 2, \ldots$, appear. These steps are known as Shapiro steps.

Comparing (9.42) and (9.43), there are similarities. However, the nonlinear nature of (9.42) will lead to more complex expressions than e.g. is the case for the AC Josephson effect.

General Solution

In simple classical mechanics problems of particles, with position x and momentum p, Hamiltonian energies $H(x,p) = T(p) + V(x)$ separate into kinetic $T(p)$ and potential $V(x)$ energy. At the turning points $\dot{x} = \partial_p H = 0$, the Hamilton function is purely determined by potential energy $H(x, p=0) = V(x)$. A similar investigation can be performed in the given case [208–210]. Through canonical transformation, the role of position and momentum variables can be exchanged. Therefore, m can be considered as the position and ϕ as the momentum variable. Thus, at the turning points $\dot{m} = -4\partial_\phi H = 0$, two momenta

$$\phi^+ = 0, \qquad \phi^- = \pi \tag{9.47}$$

are possible. In turn, this defines two potentials

$$V^\pm(m) = H(m, \phi^\pm) = -\frac{1}{8}(m^2 \mp 2\eta) + \mathcal{C}. \tag{9.48}$$

Visualizations of such potentials for varying system parameters can be seen in Fig. 9.4.

Physical solutions with energies $\varepsilon = H(m,\phi)$ must be constrained by these two potentials, $V^- < \varepsilon < V^+$. This limits the value range of m and ϕ depending on the system parameters

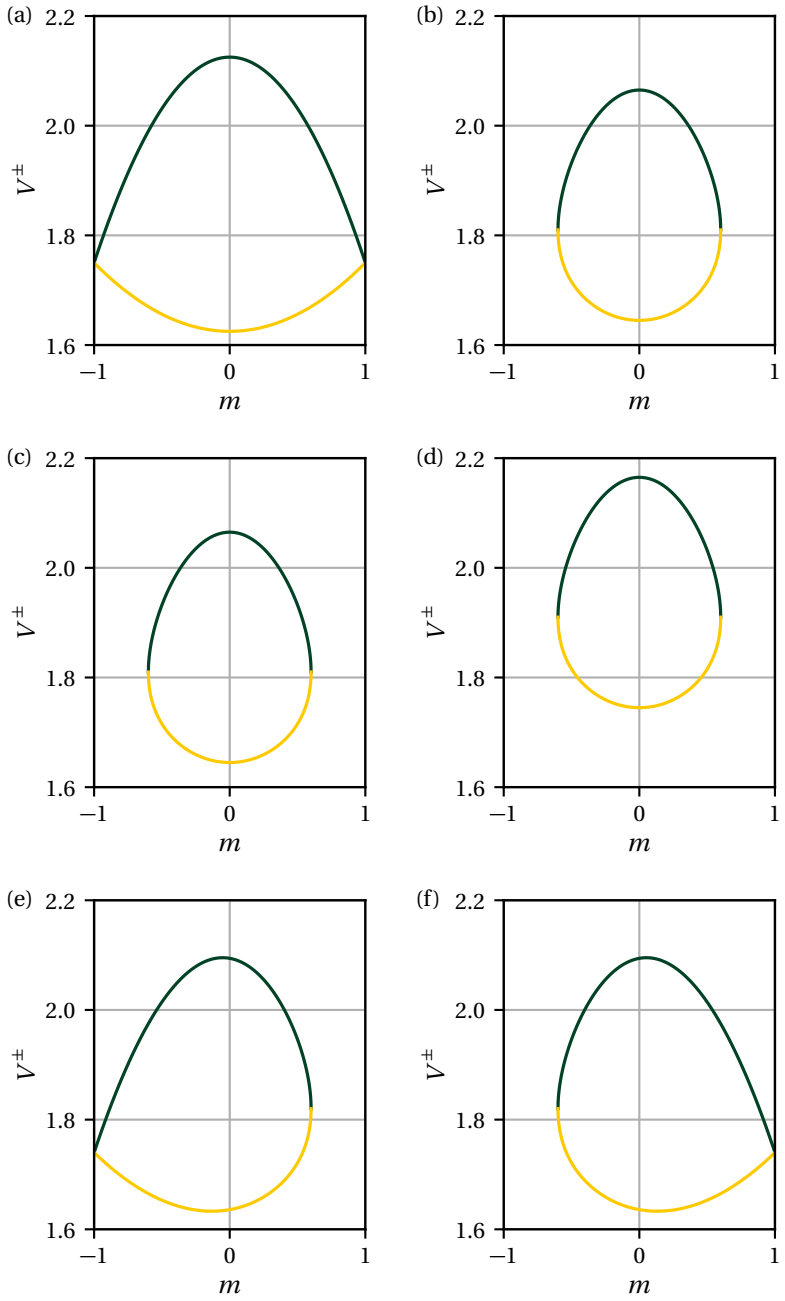

Figure 9.4: Potentials V^+ (green) and V^- (yellow) versus population difference m (9.48). (a) For $m_{12} = m_{34} = 0$, m has a possible value range from −1 to 1 governed by $V^- < \varepsilon < V^+$. (b) The value range of m is reduced for differing values ($m_{12} = m_{34} = 0.2$). (c) The potentials remain unchanged when changing the sign $m_{12} = m_{34} = -0.2$. (d) Altering the values of $\bar{\omega}_j$ (here $\bar{\omega}_j = 1.1$; all other subfigures $\bar{\omega}_j = 1$) shifts the potentials along the $V^{\pm}$-axis. (e) Choosing different values for m_{12} and m_{34} ($m_{12} = 0$, $m_{34} = 0.2$) lifts the symmetry of the potentials around the m-axis. (f) Changing the values of m_{12} and m_{34} ($m_{12} = 0.2$, $m_{34} = 0$) mirrors the potentials on the m-axis.

m_{12} and m_{34} (see Fig. 9.4). As the energy of the system is conserved, the equation of motion (9.42) for $m(\tau)$ can be expressed using the potentials $V^{\pm}$ as

$$\dot{m} = \pm 4\sqrt{(V^{+}(m)-\varepsilon)(\varepsilon - V^{-}(m))}. \tag{9.49}$$

Accordingly, the dynamical solution can be calculated as

$$\tau - \tau_0 = \int_{m_0}^{m} \frac{\pm \mathrm{d}\zeta}{4\sqrt{(V^{+}(\zeta)-\varepsilon)(\varepsilon - V^{-}(\zeta))}}. \tag{9.50}$$

This relation can be inverted piecewise to obtain $m(\tau)$.

Analytical solution for $m_{12} = m_{34} = 0$

For the special case $m_{12} = m_{34} = 0$, implying $n_1 = n_2$ and $n_3 = n_4$, an analytical expression for the dynamical solution $m(\tau)$ can be given in terms of the elliptic cosine $\mathrm{cn}(u, v)$ [211] as

$$m(\tau) = \pm\sqrt{\frac{\mu+2}{3}}\,\mathrm{cn}\big(\xi(\tau-\tau_0), \rho^2\big), \tag{9.51}$$

where $\mu = m_0^2 + 2(m_0^2-1)\cos\phi_0$, $\xi = \sqrt{6-3\mu}/2$ and $\rho^2 = (\mu+2)/(6-3\mu)$. With that, the dynamical solution of the phase $\phi(\tau)$ can be calculated by integration of (9.42), yielding

$$\begin{aligned}\phi(\tau) = 2\arctan\Big\{\sqrt{3}\tanh\Big[&\operatorname{arctanh}\big(\tan(\phi_0/2)/\sqrt{3}\big) + \ln\big(1-\rho\big) \\ &- \ln\big(\mathrm{dn}\big(\xi(\tau-\tau_0), \rho^2\big) - \rho\,\mathrm{cn}\big(\xi(\tau-\tau_0), \rho^2\big)\big)\Big]\Big\},\end{aligned} \tag{9.52}$$

with the delta amplitude $\mathrm{dn}(u, v)$ [211]. Visualizations of the Jacobian elliptic functions can be seen in Fig. 9.5.

The period of the motion can be calculated as

$$T = \frac{4\,\mathrm{K}(\rho^2)}{\xi}. \tag{9.53}$$

There, K is the complete elliptic integral of first kind [211]. The basic frequency of the oscillation $T_0 = T(m_0 = 0)$ can be calculated as

$$T_0 = \frac{4\pi}{\sqrt{12}}. \tag{9.54}$$

The analytical solutions for $m(\tau)$ and $\phi(\tau)$, as well as the dependence of T on the parameter m_0 can be seen in Fig. 9.6. The period of the FWM oscillation diverges when nearing the regime of aperiodic solutions.

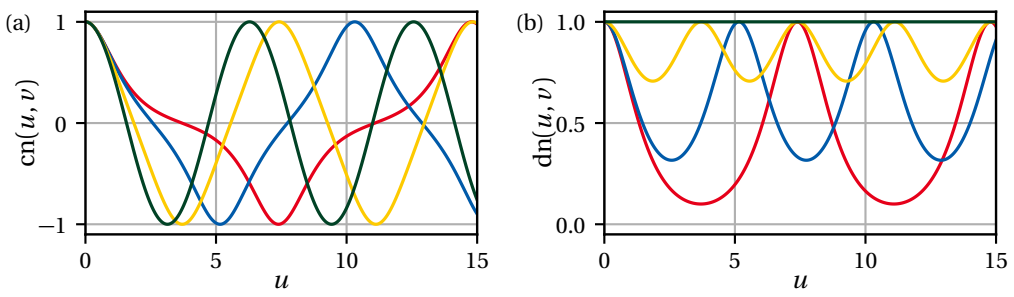

Figure 9.5: Visualizations of the Jacobian elliptic functions (a) $\mathrm{cn}(u, v)$ and (b) $\mathrm{dn}(u, v)$ versus real variable u for $v = 0$ (green), $v = 0.5$ (yellow), $v = 0.9$ (blue) and $v = 0.99$ (red).

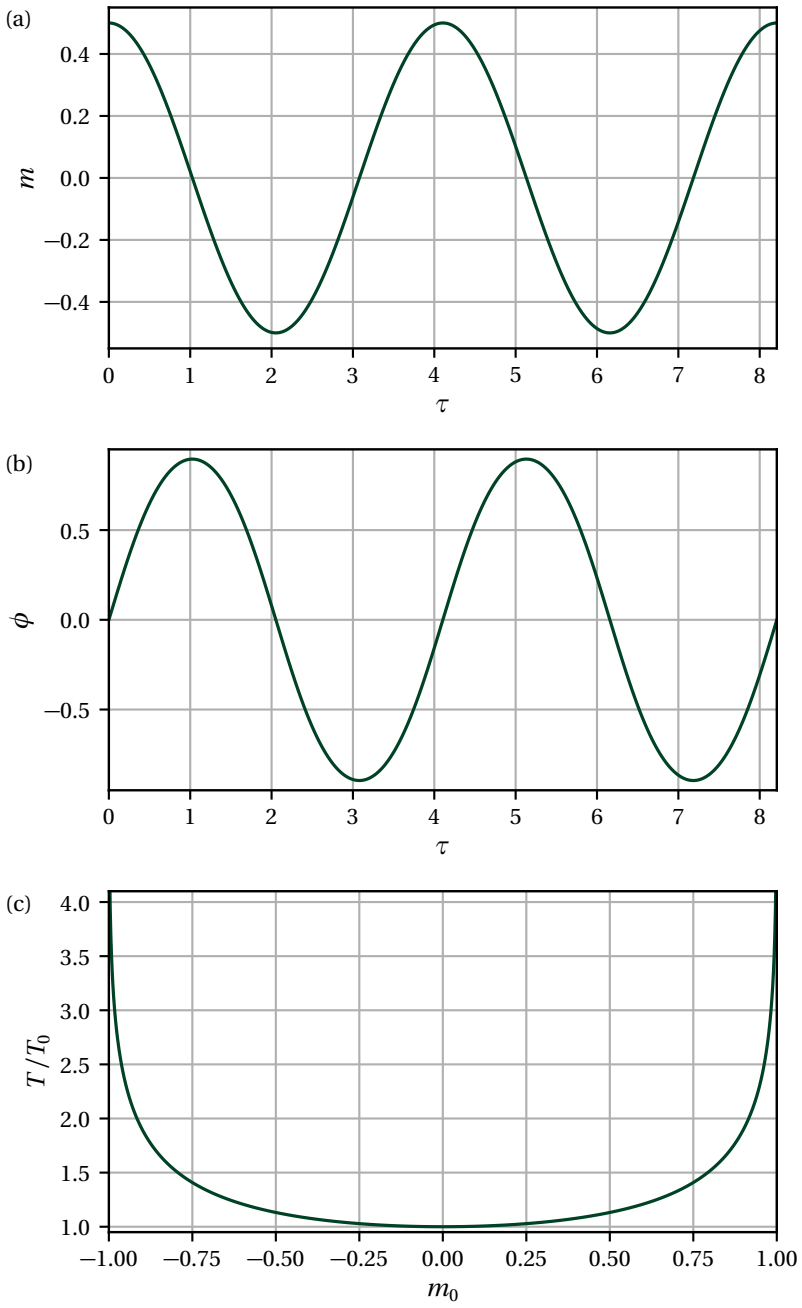

Figure 9.6: Analytical solution of FWM dynamics for $m_{12} = m_{34} = 0$. (a) Population difference m (9.51) and (b) phase difference ϕ (9.52) versus dimensionless time τ for $m_0 = 0.5$ and $\phi_0 = 0$. (c) Period of the FWM oscillation T versus initial population difference m_0 (9.53), normed to $T_0 = T(m_0)$. The period diverges when nearing the regime of transient trajectories in m-ϕ-phase-space.

9.2.3 One-Dimensional Four-Wave Mixing with Multiple Internal States

Lower-dimensional systems are of high interest in the investigation of BECs. A good platform to investigate phenomena in one-dimensional systems is the realization of trapped atoms in ring-shaped optical lattices [212]. Such a realization yields a one-dimensional implementation with periodic boundary conditions very close to common theoretical descriptions.

Using BECs in ring traps, a variety of interesting problems has been investigated including persistent flow and vortices [165, 213], realizations of the Bose-Hubbard model [214] and entanglement between two states with different circulation [215]. Also, the rotational response of two-component BECs in ring traps has been the subject of investigations [216].

We aim to find a realization of FWM with plane waves in one dimension, circumventing the restriction brought on by energy and momentum conservation (9.19) due to the dispersion relation (9.20). For this, we consider a one-dimensional trap with periodic boundary conditions, i.e. a ring trap, populated by a homogeneous two-component BEC, similar to [198]. We show that by introducing internal structure of the trapped atoms, FWM can indeed be realized in one dimension.

Two-Component Gross-Pitaevskii Equation

The onset point for the discussion of a BEC with two internal states $|1\rangle$ and $|2\rangle$ is the two-component GPE. This is given by the coupled equations for the order parameters $\psi^{(i)}(\boldsymbol{r},t)$, with $i\in\{1,2\}$, [198]

$$\mathrm{i}\hbar\partial_t\psi^{(i)}=\left(-\frac{\hbar^2}{2m}\nabla^2+U^{(i)}+\sum_{j\in\{1,2\}}g_{ij}n^{(j)}\right)\psi^{(i)}, \tag{9.55}$$

where $n^{(i)}=\left|\psi^{(i)}\right|^2$, $g_{ij}=4\pi\hbar^2a_{ij}/m$ and a_{11}, a_{22} and a_{12} are the intra- and inter-state scattering lengths, respectively. As before, we assume a shallow lattice limit, implying $U^{(i)}=0$. The complete set of coupled equations can be written in matrix form as [217]

$$\mathrm{i}\hbar\partial_t\begin{pmatrix}\psi^{(1)}\\ \psi^{(2)}\end{pmatrix}=\begin{pmatrix}-\frac{\hbar^2}{2m}\nabla^2+g_{11}n^{(1)}+g_{12}n^{(2)} & 0\\ 0 & -\frac{\hbar^2}{2m}\nabla^2+g_{22}n^{(2)}+g_{12}n^{(1)}\end{pmatrix}\begin{pmatrix}\psi^{(1)}\\ \psi^{(2)}\end{pmatrix}. \tag{9.56}$$

^{87}*Rb Hyperfine States*

As the two internal states, interacting according to (9.56), we consider the $|1\rangle=|F=1,m_F=-1\rangle$ and $|2\rangle=|F=2,m_F=1\rangle$ hyperfine states of the $5^2S_{1/2}$ ^{87}Rb ground state. The energy splitting due to the hyperfine structure, resulting from a coupling of the total electron angular momentum $\boldsymbol{J}=\boldsymbol{L}+\boldsymbol{S}$, where $\boldsymbol{L}$ is the orbital angular momentum of the outer electron and $\boldsymbol{S}$ is its spin angular momentum, with the total nuclear angular momentum $\boldsymbol{I}$, for the ground state is given by [218]

$$\Delta E_{\mathrm{hfs}}=\frac{1}{2}A_{\mathrm{hfs}}[F(F+1)-I(I+1)-J(J+1)]. \tag{9.57}$$

There, A_{hfs} is the magnetic dipole constant and $|J-I|\leq F\leq J+I$. For the $5^2S_{1/2}$ ^{87}Rb ground state, $L=0$ and $S=1/2$. Therefore, $J=1/2$ and additionally $I=3/2$. Hence, possible values for F are 1 and 2.

Each of the F hyperfine levels contains $2F+1$ magnetic sublevels, which are degenerate in the absence of an external magnetic field. If a weak external field B_z is applied, for instance along the z-direction, the hyperfine levels split up linearly according to the anomalous Zeeman effect [218]

$$\Delta E_{|F,m_F\rangle} = \mu_\mathrm{B} g_F m_F B_z. \tag{9.58}$$

There, μ_B is the Bohr magneton, $m_F \in [-F, F]$ is the magnetic quantum number and g_F is the hyperfine Landé g-factor [219]

$$g_F \simeq g_J \frac{F(F+1)-I(I+1)+J(J+1)}{2F(F+1)}, \tag{9.59}$$

with the Landé factor

$$g_J \simeq 1 + \frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)}. \tag{9.60}$$

A visualization of the energy splitting due to a weak external magnetic field of the $F=1$ and $F=2$ hyperfine states of the $5^2S_{1/2}$ ^{87}Rb ground state can be seen in Fig. 9.7 (a). Accordingly, the states $|1\rangle$ and $|2\rangle$ show a constant frequency separation ω_0 in presence of such a field.

An advantage in considering those specific hyperfine states lies in the characteristic that the intra- and inter-state scattering lengths are almost identical [220]

$$a_{12} = (5.5 \pm 0.3)\,\mathrm{nm}, \qquad a_{11} = 1.03 a_{12}, \qquad a_{22} = 0.97 a_{12}. \tag{9.61}$$

For our purpose, this justifies the approximation of setting

$$g = g_{11} \simeq g_{22} \simeq g_{12}. \tag{9.62}$$

With that, the coupled GPEs, describing the dynamics of the system, can be written as

$$\mathrm{i}\hbar\partial_t \begin{pmatrix} \psi^{(1)} \\ \psi^{(2)} \end{pmatrix} = \begin{pmatrix} -\frac{\hbar^2}{2m}\nabla^2 + gn & 0 \\ 0 & -\frac{\hbar^2}{2m}\nabla^2 + gn \end{pmatrix} \begin{pmatrix} \psi^{(1)} \\ \psi^{(2)} \end{pmatrix}, \tag{9.63}$$

where $n = n^{(1)} + n^{(2)}$.

Four-Wave Mixing Conditions

Similar to the previous discussion, we introduce a plane wave ansatz for the FWM state $|\psi_\alpha(t)\rangle$. However, in this new geometry with multiple internal levels, the momentum components constituting this state are sets of two in the respective internal states. Accordingly, the FWM state is given by the coherent superposition

$$\left|\psi_\alpha(t)\right\rangle = \sum_{i=1}^{2}\sum_{j=1}^{2} \sqrt{N}\alpha_j^{(i)}(t)\left|\boldsymbol{k}_j, i\right\rangle, \tag{9.64}$$

where i denotes the internal state, j denotes the momentum state, N is the total particle number and $\alpha_j^{(i)}(t)$ are the time-dependent complex amplitudes of the plane waves $|\boldsymbol{k}_j\rangle$ in internal state $|i\rangle$.

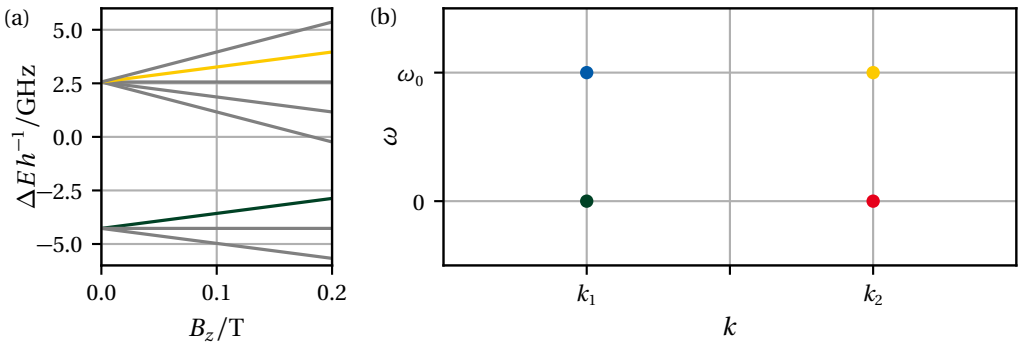

Figure 9.7: (a) Energy splitting due to the anomalous Zeeman effect of the $F = 1$ (lower) and $F = 2$ (upper) hyperfine states of the $5^2S_{1/2}$ ^{87}Rb ground state. In the presence of a weak external magnetic field B_z, the $|F = 1, m_F = -1\rangle$ (green) and $|F = 2, m_F = 1\rangle$ (yellow) states show a constant separation of ω_0 in terms of frequency. (b) Suggested FWM setup to satisfy the FWM conditions (9.18). By populating two internal states at the same momentum states $|\boldsymbol{k}_1\rangle$ and $|\boldsymbol{k}_2\rangle$, the FWM conditions can be fulfilled according to (9.66).

In this scenario, a setup can be found, which satisfies the general FWM conditions (9.18). This is given by

$$\boldsymbol{p}_1^{(1)} + \boldsymbol{p}_2^{(2)} = \boldsymbol{p}_2^{(1)} + \boldsymbol{p}_1^{(2)}, \qquad E_1^{(1)} + E_2^{(2)} = E_2^{(1)} + E_1^{(2)}. \tag{9.65}$$

A visualization of the suggested geometry can be seen in Fig. 9.7 (b). The problem of not being able to satisfy energy and momentum conservation simultaneously in the original setup is circumvented by using equal momenta in the two internal states. Thereby, momentum and energy conservation are fulfilled by pairing opposite momenta in the different internal states together. Accordingly, the conditions (9.18) are satisfied by

$$\boldsymbol{k}_1 + \boldsymbol{k}_2 = \boldsymbol{k}_2 + \boldsymbol{k}_1, \qquad \omega_1 + \omega_2 + \omega_0 = \omega_1 + \omega_0 + \omega_2, \tag{9.66}$$

where the dispersion relation (9.20) still holds.

Euler-Lagrange Dynamics

Combining the coupled two-component GPEs (9.56), the plane wave ansatz (9.64) and the suggested geometry, the Lagrangian of the system can be determined as

$$\begin{aligned} L = {} & \mathrm{i}\hbar \sum_{i,j} \left(\alpha_j^{(i)}\right)^* \partial_t \alpha_j^{(i)} - N\hbar \sum_{i,j} \omega_j n_j^{(i)} + \frac{gN^2}{2V} \sum_{i,j} \left(n_j^{(i)}\right)^2 - \frac{2gN^2}{V} \left(\left(\alpha_1^{(1)}\right)^* \left(\alpha_2^{(2)}\right)^* \alpha_2^{(1)} \alpha_1^{(2)} + \text{c.c.}\right) \\ & + \frac{2gN^2}{V} \left(n_1^{(1)} n_2^{(1)} + n_1^{(1)} n_1^{(2)} + n_1^{(1)} n_2^{(2)} + n_2^{(1)} n_1^{(2)} + n_2^{(1)} n_2^{(2)} + n_1^{(2)} n_2^{(2)}\right). \end{aligned} \tag{9.67}$$

Accordingly, the dynamics of the FWM amplitudes are given by the Euler-Lagrange equations

$$\begin{aligned}
\mathrm{i}\hbar\partial_t \alpha_1^{(1)} &= \hbar\omega_1 \alpha_1^{(1)} + \frac{gN}{V}\left[\left(n_1^{(1)} + 2\left(n_2^{(1)} + n_1^{(2)} + n_2^{(2)}\right)\right)\alpha_1^{(1)} + 2\left(\alpha_2^{(2)}\right)^* \alpha_2^{(1)}\alpha_1^{(2)}\right], \\
\mathrm{i}\hbar\partial_t \alpha_2^{(1)} &= \hbar\omega_2 \alpha_2^{(1)} + \frac{gN}{V}\left[\left(n_2^{(1)} + 2\left(n_1^{(1)} + n_1^{(2)} + n_2^{(2)}\right)\right)\alpha_2^{(1)} + 2\left(\alpha_1^{(2)}\right)^* \alpha_1^{(1)}\alpha_2^{(2)}\right], \\
\mathrm{i}\hbar\partial_t \alpha_1^{(2)} &= \hbar\omega_1 \alpha_1^{(2)} + \frac{gN}{V}\left[\left(n_1^{(2)} + 2\left(n_1^{(1)} + n_2^{(1)} + n_2^{(2)}\right)\right)\alpha_1^{(2)} + 2\left(\alpha_2^{(1)}\right)^* \alpha_1^{(1)}\alpha_2^{(2)}\right], \\
\mathrm{i}\hbar\partial_t \alpha_2^{(2)} &= \hbar\omega_2 \alpha_2^{(2)} + \frac{gN}{V}\left[\left(n_2^{(2)} + 2\left(n_1^{(1)} + n_2^{(1)} + n_1^{(2)}\right)\right)\alpha_2^{(2)} + 2\left(\alpha_1^{(1)}\right)^* \alpha_2^{(1)}\alpha_1^{(2)}\right].
\end{aligned} \tag{9.68}$$

These equations of motion are equivalent to (9.22). This implies, that the complete investigation performed in Subsections 9.2.1 and 9.2.2 can also be applied here. Therefore, introducing an additional internal state, FWM with plane waves can also be realized in one dimension.

9.3 FOUR-WAVE MIXING WITH BACKGROUND POPULATION

In the ideal FWM setting, the residual wave $|\psi_\beta\rangle$ is absent. However, additional momentum states might be populated accidentally during the initialization procedure or system evolution.

9.3.1 Numerical Simulation of Four-Wave Mixing on Discrete Periodic Grid

To investigate this scenario, the dynamics of the system, described by the GPE (8.31), are simulated using a Runge-Kutta scheme and Fast Fourier Transforms (FFT) on a discrete periodic grid. For this, a two-dimensional grid with 16×16 sites is used while setting $\gamma = 1/\mathrm{s}$ and discretizing dimensionless time with $\Delta\tau = 10^{-6}$. For implementation the geometry of FWM states described in Appendix C is chosen, yielding $\bar{\omega}_j = 1$ for $j = 1, \ldots, 4$. The populations are set to $n_1 = n_2 = 0.375$ and $n_3 = n_4 = 0.125$, resulting in $m_{12} = m_{34} = 0$ and $m_0 = 0.5$. All phases are set to $\varphi_j = 0$, yielding $\phi_0 = 0$. Also, multiple internal states can be taken into consideration in the developed simulation method.

As can be seen in Fig. 9.8, the numerical results of the GPE simulation start to deviate noticeably from the four-mode approximation (9.27) already after a few cycles. Looking at

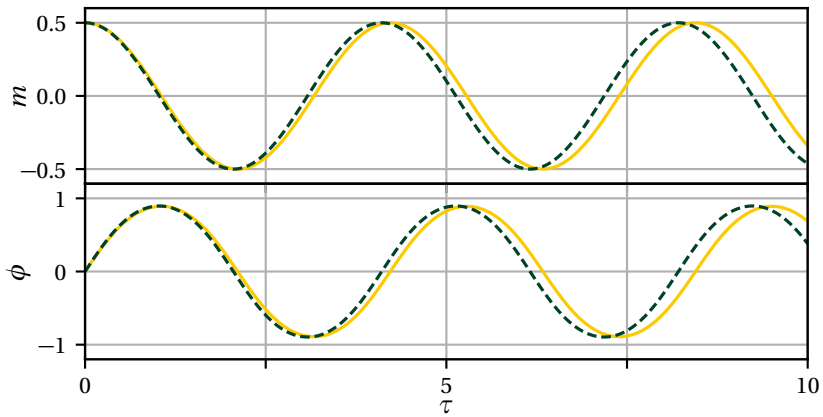

Figure 9.8: Population imbalance $m(\tau)$ and relative phase $\phi(\tau)$ versus dimensionless time τ from analytical solution (green, dashed) and numerical simulation of the GPE on a discrete periodic grid (yellow, solid).

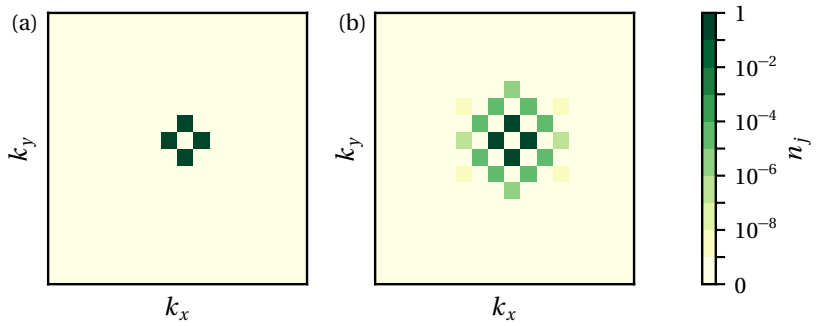

Figure 9.9: Histograms of populations on discrete 16×16 grid in k_x-k_y-plane at $\tau=0$ (a) and $\tau=5.0$ (b).

$m(\tau)$ and $\phi(\tau)$, the numerical results show a larger period of the oscillation. However, the general shape of the oscillations remains unchanged.

9.3.2 Analysis of Background Population

This behaviour is caused by an instability of the simulation due to numerical noise of the FFT producing population on the grid outside of the FWM states. As depicted in Fig. 9.9 (a), the system is prepared at $\tau=0$ with population only present in the FWM states. However, the histogram in Fig. 9.9 (b) at $\tau=5$ clearly shows that additional states in the vicinity of the FWM states have been populated. As this background population is located at the center of the lattice, the chosen grid is large enough such that no edge effects occur during the simulation.

Yet, the instability caused by accidental population of additional momentum states is not destructive in nature. Looking at Fig. 9.10 (a), the total background population

$$n_\mathrm{B}=\sum_{l>4}|\beta_l|^2 \tag{9.69}$$

grows rapidly at the beginning of the oscillation. Subsequently, the dynamics of $n_\mathrm{B}(\tau)$ stabilize and show oscillations with a maximum value of around $n_\mathrm{B}\simeq 5\cdot10^{-4}$. As can be seen in Fig. 9.10 (b), the frequency of the ensuing oscillation is about 50 times larger than the FWM frequency

$$\nu_\mathrm{F}=\frac{1}{T(m_0=0.5)}\simeq 0.244. \tag{9.70}$$

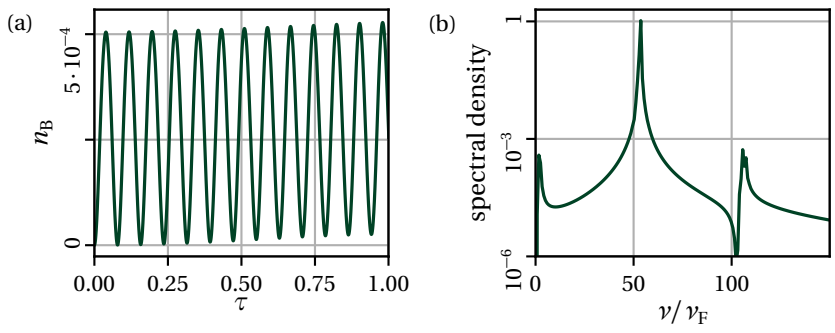

Figure 9.10: (a) Background population n_B starts oscillating and quickly reaches maximum value. (b) Oscillation frequency of n_B is about 50 times bigger than the FWM frequency ν_F.

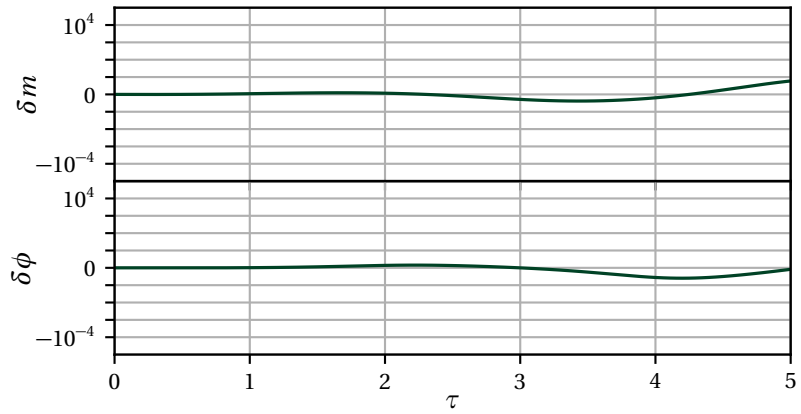

Figure 9.11: Deviations $\delta m = (m_\mathrm{n} - m_\mathrm{a})/\max(m_\mathrm{a})$ and $\delta\phi = (\phi_\mathrm{n} - \phi_\mathrm{a})/\max(\phi_\mathrm{a})$ versus dimensionless time τ from numerical (m_n, ϕ_n) to analytical solutions (m_a, ϕ_a). All population outside of FWM states is eliminated after each simulation step.

The non-negligible background population is the cause of change in the dynamics of the FWM process. Because of

$$n_\mathrm{F} + n_\mathrm{B} = 1, \qquad n_\mathrm{F} = \sum_{j=1}^{4} |\alpha_j|^2, \tag{9.71}$$

growing n_B reduces the population in the FWM states n_F in comparison to the ideal case. As the FWM process is caused by the density-density-interaction terms in the GPE (8.31), even small changes in the particle number participating in the process have profound effects on the dynamics.

The analytical solution can be recovered by eliminating all numerical noise produced by FFT after each simulation step. Using such masks in k-space, the numerical simulation and analytical solution agree within about 10^{-5} (see Fig. 9.11). However, this procedure yields a loss in total particle number of about $\Delta N/N = 10^{-6}$, far surpassing typical numerical noise.

For the implementation of the FWM neuron, we are interested in rather short time scales and more qualitative behaviour of the system. Therefore, we accept the change in frequency of the FWM oscillations and use the simulation on a discrete periodic lattice in the investigations without additionally applying a filter mask in k-space. This is beneficial due to the high flexibility of the simulation regarding initial conditions of the FWM states. However, the deviation between the ideal case and with present background population should be kept in mind, especially when looking at increasing simulation times.

10

FOUR-WAVE MIXING NEURON

In this thesis, we aim to harness the dynamics of the ideal FWM process described in the previous chapter to set up an AN. In addition to the processing of real numbers, ANs are also able to operate with complex-valued inputs and outputs [221]. As the FWM process is described in terms of complex amplitudes α_j, the presented implementation of the FWM neuron constitutes a complex-valued neuron. Due to the experimental accessibility of particle numbers and phases, we choose to describe the nonlinear activation function and the learning process in terms of absolute values and phases, rather than using real and imaginary parts of the complex amplitudes α_j [221]. The discussion of the FWM neuron is based on [36].

In Section 10.1, the implementation of an AN using the ideal FWM process is presented, describing the algorithmic procedure and the resulting nonlinear activation function. After describing steepest descent learning for complex-valued neurons in Section 10.2, the FWM neuron is used to solve the benchmark XOR problem in Section 10.3. In this context, the robustness of the FWM neuron against noise in the input data is also investigated.

10.1 IMPLEMENTATION OF FOUR-WAVE MIXING NEURON

In general, analogously to the description of ANs in Chapter 2, complex-valued neurons process an n-dimensional (complex-valued) input

$$x_j = |x_j| e^{\mathrm{i}\kappa_j}, \qquad j = 1, \ldots, n, \tag{10.1}$$

by multiplying individually with weights

$$w_j = |w_j| e^{\mathrm{i}\vartheta_j}, \tag{10.2}$$

summing up the weighted inputs $v_j = w_j x_j$ and yielding an output y via a nonlinear activation function Ω,

$$y = \Omega(u), \qquad u = \sum_{j=1}^{n} v_j. \tag{10.3}$$

10.1.1 Identifying Input and Output

We implement such a computational unit with the ideal FWM process of coherent matter-waves. The phase-flow

$$\tilde{\boldsymbol{\alpha}} = \Phi(\boldsymbol{\alpha}; \tau_{\mathrm{F}}), \tag{10.4}$$

maps the initial state $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_4)$ to the evolved state $\tilde{\boldsymbol{\alpha}} = (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_4)$ after the duration τ_F of the FWM process. Identifying the three amplitudes α_1, α_2 and α_3 as weighted inputs v_j and $\tilde{\alpha}_4$ as output y, a similar, though not identical, rule to (10.3) can be established

$$\tilde{\alpha}_4 = \Phi_4(\alpha_1, \alpha_2, \alpha_3, 0; \tau_\mathrm{F}). \tag{10.5}$$

The fourth component of the phase-flow map constitutes a nonlinear activation function of a complex-valued FWM neuron with three input channels. In an experiment, we suggest to use externally stored weights w_j of the neuron and the classical input data x_j to prepare the weighted input amplitude

$$\alpha_j = w_j x_j \tag{10.6}$$

by a sequence of Bragg pulses (see Appendix C). A diagrammatic visualization of the described implementation algorithm of the FWM neuron can be seen in Fig. 10.1.

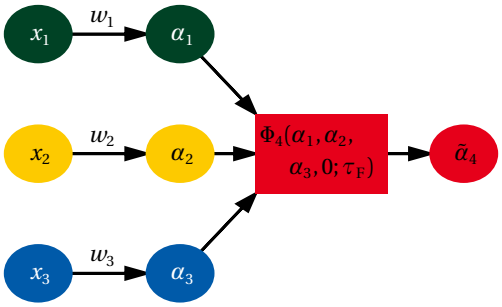

Figure 10.1: Initialisation sequence for a FWM neuron. Classical inputs x_j are weighted with w_j, yielding amplitudes α_j. The nonlinear relation $\Phi_4(\alpha_1, \alpha_2, \alpha_3, 0; \tau_\mathrm{F})$ yields the output $\tilde{\alpha}_4$.

10.1.2 Nonlinear Activation Function

In order to quantify the nonlinear activation function, τ_F has to be determined. To do so, we fix the total particle number of every neuron implementation by setting $\gamma = 1/\mathrm{s}$ and choose $n_1 = n_2 = 0.45$ and $n_3 = 0.1$, while $\varphi_j = 0$. The resulting FWM oscillation can be seen in Fig. 10.2. To maximize the output in terms of $\tilde{n}_4$ for this scenario, we set

$$\tau_\mathrm{F} = T/2, \tag{10.7}$$

where T is the oscillation period as in (9.53).

To obtain a quantitative value for $t_\mathrm{F} = \tau_\mathrm{F}/\gamma$, experimental parameters have to be taken into account. In their paper, *Chomaz et al.* [139] consider a quasi two-dimensional box trap with $N = 100{,}000$ ^{87}Rb atoms inside. They are able to produce traps with an area of $A = 200\,\mu\mathrm{m}^2$ while enclosing the BEC on a length scale of $h = 0.3\,\mu\mathrm{m}$ along the third direction, yielding a volume of $V = 60\,\mu\mathrm{m}^3$. With the scattering length a (9.61) and mass m (cf. Appendix B) of ^{87}Rb, the FWM neuron duration can be determined as

$$t_\mathrm{F} \simeq 43.2\,\mu\mathrm{s}. \tag{10.8}$$

The FWM neuron's response to varying weighted inputs is calculated numerically (cf. Section 9.3.1). The population probabilities n_j are tuned from 0 to 1 while enforcing the constraint

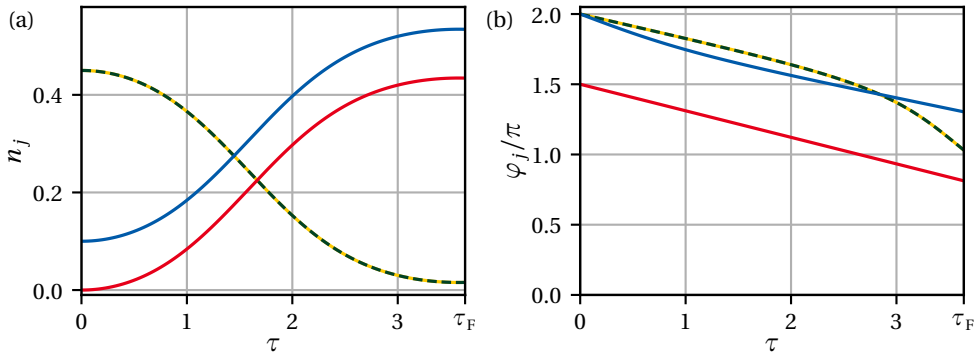

Figure 10.2: FWM dynamics in terms of (a) population probabilities n_j and (b) phases φ_j versus dimensionless time τ for $\gamma = 1/\mathrm{s}$, $n_1 = n_2 = 0.45$ and $n_3 = 0.1$, while $\varphi_j = 0$ (α_1: green, dashed; α_2: yellow, solid; α_3: blue; α_4: red). Neuron algorithm duration $\tau_F = T/2$ is determined as a half-oscillation period leading to maximal response.

$\sum_{j=1}^{4} n_j = 1$. Due to probability (number) conservation, all admissible combinations of n_j form a plane in n_1-n_2-n_3-space. The input phases φ_j are varied from 0 to 2π.

The resulting nonlinear activation function can be seen in Fig. 10.3. The output particle number

$$\tilde{n}_4 = |\Phi_4(\alpha_1, \alpha_2, \alpha_3, 0; \tau_F)|^2 \tag{10.9}$$

is independent of the input phases φ_j. Hence, only the input particle numbers n_j determine this part of the output. While we were not able to recover an analytical expression for the relation, it can be extracted from Fig. 10.3 (a), that there has to be an exchange symmetry regarding n_1 and n_2.

The output phase

$$\tilde{\varphi}_4 = \arg[\Phi_4(\alpha_1, \alpha_2, \alpha_3, 0; \tau_F)] \tag{10.10}$$

exhibits a remarkably simple behaviour. By analyzing Fig. 10.3 (b), one finds

$$n_\varphi = 3n_1 + 3n_2 + 5n_3, \qquad \varphi_\varphi = \varphi_1 + \varphi_2 - \varphi_3. \tag{10.11}$$

Accordingly, the input-output-relation reads

$$\tilde{\varphi}_4 = s\, n_\varphi + \varphi_\varphi + d, \tag{10.12}$$

where the slope and offset of phase were determined from a fit as $s = (-1.77 \pm 0.01)$ and $d = (2.67 \pm 0.04)$.

The numerical results in Fig. 10.3 can be used to determine the partial derivatives $\partial \tilde{n}_4/\partial n_j$, $\partial \tilde{\varphi}_4/\partial n_j$ and $\partial \tilde{\varphi}_4/\partial \varphi_j$. These are needed to be able to train the neuron according to a steepest descent method.

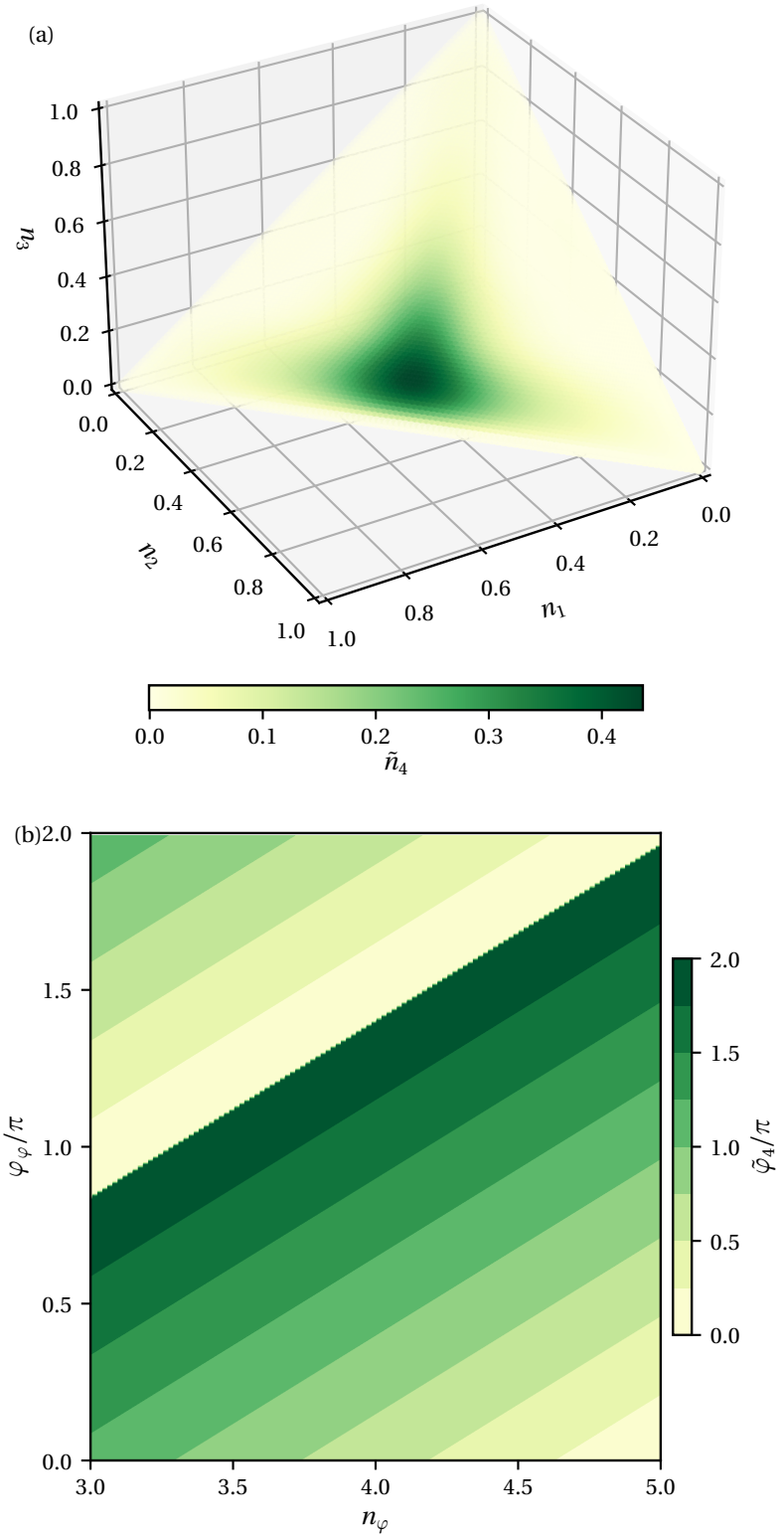

Figure 10.3: Nonlinear activation function $\Phi_4(\alpha_1, \alpha_2, \alpha_3, 0; \tau_F)$ of the FWM neuron in terms of (a) $\tilde{n}_4$ (10.9) versus n_1, n_2 and n_3 and (b) $\tilde{\varphi}_4$ (10.10) versus n_φ and φ_φ (10.11).

10.2 STEEPEST DESCENT LEARNING FOR COMPLEX-VALUED NEURONS

The steepest descent learning method for single-layer feedforward networks, in this case for a single output neuron, has been introduced in Section 3.1. Here, the developed training method via iterative weight updates is extended to complex-valued neurons.

As weights, inputs and outputs are complex-valued quantities, this characteristic is also imposed on the cost function $\mathcal{E}$ of the training procedure. Again, we consider the squared error averaged over a training sample set (3.2) as the cost function. For the FWM neuron, the total instantaneous error measure is given by

$$\mathcal{E}^{(i)} = \frac{1}{2}\left|\tilde{\alpha}_4^{(i)} - \hat{\alpha}_4^{(i)}\right|^2. \tag{10.13}$$

There, $\hat{\alpha}_4^{(i)}$ is the desired response associated with $\boldsymbol{x}^{(i)}$ and $\tilde{\alpha}_4^{(i)}$ is the neuron response to this stimulus.

As before, we perform the training procedure in an on-line manner (cf. Section 3.1.2). The absolute values and phases of the weights can be updated independently [221]

$$\Delta|w_j^{(i)}| = -\eta_\mathrm{a}\partial_{|w_j|}\mathcal{E}^{(i)}, \qquad \Delta\vartheta_j^{(i)} = -\eta_\mathrm{p}\partial_{\vartheta_j}\mathcal{E}^{(i)}, \tag{10.14}$$

where η_a and η_p are the learning rates for absolute value and phase, respectively. The required gradients for the update rules (10.14), keeping in mind the variable dependencies of the nonlinear activation function, are calculated using the chain rule as

$$\begin{aligned}
\frac{\partial\tilde{n}_4}{\partial|w_j|} &= \frac{\partial\tilde{n}_4}{\partial n_j}\frac{\partial n_j}{\partial|w_j|} = |x_j|\frac{\partial\tilde{n}_4}{\partial n_j},\\
\frac{\partial\tilde{\varphi}_4}{\partial|w_j|} &= \frac{\partial\tilde{\varphi}_4}{\partial n_j}\frac{\partial n_j}{\partial|w_j|} = |x_j|\frac{\partial\tilde{\varphi}_4}{\partial n_j},\\
\frac{\partial\tilde{\varphi}_4}{\partial\vartheta_j} &= \frac{\partial\tilde{\varphi}_4}{\partial\varphi_j}\frac{\partial\varphi_j}{\partial\vartheta_j} = \frac{\partial\tilde{\varphi}_4}{\partial\varphi_j}.
\end{aligned} \tag{10.15}$$

Hence, the update rules for $|w_j|$ and ϑ_j are

$$\Delta|w_j^{(i)}| = -\eta_\mathrm{a}\Big[\Big(\tilde{n}_4^{(i)} - \hat{n}_4^{(i)}\cos\Big(\tilde{\varphi}_4^{(i)} - \hat{\varphi}_4^{(i)}\Big)\Big)\partial_{n_j}\tilde{n}_4^{(i)} + \tilde{n}_4^{(i)}\sin\Big(\tilde{\varphi}_4^{(i)} - \hat{\varphi}_4^{(i)}\Big)\partial_{n_j}\tilde{\varphi}_4^{(i)}\Big]|x_j^{(i)}|, \tag{10.16}$$

$$\Delta\vartheta_j^{(i)} = -\eta_\mathrm{p}\tilde{n}_4^{(i)}\hat{n}_4^{(i)}\sin\Big(\tilde{\varphi}_4^{(i)} - \hat{\varphi}_4^{(i)}\Big)\partial_{\varphi_j}\tilde{\varphi}_4^{(i)}. \tag{10.17}$$

10.3 XOR PROBLEM

To test its learning capabilities, the FWM neuron is used to solve the XOR problem as introduced in Section 3.3.1. As discussed before, a single real-valued neuron is not sufficient to solve this problem, i.e. hidden layers are required [42]. However, a single complex-valued neuron is indeed able to solve this problem [222].

10.3.1 Input and Output Encoding

To use the full value range of the nonlinear activation function of the FWM neuron to solve the XOR problem, an encoding scheme for the inputs and the output has to be developed. The inputs $x_{1,2}$ are chosen to lie on the positive real axis ($\kappa_j = 0$). While an input 0 is identified by $|x_j| = 0.3$, an input 1 is given by $|x_j| = 0.45$.

The weights w_j of the neuron are still allowed to possess non-vanishing phases ϑ_j. Therefore, the weighted inputs presented to the FWM neuron will be given by

$$\sqrt{n_j} = |w_j||x_j|, \qquad \varphi_j = \vartheta_j. \tag{10.18}$$

As two input particle numbers, chosen to be n_1 and n_2, of the FWM neuron are set using this encoding, the third, in this case n_3, is automatically determined to ensure $\sum_j n_j = 1$. Consequently, the combinations of inputs n_1 and n_2 are constrained by $0 \leq n_1 + n_2 \leq 1$.

The particle number response of the FWM neuron to the inputs is completely determined by the input particle numbers $\tilde{n}_4(n_1, n_2)$. The neuron response in terms of the phase follows

$$\tilde{\varphi}_4(n_1, n_2, \varphi_1, \varphi_2) = \tilde{\varphi}_4(n_1, n_2, 0, 0) + \varphi_1 + \varphi_2 \tag{10.19}$$

These input-output relations can be seen in Fig. 10.4.

The possible outputs of the XOR problem are encoded in a similar fashion. An output 0 is encoded via $\tilde{n}_4 = 0.125$ and $\tilde{\varphi}_4 = 1.5$ or $\tilde{n}_4 = 0.435$ and $\tilde{\varphi}_4 = 2.5$ for the input cases $[0,0]$ and $[1,1]$ respectively. The output 1 is always encoded as $\tilde{n}_4 = 0.155$ and $\tilde{\varphi}_4 = 2$. The complete encoding of the XOR problem for the FWM neuron can be seen in Table 10.1. The presented encoding is completely equivalent to the original XOR problem. Hence, it can be used to solve the problem by means of the FWM neuron.

10.3.2 Training Results

Starting from random initial weights, the update rules (10.16) and (10.17) are used to train the FWM neuron to solve the XOR problem. Training epochs are performed with 1,000 random samples. The learning rate of the phase $\eta_\mathrm{p} = 10^{-8}$ is kept constant for all epochs while the absolute value learning rate η_a is gradually reduced from 10^{-3} to 10^{-4} during the training. After each epoch, the performance of the neuron is evaluated by calculating the averaged

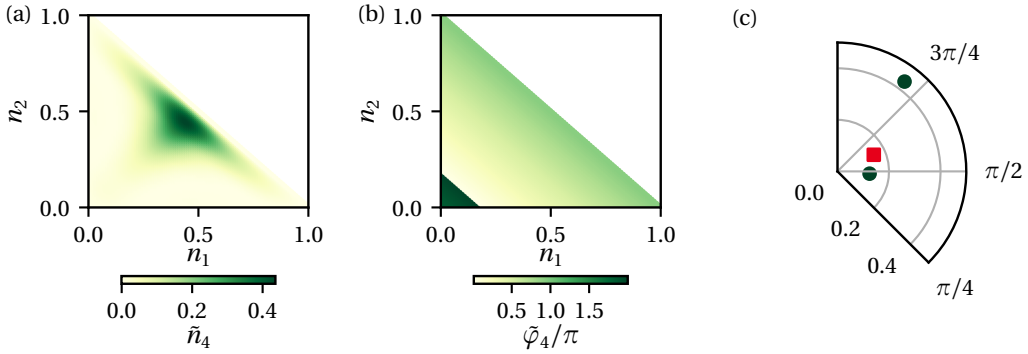

Figure 10.4: Input-output relations (a) $\tilde{n}_4(n_1, n_2)$ and (b) $\tilde{\varphi}_4(n_1, n_2, 0, 0)$ of the FWM neuron to solve the XOR problem. (c) By choosing the outputs of the individual cases according to Table 10.1 (green, output 0; red, output 1), the XOR problem is solvable using a single FWM neuron.

Table 10.1: Encoded input-output mapping for the XOR problem using the FWM neuron.

Input 1	Input 2	$\lvert x_1\rvert$	$\lvert x_2\rvert$	Output	$\tilde{n}_4$	$\tilde{\varphi}_4$
0	0	0.3	0.3	0	0.125	1.5
0	1	0.3	0.45	1	0.155	2.0
1	0	0.45	0.3	1	0.155	2.0
1	1	0.45	0.45	0	0.435	2.5

squared error $\mathcal{E}$ according to (3.2) using all $m = 4$ possible input-output pairs of the XOR problem.

As can be seen in Fig. 10.5, the FWM neuron is able to learn to solve the XOR problem. After 100 training epochs, the initial error is reduced to $\mathcal{E} = 7.8 \cdot 10^{-6}$. A sample is categorized as being identified correctly, if the neuron output is within ±0.01 in terms of particle number and within ±0.1 in terms of phase of the desired value. At the end of the training procedure, every test sample is identified correctly.

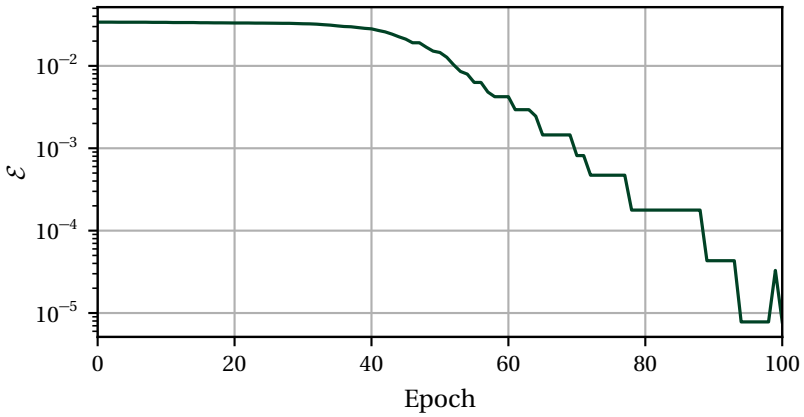

Figure 10.5: Averaged squared error $\mathcal{E}$ (3.2) over all four possible input-output pairs of the XOR problem versus number of training epochs.

10.3.3 Imperfect Training and Testing Data

In real-world application, input data tends to be imperfect. To take this aspect into account, random noise is added to the neuron inputs. At first, this will be limited to test samples to investigate the influence of such noisy data on test results for a neuron trained with ideal data. Subsequently, the neuron is also subjected to noisy data during the training process.

Noisy Testing Data

The FWM neuron trained on the XOR problem as described in Section 10.3.2 is to be tested in terms of robustness against noise in the input data. To do this, the network setup at the end of the training procedure is presented with noisy test samples. This additive noise is assumed to be Gaussian with zero mean and standard deviation σ, influencing both inputs $\lvert x_1\rvert$ and $\lvert x_2\rvert$.

The standard deviation is tuned from $\sigma = 0$ to $\sigma = 0.05$. For each standard deviation, the averaged squared error $\mathcal{E}$ is calculated according to (3.2) for $m = 10{,}000$ noisy inputs. Additionally, the percentage of correctly identified test samples p_c is determined for each

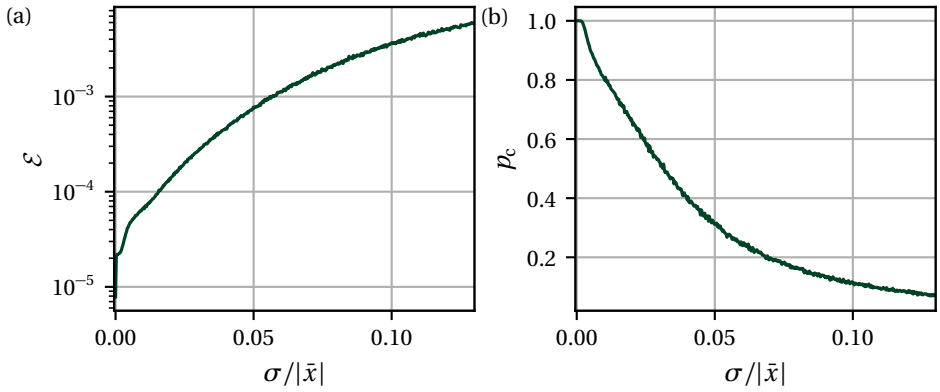

Figure 10.6: (a) Averaged squared error $\mathcal{E}$ and (b) percentage of correctly identified samples p_c versus standard deviation σ for testing with noisy data on the FWM neuron trained for 100 epochs with ideal data to solve the XOR problem. The additive noise on the inputs is Gaussian with zero mean and standard deviation σ. σ is normed to the mean input $|\bar{x}|$=0.375.

standard deviation. As in Section 10.3.2, a sample is categorized as being identified correctly if the neuron output is within ±0.01 in terms of particle number and within ±0.1 in terms of phase of the desired value.

As can be seen in Fig. 10.6, both $\mathcal{E}$ and p_c indicate a decrease in performance with increasing σ. Looking at p_c, 90 % of the test samples are identified correctly for $\sigma \simeq 0.05|\bar{x}|$, where $|\bar{x}| = 0.375$ is the mean input. For a standard deviation of 1 % of the mean input $\sigma = 0.01|\bar{x}|$, the neuron identifies 79.9 % of the samples correctly, dropping to 31.1 % for $\sigma = 0.05|\bar{x}|$ and 11.3 % for $\sigma = 0.1|\bar{x}|$.

Noisy test data clearly yields a loss in performance of the FWM neuron. However, typically not only testing data is subject to noise. Such noisy inputs will also be present during the training procedure.

Training and Testing on Noisy Data

To investigate the influence of noisy training inputs on the performance of the FWM neuron in solving the XOR problem, an additive Gaussian noise with zero mean and standard deviation σ_{train} is applied on every training sample. Equivalently to Section 10.3.2, the training is performed over 100 epochs with 1,000 noisy samples per epoch.

As before, the performance of the FWM neuron in solving the XOR problem is tested after the training procedure on noisy data in terms of the averaged squared error $\mathcal{E}$ and the percentage of correctly identified samples p_c using $m = 10{,}000$ test samples. The additive noise applied to the testing inputs is Gaussian with zero mean and standard deviation σ_{test}.

As can be seen in Fig. 10.7, training on noisy data has a similar effect on the performance of the FWM neuron as testing with such data. For constant testing noise, p_c decreases with increasing σ_{train}. In particular, for $\sigma_{\text{test}} = 0$, p_c is constant at 100 % up to a certain degree in training noise and subsequently exhibits a sharp drop.

From the calculated results, there does not seem to be an exceptional point where noise in training and testing data constructively produce better performance. This intuition is reinforced by the results in Fig. 10.8. For equal noise in training and testing $\sigma = \sigma_{\text{train}} = \sigma_{\text{test}}$, the FWM neuron shows nearly identical performance as in testing on noisy data after training on ideal samples. This reveals a robustness of the training procedure of the FWM neuron against noise.

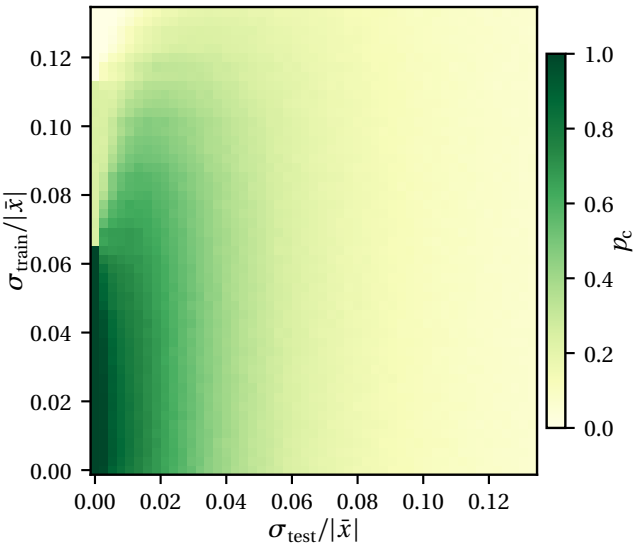

Figure 10.7: Percentage of correctly identified samples p_c versus standard deviations of the testing σ_{test} and the training inputs σ_{train}. Standard deviations are normed to the mean input $|\bar{x}|$=0.375.

10.3.4 Discussion

The FWM neuron proves that coherent matter waves and their intrinsic nonlinear interactions can be used to set up an AN. Identifying three complex amplitudes of the FWM setup as input and the fourth amplitude as output, a new implementation for a complex-valued neuron can be introduced. Through in-detail investigation of the nonlinear response of the FWM neuron to complex-valued input data, its learning capabilities can be tested using steepest descent learning for complex-valued neurons.

Due to its complex nature, the FWM neuron is able to solve the XOR problem with just a single neuron. At the end of a learning procedure, the FWM neuron is able to identify all test samples correctly. Even when presented with noisy data, the FWM neuron shows decent performance up to a certain degree of noise.

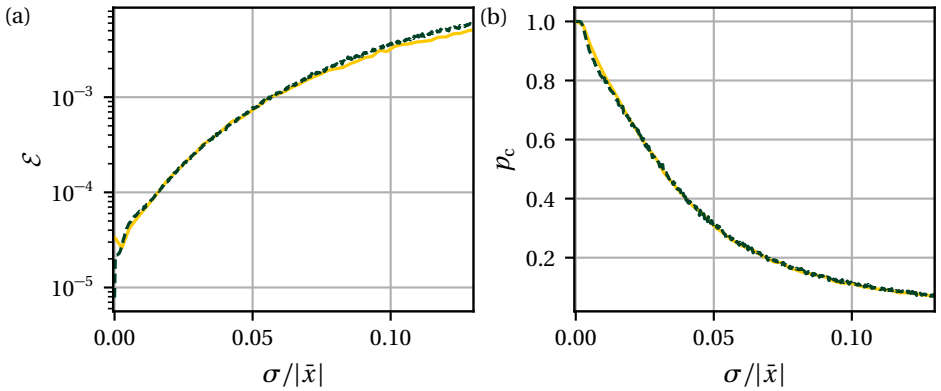

Figure 10.8: (a) Averaged squared error $\mathcal{E}$ and (b) percentage of correctly identified samples p_c versus standard deviation σ. Training on ideal samples and testing on noisy data (green, dotted) yields nearly identical results as applying equal noise during training and testing (yellow, solid).

Additionally, with a duration of the FWM neuron implementation of $t_F = 43.2\,\mu$s, the input processing is much faster than it was the case for the thermal cloud neuron. Hence, an actual implementation of the FWM neuron seems within the realm of possibility.

To be able to tackle more sophisticated problems, other than the XOR problem, one has to go beyond a single FWM neuron. Therefore, a logical step forward is the implementation of a FWM neuron in a feedforward neural network architecture.

11

FOUR-WAVE MIXING NEURAL NETWORK

To implement the FWM neuron in a deep feedforward neural network, two key aspects have to be investigated: parallelization ability and communication between layers of the network. As two-dimensional FWM and one-dimensional FWM with multiple internal levels are equivalent (cf. Section 9.2.3), both variants are taken into account for network implementation.

To parallelize multiple FWM neuron implementations, different pathways can be considered: spatial parallelization, parallelization in momentum space or using the extended internal structure of the atoms. Those investigations are performed in Section 11.1. It is shown, that especially parallelizing FWM neurons using multiple internal states lends itself well to set up a network. A scheme to implement communicating layers of a network is shown in Section 11.2, realizing a tree-like neural network. Subsequently, this implementation is trained on an example dataset using a six-dimensional input, three FWM neurons in a hidden layer and one FWM neuron in the output layer to display the learning capabilities of the FWM neural network.

11.1 NEURON PARALLELIZATION

In an ANN, the individual neurons inside one layer are supposed to operate independently of each other. Therefore, the goal of parallelization is to find a setup such that the dynamics of a single neuron are not disturbed by the presence of other neurons.

11.1.1 Spatial Separation

A convenient way to parallelize the FWM neuron is to implement spatially separated realizations of the neuron to form a layer of a network. In this case, the coherent matter waves participating in the FWM process of the individual neurons do not overlap and, as a result, do not interact with each other. Hence, the neurons act completely independent of each other.

In the two-dimensional realization of the FWM setup, spatial separation can be achieved by spatially stacking two-dimensional box traps (cf. Section 6.1.2). Analogously, the one-dimensional implementation with multiple internal states can be parallelized spatially by using individual ring potentials for the individual neurons. A visualization of spatially stacking two-dimensional box traps can be seen in Fig. 11.1.

The downside of spatially separating individual neuron implementations lies in the next step of setting up a network, transporting the output information into the next layer. As the neurons are spatially separated, the output information of a single neuron (complex amplitude $\tilde{\alpha}_4$ of the fourth momentum component of the FWM process) would have to be transported through space and merged with outputs from other neurons to implement the input of the next layer. This seems to be quite cumbersome and experimentally difficult to achieve. Hence,

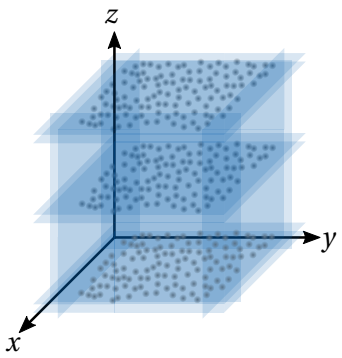

Figure 11.1: Visualization of box traps in the x-y-plane spatially stacked along the z-direction.

alternative approaches to obtain neuron parallelization are considered where the individual FWM processes do not show spatial separation.

11.1.2 Stacked Four-Wave Mixing in Momentum Space

An alternative to spatially separating FWM neuron setups is to implement multiple realizations stacked in momentum space. To investigate this scenario, consider a two-dimensional FWM configuration with wave vectors $\boldsymbol{k}_1$, $\boldsymbol{k}_2$, $\boldsymbol{k}_3$ and $\boldsymbol{k}_4$, satisfying the FWM conditions (9.19). This configuration can be arbitrarily chosen to lie in the k_x-k_y-plane. For parallelization, we stack a second FWM configuration with wave vectors $\boldsymbol{k}_5$, $\boldsymbol{k}_6$, $\boldsymbol{k}_7$ and $\boldsymbol{k}_8$ on top of the original one in k-space by setting

$$\boldsymbol{k}_{j+4} = \boldsymbol{k}_j + \Delta k_z, \qquad j = 1,\dots,4, \tag{11.1}$$

where Δk_z is the separation of the two configurations along the k_z-axis. A visualization of the proposed geometry can be seen in Fig. 11.2. The wave vectors of the second FWM configuration also satisfy the FWM conditions (9.19).

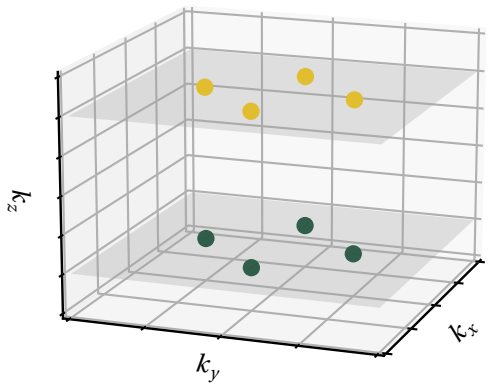

Figure 11.2: Visualization of geometry of the FWM configurations stacked in momentum space. The second FWM configuration (yellow) is a copy of the original one (green), shifted along the k_z-axis by Δk_z.

Dynamics of Stacked Setup

The dynamics for the complex amplitudes $\alpha_j(t)$, $j \in [1,8]$, can be recovered by determining the Lagrangian L according to (9.12) for the stacked FWM ansatz

$$\left|\psi_\alpha\right\rangle = \sum_{j=1}^{8} \sqrt{N}\alpha_j \left|\boldsymbol{k}_j\right\rangle. \tag{11.2}$$

Evaluating the dimensionless Euler-Lagrange equations (see Section 9.2.1) reveals, that the system of differential equations describing the dynamics of the amplitudes can be written in block structure as

$$\mathrm{i}\dot{\boldsymbol{\alpha}} = \begin{pmatrix} H_\ell & 0 \\ 0 & H_\mathrm{u} \end{pmatrix} \boldsymbol{\alpha}, \tag{11.3}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_8)^\mathrm{T}$. The matrices governing the dynamics of the lower H_ℓ and upper configuration H_u, respectively, are given by

$$H_\ell = 2\begin{pmatrix} \frac{\bar{\omega}_1}{2} + m_1 & \alpha_6^*\alpha_5 & \alpha_2^*\alpha_4 + \alpha_6^*\alpha_8 + \alpha_7^*\alpha_5 & \alpha_6^*\alpha_7 + \alpha_8^*\alpha_5 \\ \alpha_5^*\alpha_6 & \frac{\bar{\omega}_2}{2} + m_2 & \alpha_5^*\alpha_8 + \alpha_7^*\alpha_6 & \alpha_1^*\alpha_3 + \alpha_5^*\alpha_7 + \alpha_8^*\alpha_6 \\ \alpha_4^*\alpha_2 + \alpha_5^*\alpha_7 + \alpha_8^*\alpha_6 & \alpha_6^*\alpha_7 + \alpha_8^*\alpha_5 & \frac{\bar{\omega}_3}{2} + m_3 & \alpha_8^*\alpha_7 \\ \alpha_5^*\alpha_8 + \alpha_7^*\alpha_6 & \alpha_3^*\alpha_1 + \alpha_7^*\alpha_5 + \alpha_6^*\alpha_8 & \alpha_7^*\alpha_8 & \frac{\bar{\omega}_4}{2} + m_4 \end{pmatrix} \tag{11.4}$$

and

$$H_\mathrm{u} = 2\begin{pmatrix} \frac{\bar{\omega}_1 + \bar{\omega}_z}{2} + m_5 & \alpha_2^*\alpha_1 & \alpha_2^*\alpha_4 + \alpha_3^*\alpha_1 + \alpha_6^*\alpha_8 & \alpha_2^*\alpha_3 + \alpha_4^*\alpha_1 \\ \alpha_1^*\alpha_2 & \frac{\bar{\omega}_2 + \bar{\omega}_z}{2} + m_6 & \alpha_1^*\alpha_4 + \alpha_3^*\alpha_2 & \alpha_1^*\alpha_3 + \alpha_4^*\alpha_2 + \alpha_5^*\alpha_7 \\ \alpha_4^*\alpha_2 + \alpha_1^*\alpha_3 + \alpha_8^*\alpha_6 & \alpha_4^*\alpha_1 + \alpha_2^*\alpha_3 & \frac{\bar{\omega}_3 + \bar{\omega}_z}{2} + m_7 & \alpha_4^*\alpha_3 \\ \alpha_3^*\alpha_2 + \alpha_1^*\alpha_4 & \alpha_3^*\alpha_1 + \alpha_2^*\alpha_4 + \alpha_7^*\alpha_5 & \alpha_3^*\alpha_4 & \frac{\bar{\omega}_4 + \bar{\omega}_z}{2} + m_8 \end{pmatrix}, \tag{11.5}$$

where

$$m_i = \sum_{j=1}^{8} n_j^2 - \frac{n_i}{2}, \qquad \bar{\omega}_z = \frac{\hbar|\Delta k_z|^2}{2m\gamma}. \tag{11.6}$$

The block structure of (11.3) implies that the total particle numbers in the individual configurations

$$N_\ell = \sum_{j=1}^{4} n_j N, \qquad N_\mathrm{u} = \sum_{j=5}^{8} n_j N \tag{11.7}$$

are conserved. Hence, there is no effective particle interchange between the two configurations. However, there are terms present in H_ℓ as well as in H_u which generate coherent coupling between the two configurations. Consequently, the dynamics of the configurations have the potential to influence each other, which is supposed to be avoided in the setup of a network structure.

Analysis of Independence of Stacked Configurations

The aim of parallelization is that the dynamics of a FWM neuron remain undisturbed in the presence of another neuron. Hence, we can define an ideal target state $|\psi_{\mathrm{id}}\rangle$, which is given by the state of an individual neuron, described by the wave function in the k_x-k_y-plane at time τ_{F}, and depends on the input amplitudes α_1, α_2 and α_3 (see Section 10.1). Analogously, we define test states $|\psi_{\mathrm{te}}\rangle$ as the state of an identical neuron in the k_x-k_y-plane, however now in the presence of a second neuron stacked in momentum space. This second neuron is characterized by the same implementation algorithm described in Section 10.1. In particular, this implies, that the total particle number of the second neuron is the same as the one of the reference neuron. However, the second neuron might show varying ratios of particle number inputs $n_{\mathrm{u},1}$, $n_{\mathrm{u},2}$ and $n_{\mathrm{u},3}$. Additionally, the separation of the neurons along the k_z-axis Δk_z might be varied.

To quantify possible deviations between the ideal and test states, we introduce the fidelity

$$\mathcal{F} = \left|\langle \psi_{\mathrm{id}} | \psi_{\mathrm{te}} \rangle\right|^2 . \tag{11.8}$$

This is a measure which characterizes the overlap between the two states. The value range of the fidelity is $\mathcal{F} \in [0,1]$, where $\mathcal{F} = 1$ if the two states perfectly overlap. Hence, if the neuron activity is completely undisturbed by the presence of an additional neuron, the ideal $|\psi_{\mathrm{id}}\rangle$ and test state $|\psi_{\mathrm{te}}\rangle$ should yield $\mathcal{F} = 1$.

PARAMETERS OF NUMERICAL SIMULATION Throughout this section, we investigate numerical simulations of the GPE on a three-dimensional discrete grid with $16 \times 16 \times 16$ sites. In the simulations, dimensionless time is discretized with $\Delta\tau = 10^{-4}$, while the simulation time is taken to be τ_{F} from Section 10.1. Again, the geometry described in Appendix C is chosen, yielding $\omega_j = 1/\mathrm{s}$, $j = 1,\ldots,4$, for the momentum components in the k_x-k_y-plane. For the isolated reference neuron, we choose $\gamma = 1/\mathrm{s}$ and set $n_1 = n_2 = 0.45$ and $n_3 = 0.1$ while keeping all phases at $\varphi_j = 0$. To compensate for the increased total particle number in the investigation of the test neuron due to the presence of the stacked neuron, the frequency is accordingly set to $\gamma = 2/\mathrm{s}$.

INFLUENCE OF SEPARATION IN MOMENTUM SPACE As a first benchmark, we consider the case, where the stacked FWM neuron is identical to the reference neuron. This implies $n_{\mathrm{u},1} = n_{\mathrm{u},2} = 0.45$ and $n_{\mathrm{u},3} = 0.1$ while $\varphi_{\mathrm{u},j} = 1$. The two configurations are separated by

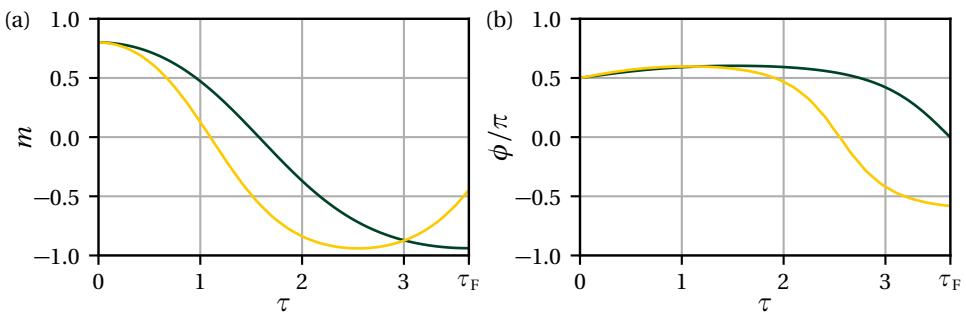

Figure 11.3: Dynamics of FWM neuron setup with (yellow) and without (green) the presence of an additional identical stacked neuron with separation $\Delta\kappa_z = 8$ in terms of (a) population imbalance m and (b) phase difference ϕ versus dimensionless time τ.

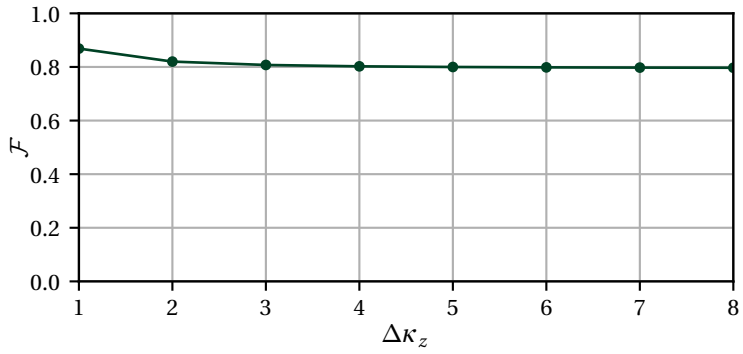

Figure 11.4: Fidelity $\mathcal{F}$ versus separation along the k_z-axis in terms of grid sites $\Delta\kappa_z$ between the ideal case of an isolated FWM neuron and a FWM neuron in the presence of an additional identical stacked neuron.

$\Delta\kappa_z = 8$, where $\Delta\kappa_z$ is the separation of configurations along the k_z-direction in terms of discrete grid sites. The dynamics of the reference and test FWM neurons are compared in terms of population imbalance m (9.29) and phase difference ϕ (9.28). As can be seen in Fig. 11.3, the dynamics of the FWM neuron with and without the presence of a second neuron differ visibly. At time τ_F, the fidelity is given by $\mathcal{F} = 0.797$.

Clearly, the dynamics of the FWM neuron are influenced by the presence of the second configuration. To investigate whether this influence is dependent on the separation of the two configurations along the k_z-direction, the fidelity $\mathcal{F}$ is calculated for varying $\Delta\kappa_z$. As can be seen in Fig. 11.4, the fidelity decreases with increasing separation, however not in a drastic manner. Yet, there is no separation at which the influence of the second configuration can be neglected.

INFLUENCE OF PARTICLE NUMBER RATIOS Already from the first example in the previous paragraph, the presence of a stacked neuron seems to influence the dynamics of a test neuron. However, in a neural network, the input states of two neurons in a layer will typically not be identical. To investigate the effect of varying input states of a stacked neuron on a test neuron,

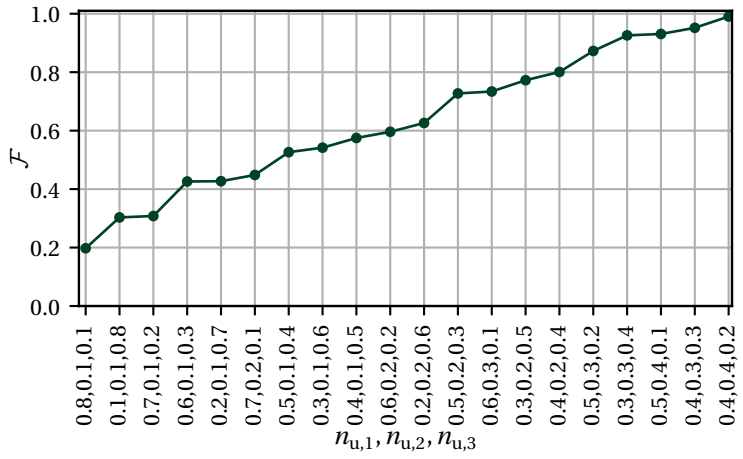

Figure 11.5: Fidelity $\mathcal{F}$ versus relative particle number ratios $n_{\mathrm{u},1}$, $n_{\mathrm{u},2}$, $n_{\mathrm{u},3}$ of stacked FWM neuron.

we vary the relative particle number ratios $n_{u,1}$, $n_{u,2}$, $n_{u,3}$ and calculate the fidelity between the state of the test neuron and an isolated reference neuron.

As can be seen in Fig. 11.5, the fidelity changes heavily with varying combinations $n_{u,1}$, $n_{u,2}$, $n_{u,3}$. Spanning a range from $\mathcal{F} \approx 0.2$ to near perfect matching between test and ideal state, the fidelity is typically higher for particle number ratios which are similar to the one of the reference neuron ($n_1 = n_2 = 0.45$, $n_3 = 0.1$). This implies that the more different the input state of the stacked neuron is compared to the reference neuron in terms of particle number ratios, the more it will influence the dynamics of the latter.

Discussion

The presented investigation shows, that the dynamics of a FWM neuron are clearly influenced by the presence of a second FWM neuron stacked in momentum space. The severity of this influence in terms of the fidelity $\mathcal{F}$ after a neuron implementation time τ_F depends on the ratios of relative particle numbers of the stacked neuron. However, as all combinations should be realizable to set up a layer of a functioning neural network, any influence at all prohibits an ANN setup using this implementation method.

While the coupling of the FWM configurations stacked in momentum space prevents the usage of this setup in terms of deep feedforward neural networks, the mutual influence of the configurations might be exploited as a feature in an alternative approach to machine learning, namely reservoir computing [223–225]. In reservoir computing, the input data can be encoded in the dynamics of a physical system [226–229]. A pattern analysis is performed in the readout of the system. In contrast to neural networks, input weights w_{in} and weights w of possible recurrent connections inside the system are not updated during the training procedure. Only the output weights w_{out} of the readout are changed during the training, for example by linear regression. This implies, that the parameters of the physical system are kept fixed. Hence, hardly any control over the system dynamics is required. A schematic visualization of reservoir computing can be seen in Fig. 11.6.

A setup of d stacked FWM configurations offers the possibility to realize reservoir computing with up to $4d$ inputs. The more FWM configurations are stacked in momentum space, the more mutual interaction is introduced into the system, implying more possible recurrent connections inside the system. Performing a readout after some time of system evolution and using established methods from reservoir computing, this approach might offer an interesting implementation opportunity for reservoir computing worth investigating.

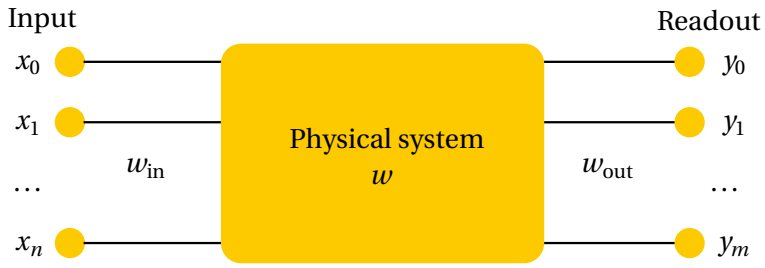

Figure 11.6: Schematic visualization of reservoir computing. The n-dimensional input is prepared as the state of a physical system, where a m-dimensional readout is performed after some time of system evolution. During the training procedure, only the output weights w_{out} are modified such that the required control over the physical system is minimal. Graphic adapted from [226].

11.1.3 Parallel Four-Wave Mixing using Internal States

Another approach to parallelize the FWM neuron implementation is to utilize the extended internal structure of the BEC. As seen in Section 9.2.3 the dynamics of a two-component BEC can be described by a two-component GPE (9.55). By analyzing the matrix representation (9.56) of this equation, one can identify a way to decouple the dynamics of the two components: for vanishing inter-component scattering length $a_{12} = 0$ no coupling between the two components arises and the dynamics will be independent of each other.

We envision a situation of two two-dimensional overlapping FWM configurations with identical geometry (see Appendix C), one in internal state $|1\rangle$ and the other in internal state $|2\rangle$. In an ideal scenario, the intra-component scattering lengths are equal $a_{11} = a_{22}$, while the inter-component scattering length vanishes $a_{12} = 0$. To find such a configuration, either one has to make use of a favorable configuration in the hyperfine structure of, for example, ^{87}Rb, or tune the intra- and inter-component interaction strengths via Feshbach resonances.

PARAMETERS OF NUMERICAL SIMULATION To investigate the feasibility of this decoupling ansatz, we compare the dynamics of an isolated reference FWM neuron with those of an identical test neuron in the presence of another FWM configuration in an additional internal state, similar to the procedure in Section 11.1.2. In this section, we use numerical simulations of the two-component GPE on a two-dimensional discrete grid with 16×16 sites. Again, dimensionless time is discretized with $\Delta\tau = 10^{-4}$ while the simulation time is taken to be τ_F from Section 10.1. Using the two-dimensional FWM geometry from Appendix C yields $\omega_j = 1/\mathrm{s}$, $j = 1, \ldots, 4$. For the isolated reference FWM configuration we set $\gamma = 1/\mathrm{s}$ and choose $n_1 = n_2 = 0.45$ and $n_3 = 0.1$ while keeping $\varphi_j = 0$. In investigating the test neuron, the increased particle number in the presence of a second FWM configuration has to be taken into account regarding γ only if the inter-component scattering length is non-vanishing. Accordingly, we set $\gamma = (1 + \tilde{a}_{12})/\mathrm{s}$, where $\tilde{a}_{12} = a_{12}/a_{11}$.

INFLUENCE OF INTER-COMPONENT SCATTERING LENGTH At first, we consider the case where the second FWM configuration is identical to the test neuron, yielding $n_{2,1} = n_{2,2} = 0.45$ and $n_{2,3} = 0.1$ with $\varphi_{2,j} = 0$. To investigate the influence of the relative inter-component scattering on the dynamics of the test neuron, we vary $\tilde{a}_{12}$. The dynamics of the reference and

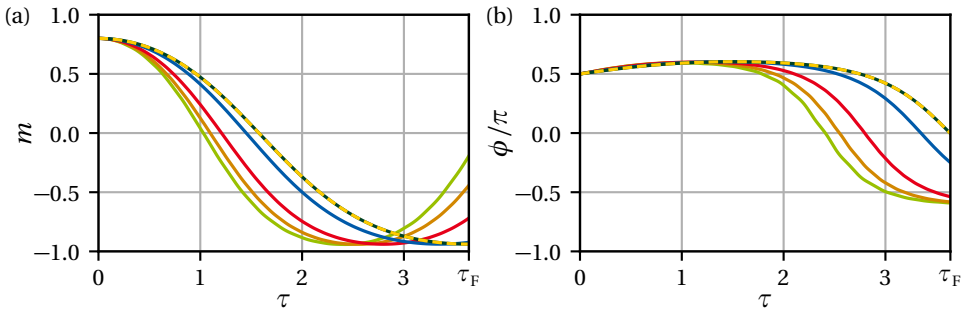

Figure 11.7: (a) Population imbalance m and (b) phase difference ϕ versus dimensionless time τ for an isolated reference FWM neuron (green) and a test FWM neuron with a second configuration present in an additional internal state with varying relative inter-component scattering lengths: $\tilde{a}_{12} = 0$ (yellow, dashed), $\tilde{a}_{12} = 0.1$ (blue), $\tilde{a}_{12} = 0.5$ (red), $\tilde{a}_{12} = 1$ (orange) and $\tilde{a}_{12} = 2$ (lime).

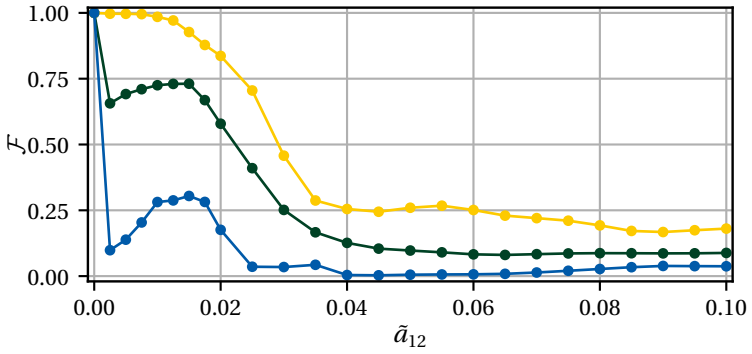

Figure 11.8: Mean (green), maximum (yellow) and minimum (blue) fidelity $\mathcal{F}$ versus relative inter-component scattering length $\tilde{a}_{12}$ for the combinations of input relative particle numbers $n_{2,1}$, $n_{2,2}$ and $n_{2,3}$ from Fig. 11.5.

test FWM configurations in terms of the population imbalance m and the phase difference ϕ can be seen in Fig. 11.7.

Clearly, for non-negligible inter-component scattering, the dynamics of the test FWM neuron are influenced significantly. However, with $\tilde{a}_{12}$ nearing lower values, the dynamics of the test neuron approach those of the reference neuron. To possibly confirm this impression, we deepen the present investigation by calculating the fidelity $\mathcal{F}$ (11.8) for varying relative inter-component scattering lengths $\tilde{a}_{12}$ and varying relative particle numbers $n_{2,1}$, $n_{2,2}$ and $n_{2,3}$ of the FWM configuration in the second internal state.

As can be seen in Fig. 11.8, the fidelity $\mathcal{F}$ is generally reduced below the desired value $\mathcal{F}=1$ for non-vanishing $\tilde{a}_{12}$. However, for $\tilde{a}_{12}=0$, the fidelity is $\mathcal{F}=1$ for every combination of relative particle numbers $n_{2,1}$, $n_{2,2}$ and $n_{2,3}$. Hence, in this case, the dynamics of the test FWM neuron are indeed undisturbed by the presence of a second configuration.

Using two-dimensional FWM in multiple internal states with vanishing inter-component scattering yields the possibility to parallelize the FWM neuron while ensuring that the dynamics of the individual neurons remain undisturbed. Hence, we use this parallelization method to set up a network of FWM neurons with multiple layers.

11.2 COMMUNICATION BETWEEN LAYERS - NETWORK SETUP

To test whether the FWM neuron can be implemented in a network structure, we examine an example case, namely a two-layer tree-like feedforward network. In contrast to fully-connected networks, neurons are not connected to every neuron in the adjacent layers in tree-like networks (see network structure in Fig. 11.9). They rather form branches leading to one or multiple roots. We realize a 9-3-1 network, implying a six-dimensional input, three neurons in the hidden layer and one neuron in the output layer.

11.2.1 Network Setup using Multiple Internal States

Every FWM neuron is realized via a two-dimensional FWM process in a distinguished internal state $|i\rangle$. Analogously to the investigations in Section 11.1.3, the intra-component scattering lengths a_{ii} are assumed to be equal for all internal components, while all inter-component scattering lengths a_{ij} are assumed to vanish. Using the wave vector constellations shown in

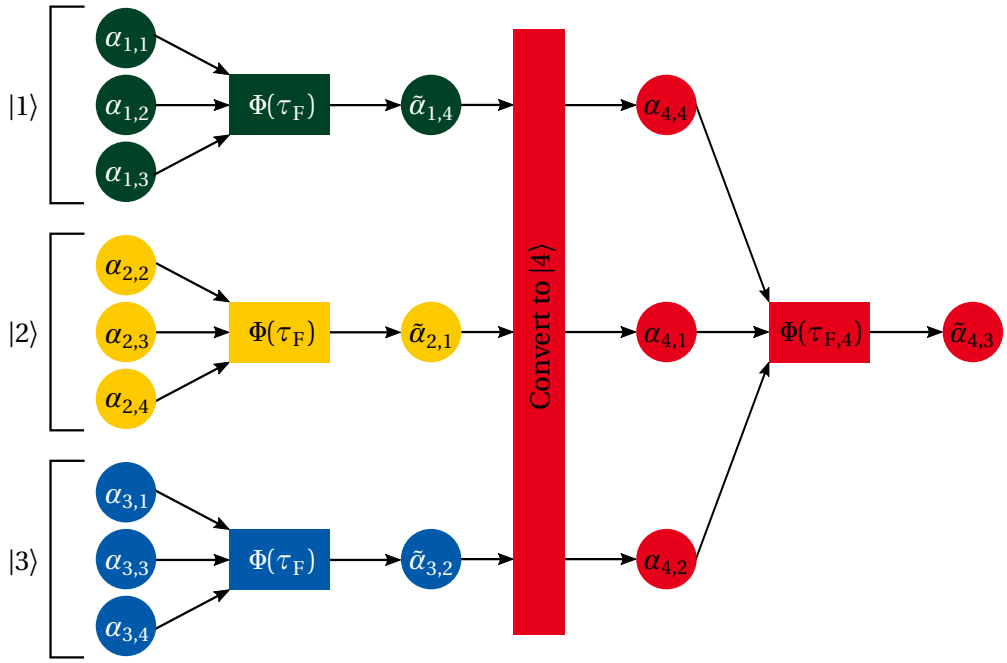

Figure 11.9: Setup of a tree-like feedforward FWM neural network. Inputs are prepared as described in Section 10.1 in three internal states $|1\rangle$, $|2\rangle$ and $|3\rangle$. For the case of vanishing inter-component scattering lengths a_{ij}, those neurons are independent of each other. The momentum components of the two-dimensional FWM process constituting the inputs and output are cyclically permuted to ensure a proper FWM configuration for the output layer. Outputs $\tilde{\alpha}_{1,4}$, $\tilde{\alpha}_{2,1}$ and $\tilde{\alpha}_{3,2}$ of the hidden layer are transported to the output layer via a conversion to internal state $|4\rangle$ using Rabi oscillations. The total output of the network is given by $\tilde{\alpha}_{4,3}$ with the implementation algorithm of the neuron in the output layer described in Section 11.2.2.

Fig. 11.9, a propagation from the output signals of the neurons in the hidden layer to an input of the output neuron is possible. There, transport of population between different internal states can be achieved via Rabi oscillations [128].

The three neurons in the hidden layer are implemented according to the algorithm described in Section 10.1. Accordingly, the total particle number of the input of one neuron is always fixed (all particle numbers in the implementation of the FWM neural network are normalized to this total input particle number of an individual neuron in the hidden layer). Therefore, by setting two particle numbers, the third one is implicitly determined to ensure $\sum_j n_{i,j} = 1$, $i = [1,2,3]$. Hence, the network is able to process six independent inputs. While the phase output of a FWM neuron, realized according to Section 10.1, has a value range $\tilde{\varphi}_4 \in [0, 2\pi]$ which is equal to the input value range of a FWM neuron, the output particle number value range is $\tilde{n}_4 \in [0, 0.43]$.

The roles of the momentum components of the FWM process as input and output of the neurons are cyclically permuted to ensure that the neuron in the output layer receives input in three different momentum states, again enabling FWM. This cyclic permutation leads to slight changes in the implementation algorithm of the FWM neurons in some internal states. For the case of internal state $|1\rangle$, where the fourth momentum component is identified as output, the formulation of the FWM neuron is exactly given by the description in Section 10.1. For internal states $|2\rangle$ and $|3\rangle$, where the output is given by the first and second momentum component, respectively, minor adjustments have to be taken into account. The particle number outputs

$\tilde{n}_{2,1}$ and $\tilde{n}_{3,2}$ show an exchange symmetry regarding $n_{i,3}$ and $n_{i,4}$, $i \in [2,3]$. Additionally, the phase outputs $\tilde{\varphi}_{2,1}$ and $\tilde{\varphi}_{3,2}$ respond to the linear combinations

$$n_{\varphi,2} = 3n_{2,3} + 3n_{2,4} + 5n_{2,2}, \qquad \varphi_{\varphi,2} = \varphi_{2,3} + \varphi_{2,4} - \varphi_{2,2} \tag{11.9}$$

and

$$n_{\varphi,3} = 3n_{3,3} + 3n_{3,4} + 5n_{3,1}, \qquad \varphi_{\varphi,3} = \varphi_{3,3} + \varphi_{3,4} - \varphi_{3,1}, \tag{11.10}$$

respectively. Otherwise, the implementations are completely equivalent.

11.2.2 Nonlinear Activation Function of Second-Layer Neuron

The FWM neuron in the output layer has to be implemented using a modified algorithm as, in contrast to the neurons in the hidden layer, the total input particle number will not be fixed. To maximize the output of the FWM neuron in the output layer in terms of $\tilde{n}_{4,3}$ for the case of maximum particle number input, $n_{4,1} = n_{4,2} = n_{4,4} = 0.43$, the implementation duration $\tau_{\mathrm{F},4}$ of this neuron is redetermined as

$$\tau_{\mathrm{F},4} = T_4/2, \tag{11.11}$$

where T_4 is the oscillation period for the given configuration. The resulting FWM oscillation can be seen in Fig. 11.10. Considering the same experimental parameters as in Section 10.1.2 yields

$$t_{\mathrm{F},4} \simeq 23.4\,\mu\mathrm{s}. \tag{11.12}$$

The nonlinear activation function of a FWM neuron in the second layer of a 9-3-1-network can be seen in Fig. 11.11. Again, the output particle number $\tilde{n}_{4,3}$ is independent of the input phases $\varphi_{4,j}$. As was the case before, we were not able to extract an analytical expression from the data. However, the exchange symmetry regarding $n_{4,1}$ and $n_{4,2}$ persists. The output phase $\tilde{\varphi}_{4,3}$ shows the same linear dependence from (10.12) with

$$n_{\varphi,4} = 3n_{4,1} + 3n_{4,2} + 5n_{4,4}, \qquad \varphi_{\varphi,4} = \varphi_{4,1} + \varphi_{4,2} - \varphi_{4,4}, \tag{11.13}$$

and parameters $s = (-0.97 \pm 0.01)$ and $d = (3.50 \pm 0.04)$ determined via a least square fit. The numerical results can be used to determine the partial derivatives $\partial\,\tilde{n}_{4,3}/\partial\,n_{4,j}$, $\partial\,\tilde{\varphi}_{4,3}/\partial\,n_{4,j}$ and $\partial\,\tilde{\varphi}_{4,3}/\partial\,\varphi_{4,j}$ needed to train the network using the backpropagation algorithm.

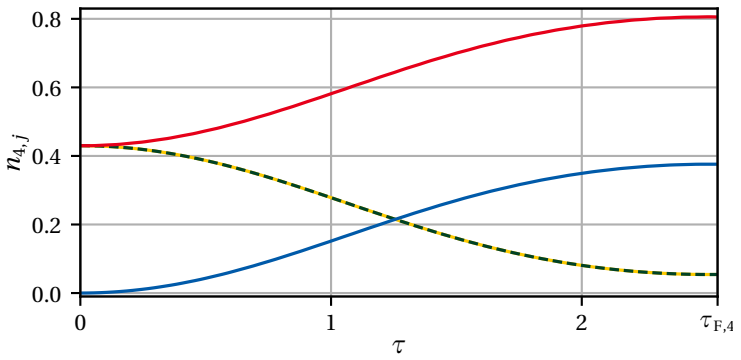

Figure 11.10: Dynamics of population probabilities $n_{4,j}$ versus dimensionless time τ for a FWM neuron in the second layer of a network. The initial state is given by $n_{4,1} = n_{4,2} = n_{4,4} = 0.43$, while $\varphi_{4,j} = 0$ (α_1: green, dashed; α_2: yellow, solid; α_3: blue; α_4: red). The neuron duration $\tau_{\mathrm{F}4}$ is determined as a half-oscillation period leading to maximal response.

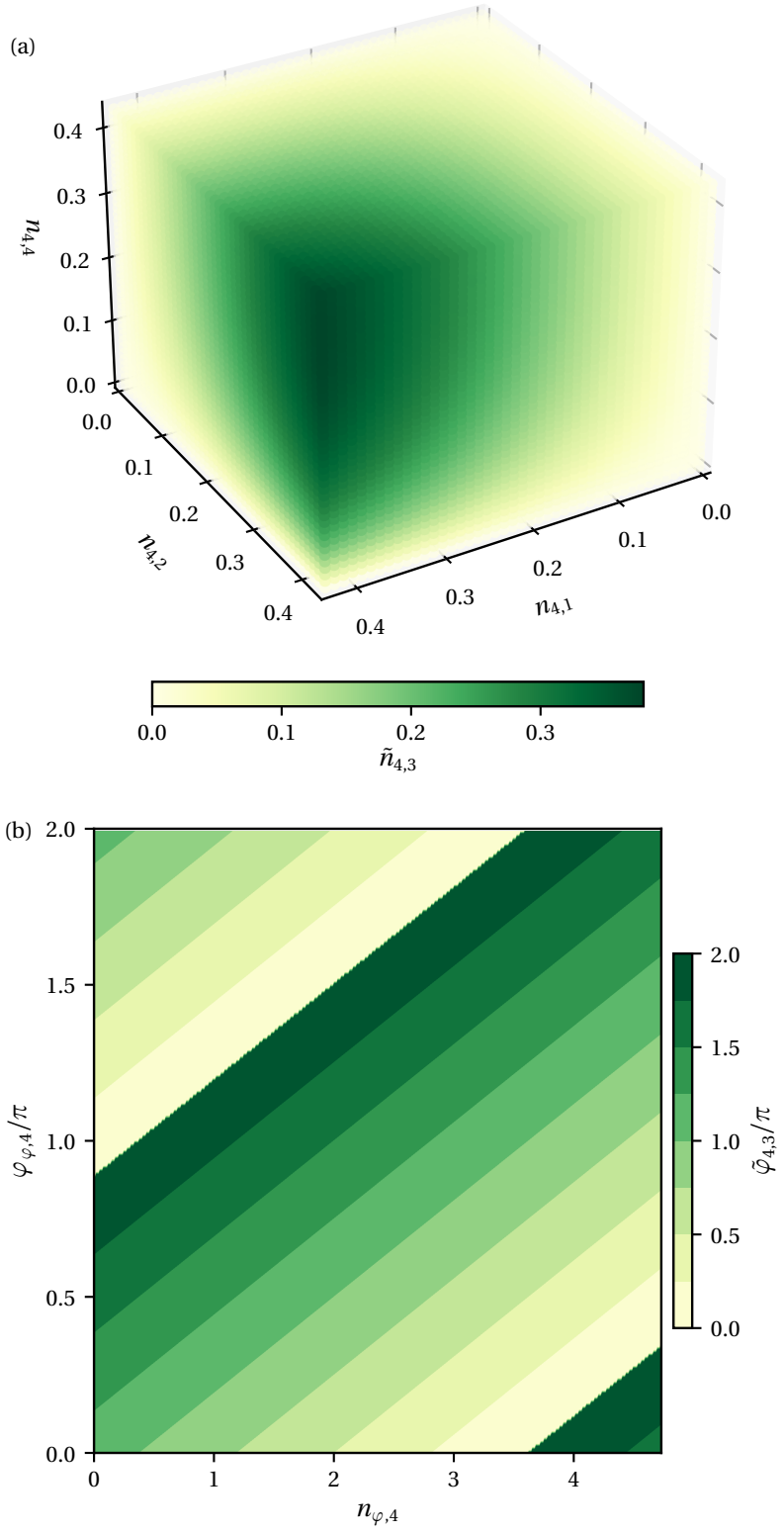

Figure 11.11: Nonlinear activation function of a FWM neuron in the second layer of a 9-3-1-network in terms of (a) $\tilde{n}_{4,3}$ versus $n_{4,1}$, $n_{4,2}$ and $n_{4,4}$ and (b) $\tilde{\varphi}_{4,3}$ versus $n_{\varphi,4}$ and $\varphi_{\varphi,4}$.

11.2.3 Learning Capability of Four-Wave Mixing Neural Network

To test the learning capabilities of the FWM neural network described in the previous section, an exemplary task is presented to the network. A six-dimensional input to the network is given by a sequence of six logical bits $x_i \in [0,1]$, $i = 1,\ldots,6$. The goal of the training procedure is to perform classification tasks such as "Has the sequence an even/odd number of ones?" or other examples of that kind. To do so, the network is trained to distinguish the different possible input patterns. The complexity of the training is reduced by restricting inputs, weights and outputs to real values given by relative particle numbers n. The presented procedure is a proof of concept and readily expandable to complex numbers.

To make use of the full range of the nonlinear activation functions of the neurons, a logical input 0 is encoded as a particle number $n = 0.35$, while a logical 1 is given by $n = 0.45$. As input channels for the FWM neurons in the hidden layer, those relative particle numbers are chosen under which the output particle numbers of the respective neurons show an exchange symmetry. Hence, the input channels are $n_{1,1}$, $n_{1,2}$, $n_{2,3}$, $n_{2,4}$, $n_{3,3}$ and $n_{3,4}$. Also, the desired output values $\tilde{n}_{4,3}$ have to be chosen carefully to make use of the nonlinear activation function of the FWM neuron in the second layer. The complete encoding scheme used for training and testing of the network can be seen in Appendix D.

The network is trained using the backpropagation algorithm from Section 3.2, with the update rule for the absolute value of a complex weight from Section 10.2, in accordance with the proposed real-valued implementation using relative particle numbers. On-line training is performed over 1,000 epochs with 1,000 random samples per epoch. The learning rate is set to $\eta_a = 10^{-3}$ for the first 100 epochs, to $\eta_a = 5 \cdot 10^{-4}$ for epochs 100 through 500, to $\eta_a = 2.5 \cdot 10^{-4}$ for epochs 500 through 850 and to $\eta_a = 10^{-4}$ for the final 150 epochs. After each epoch, the performance of the network in distinguishing the different bit sequences is evaluated by calculating the averaged squared error $\mathcal{E}$ according to (3.2) for $m = 1{,}000$ random samples.

As can be seen in Fig. 11.12, the FWM neural network is able to improve its performance in distinguishing the presented samples via backpropagation training. Starting from random weights, the averaged squared error is reduced to $\mathcal{E} = 9.01 \cdot 10^{-6}$ after 1,000 epochs. By increasing its performance, the network can be applied to a variety of classification tasks.

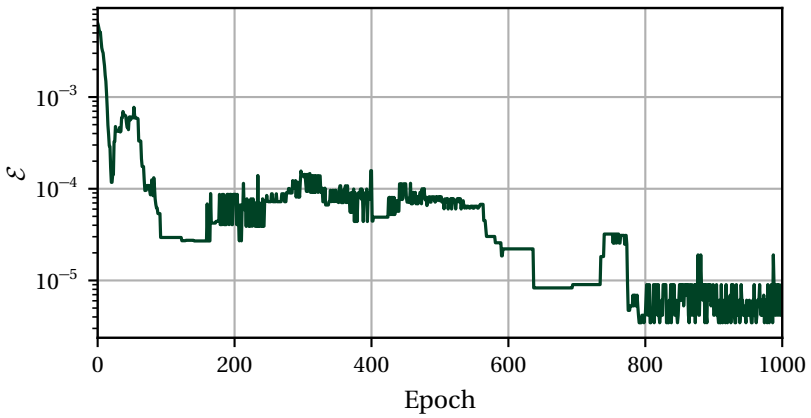

Figure 11.12: Averaged squared error $\mathcal{E}$ (3.2) over $m = 1{,}000$ random samples versus epochs for training the FWM neural network to distinguish a logical 6-bit sequence.

11.2.4 Discussion

In the pursuit of setting up a FWM neural network, we investigated different possible ways to achieve this. While spatially separating FWM neurons yields a perfect parallelization possibility, this same spatial separation prevents transport of output information of one layer to the next and therefore the establishment of a feedforward network.

In contrast, FWM configurations stacked in momentum space do not show spatial separation and overlap perfectly. However, interactions between the stacked configurations are non-negligible. Accordingly, the FWM dynamics are influenced by the presence of stacked configurations, preventing parallelization in the sense of neural networks. Yet, the presented constellation might still find an application in the realm of machine learning, specifically in reservoir computing. There, one only has to have control over readout weights and recurrent interactions inside the physical system might actually be beneficial. As the FWM setup stacked in momentum space possesses clearly defined input and output channels, it might be a good candidate to implement reservoir computing.

A suitable method yielding independent parallel FWM neurons and the possibility to transport output information to the next layer is given by using multiple internal states of the BEC. In an ideal scenario, the intra-component scattering lengths a_{ii} are equal for all states constituting the network and all inter-component scattering lengths a_{ij} vanish. As noted before, we assume such a configuration can be found in the extended internal structure of a BEC isotope or can be artificially produced using, for example, Feshbach resonances. Given one finds such a configuration, we are able to set up a tree-like feedforward neural network with six inputs and one output using four internal states. The backpropagation algorithm can be used to train this network to distinguish different patterns of a logical 6-bit sequence, demonstrating the learning capabilities of the network. This example serves as a proof of concept that FWM neurons can indeed be embedded and trained in a network structure, yielding the FWM neural network.

Part IV

SILENT WHITE LIGHT

12

EMISSION PROPERTIES OF QUANTUM DOT SUPERLUMINESCENT DIODES

Characteristic emission properties have made superluminescent diodes (SLDs) an attractive light source for optical implementations of artificial neural networks [19–23]. Their combination of high output power, broad optical spectra and spatially directed emission can be utilized efficiently in such setups.

In this thesis, we aim to aid the investigation of such light sources, and in particular quantum dot superluminescent diodes (QDSLDs), by investigating first- and second-order temporal correlation properties of these emitters. Especially their behaviour of showing temperature-dependent intensity noise suppression, as demonstrated by *Blazek et al.* [24], is of interest. Adding suppressed intensity noise to the characteristics of such diodes, implying emission which is first-order incoherent and second-order coherent, introduces a new class of light sources.

In Chapter 12 we investigate the emission properties of QDSLDs at room temperature. We show, that first- and second-order temporal correlation properties of these diodes can be described by a superposition of independent stochastic emitters, enabling access to statistical properties of the emitted light. Subsequently, in Chapter 13, we investigate the transformation of photon statistics of the QDSLD emission through interaction with a pumped atomic three-level-system. Introducing temperature dependence into our description via measured output intensities, we are able to reproduce experimentally observed intensity noise suppression.

12.1 QUANTUM DOT SUPERLUMINESCENT DIODES

QDSLDs are semiconductor-based light sources that can be classified between coherent laser diodes (LDs) and incoherent light emitting diodes (LEDs). Using specialized waveguide geometries and an active medium consisting of inhomogeneously broadened InAs/InGaAs quantum dot (QD) layers, they are able to combine high output intensities, spatially directed emission and spectral widths in the THz regime [230].

The setup of QDSLDs is very comparable to all other sorts of opto-electronic semiconductor light sources. The active medium of the diode is positioned inside a waveguide and is embedded in a forward biased p-n-junction. Under the emission of a photon, electrons from the n-doped region and holes from the p-doped region recombine in the active region of the material [147]. In contrast to LDs and LEDs, QDSLDs emit amplified spontaneous emission (ASE). In this process, an initially spontaneously emitted photon gets amplified due to stimulated emission along its propagation through the waveguide. The consequence is emission with large optical bandwidth, small coherence length, yet enhanced output power compared to LEDs [231].

Since being proposed by *Lee et al.* in 1973 [230], the emission properties of QDSLDs have become an intensely researched topic in recent years. This was made possible after *Boitier et al.* enabled the direct measurement of coherence times in the femtosecond regime using two-photon absorption in semiconductors in 2009 [232]. From a theoretical point of view, the understanding of light generation processes inside the diode is a main focus. Approaches utilized in those investigations include rate equation models [233–235], travelling wave approaches [236], finite element methods [237] and quantized treatments [238, 239]. Also experimentalists put a lot of effort into analyzing first- and second-order temporal correlation properties of QDSLDs [240, 241].

In 2011, *Blazek et al.* were able to observe a temperature-dependent suppression of intensity fluctuations using a broadband emitting QDSLD [24], thereby introducing the term "hybrid light". At around 190 K, they measured the intensity fluctuations in terms of the second-order correlation coefficient $g^{(2)}(0) = \langle I^2\rangle / \langle I\rangle^2$ to be $g^{(2)}(0) = 1.33$. Such emission being first-order incoherent (implying a broadband emission spectrum) and second-order coherent (implying reduced intensity fluctuations more similar to the photon statistics of a laser), might also be called "silent white light". Previous theoretical investigations of this behaviour focused on the quantum nature of the diode material [238, 242].

In this chapter, based on [37], we analyze the emission properties of QDSLDs in terms of optical emission spectrum and the second-order temporal correlation function by means of a stochastic model of the diode emission. For this, a general approach of modelling arbitrary emission spectra via independent stochastic processes is introduced in Section 12.2. Subsequently, this model is applied to the QDSLD in Section 12.3, also determining the intensity distribution and second-order temporal correlation function of the emission.

12.2 STOCHASTIC MODELLING OF EMISSION SPECTRA

12.2.1 Complex Ornstein-Uhlenbeck Process

The classical electric field emitted by a diode is modelled as a superposition of stochastic fields. Therefore, the electric field outside of the diode can be written as

$$\varepsilon_{\mathrm{d}}(t) = \sum_{m=1}^{M} \varepsilon_m(t), \tag{12.1}$$

where M is the number of stochastically fluctuating fields involved in the emission and $\varepsilon_m(t)$ is the complex field amplitude of the m-th field.

An individual classical field $\varepsilon \in \mathbb{C}$ is modelled as a complex Ornstein-Uhlenbeck process [243], which is described by the Ito stochastic differential equation (see Section 5.1.1) [116]

$$\mathrm{d}\varepsilon(t) = (\mathrm{i}\nu_0 - \gamma)\varepsilon(t)\mathrm{d}t + \sqrt{\gamma I}\mathrm{d}W(t), \tag{12.2}$$

where ν_0 is the carrier frequency of the field, γ is the linewidth, $I = \lim_{t\to\infty}\langle|\varepsilon(t)|^2\rangle$ is the mean intensity of the electric field and $\mathrm{d}W(t) \in \mathbb{C}$ is a complex Wiener noise increment with $\langle \mathrm{d}W(t)\rangle = 0$ and $\langle|\mathrm{d}W(t)|^2\rangle = \mathrm{d}t$. Due to the complex nature of the quantities, this description is identical to a two-dimensional Ornstein-Uhlenbeck process for the real and imaginary parts of the electric field amplitude.

To calculate temporal correlation properties of the electric field, the definitions and calculation rules for stochastic integrals introduced in Section 5.3 are used. The stationary first-order temporal correlation function of the process reads [116, 244]

$$G^{(1)}(\tau) = \lim_{t\to\infty} \left\langle \varepsilon^*(t)\varepsilon(t+\tau)\right\rangle = I e^{-\gamma|\tau| - \mathrm{i}\nu_0\tau}. \tag{12.3}$$

In accordance to the Wiener-Khintchine theorem [245, 246], the spectral power density is given by the Fourier transform of the stationary first-order temporal correlation function. Accordingly, the spectral power density for the complex Ornstein-Uhlenbeck process can be calculated as

$$S(\nu) = \sqrt{\frac{2}{\pi}} \frac{I\gamma}{(\nu-\nu_0)^2+\gamma^2}. \tag{12.4}$$

The spectral power density shows a Lorentzian shape and is normed to the mean intensity

$$I = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}\nu\, S(\nu). \tag{12.5}$$

The stationary normalized second-order temporal correlation function can be determined using the Siegert relation as [244, 247]

$$g^{(2)}(\tau) = \lim_{t\to\infty} \frac{\langle \varepsilon^*(t)\varepsilon^*(t+\tau)\varepsilon(t+\tau)\varepsilon(t)\rangle}{\langle \varepsilon^*(t)\varepsilon(t)\rangle \langle \varepsilon^*(t+\tau)\varepsilon(t+\tau)\rangle} = 1 + \left|\frac{G^{(1)}(\tau)}{G^{(1)}(0)}\right|^2 = 1 + e^{-2\gamma|\tau|}. \tag{12.6}$$

12.2.2 Stochastic Simulation

In addition to analytical investigations, numerical simulations of (12.2) can be used to determine temporal correlation features of the emission field. For the benefit of efficient simulation, the rapidly oscillating carrier frequency is separated from the dynamics of the field via the transformation $\varepsilon(t) = \eta(t)e^{-\mathrm{i}\nu_0 t}$. This yields the Ito stochastic differential equation

$$\mathrm{d}\eta(t) = -\gamma\eta(t)\mathrm{d}t + \sqrt{\gamma I}\,\mathrm{d}W(t). \tag{12.7}$$

As the diffusion coefficient $\sqrt{\gamma I}$ is independent of the electric field amplitude η itself, the Euler scheme can be used to achieve strong convergence of order 1.0 (see Appendix A.2). Therefore, an iterative time-discrete simulation scheme for the electric field amplitude can be constructed as [127]

$$\eta(t_{i+1}) = \eta(t_i) - \gamma\eta(t_i)\Delta t + \sqrt{\gamma I}\,\Delta W. \tag{12.8}$$

The discrete time step of the simulation scheme is $\Delta t = t_{i+1} - t_i$. The term $\Delta W \in \mathbb{C}$ symbolizes a complex Gaussian random process with mean $\langle \Delta W\rangle = 0$ and variance $\langle |\Delta W|^2\rangle = \Delta t$.

By sampling over N realizations of the electric field amplitude, the first-order temporal correlation function can be computed using this simulation scheme as

$$G^{(1)}(\tau) = \frac{1}{N}\sum_{n=1}^{N} \varepsilon_n^*(t_\mathrm{s})\varepsilon_n(t_\mathrm{s}+\tau), \tag{12.9}$$

where ε_n is the n-th realization of the field and $t_\mathrm{s} \gg 1/\gamma_n$ is a point in time, long after an initial transient phase is surpassed. Using a Fast Fourier Transform, (12.9) can be used to calculate the spectral power density of the field.

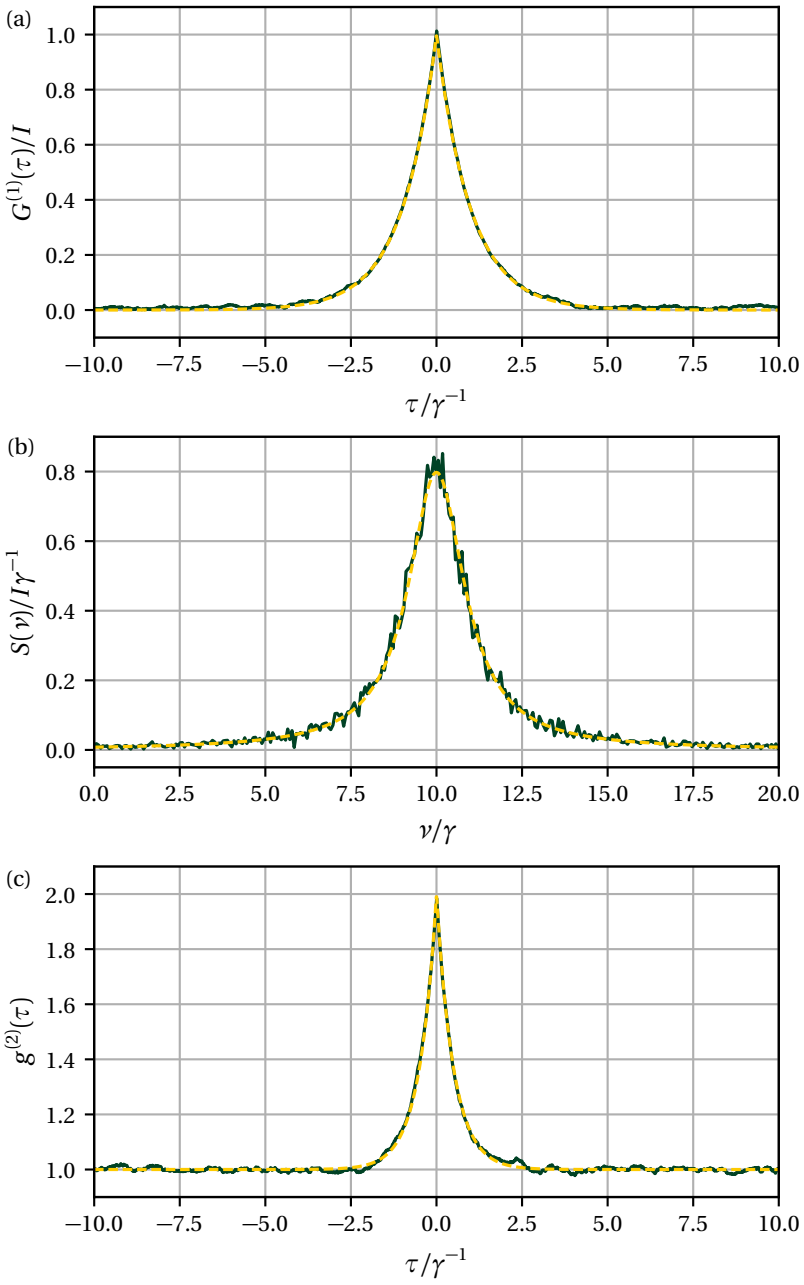

Figure 12.1: (a) First-order temporal correlation function $G^{(1)}(\tau)$ versus time τ, (b) spectral power density $S(\nu)$ versus frequency ν for central frequency $\nu_0 = 10\gamma$ and (c) normalized second-order temporal correlation function $g^{(2)}(\tau)$ versus time τ for the electric field described by (12.2). Simulation results (green, solid) match the analytical expression from (12.3), (12.4) and (12.6) (yellow, dashed). The numerical simulations were performed with $\Delta t = 0.01/\gamma$, while the correlation functions were calculated using $N = 10{,}000$ realizations.

The normalized second-order temporal correlation function (12.6) can be determined by separate calculations for the denominator and the numerator. While the first-order temporal correlation functions in the denominator can be simulated according to (12.9), the second-order correlation function in the numerator can be calculated as

$$G^{(2)}(\tau) = \frac{1}{N}\sum_{n=1}^{N} |\varepsilon_n(t_\mathrm{s}+\tau)\varepsilon_n(t_\mathrm{s})|^2. \tag{12.10}$$

The normalized second-order temporal correlation function results as

$$g^{(2)}(\tau) = \frac{G^{(2)}(\tau)}{G^{(1)}(0)G^{(1)}(\tau)}. \tag{12.11}$$

A comparison between analytical expression and simulation results of the first- and second-order temporal correlation properties of an individual field can be seen in Fig. 12.1.

12.2.3 Modelling Arbitrary Spectra as Sums of Lorentzians

All individual classical electric fields ε_m contributing to a total emission according to (12.1) are assumed to be independent and to be described by a complex Ornstein-Uhlenbeck process according to (12.2). Meanwhile, each individual field is characterized by a specific linewidth γ_m, mean intensity I_m and central frequency ν_m. As all fields are independent, the stationary first-order temporal correlation function is given by

$$G_\mathrm{d}^{(1)}(\tau) = \lim_{t\to\infty}\left\langle \varepsilon_\mathrm{d}^*(t)\varepsilon_\mathrm{d}(t+\tau)\right\rangle = \lim_{t\to\infty}\sum_{m=1}^{M}\left\langle \varepsilon_m^*(t)\varepsilon_m(t+\tau)\right\rangle. \tag{12.12}$$

Thus, the spectral power density is the incoherent sum of the individual spectra

$$S_\mathrm{d}(\nu) = \sum_{m=1}^{M} S_m(\nu). \tag{12.13}$$

For the case that the emission of a diode is described as the sum of complex Gaussian fields, this implies, that arbitrary emission spectra $S_\mathrm{t}(\nu)$ of the diode can be modelled as the sum of Lorentzian lineshapes with individual central frequencies, amplitudes and widths. For example, this can be done via the adjustment of the $3M$ free parameters of the model using a least square fit, minimizing the error functional

$$e = \sum_i (S_\mathrm{t}(\nu_i) - S_\mathrm{d}(\nu_i))^2 \tag{12.14}$$

at discrete frequencies ν_i. The capabilities of this approach are demonstrated on different important line shapes (all spectra are normalized to $\int_{-\infty}^{\infty}\mathrm{d}\nu\, S(\nu) = \sqrt{2\pi}$):

- Gaussian spectrum (center ν_0, standard deviation σ) [211]:

$$S_\mathrm{g}(\nu) = \frac{1}{\sqrt{\sigma^2}} e^{-\frac{(\nu-\nu_0)^2}{2\sigma^2}} \tag{12.15}$$

- Lorentzian spectrum (center ν_0, half width at half mean γ) [211]:

$$S_\mathrm{l}(\nu) = \sqrt{\frac{2}{\pi}}\frac{\gamma}{(\nu-\nu_0)^2+\gamma^2} \tag{12.16}$$

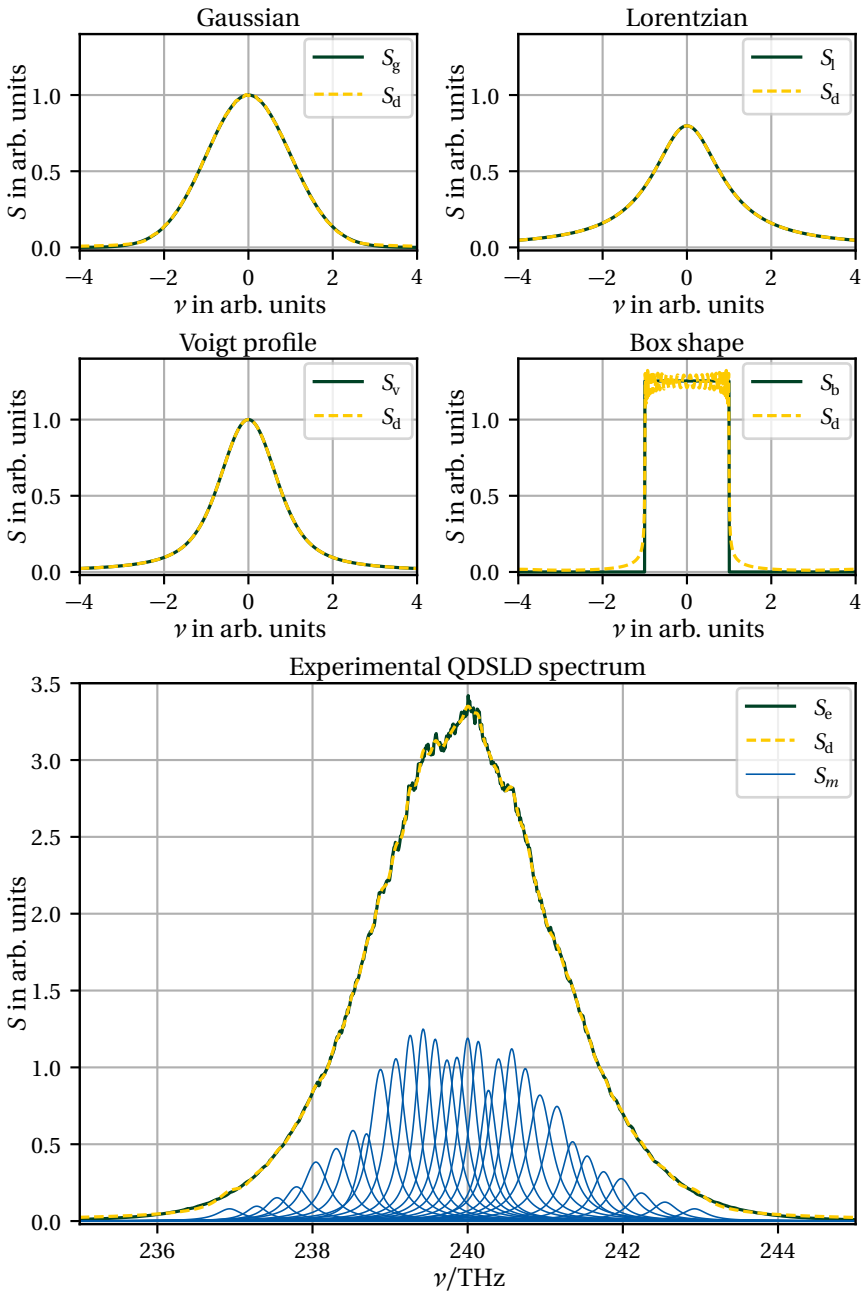

Figure 12.2: Gaussian spectrum $S_g(\nu)$ (12.15), Lorentzian spectrum $S_l(\nu)$ (12.16), Voigt profile $S_v(\nu)$ (12.17), bandwidth limited box shape $S_b(\nu)$ (12.18) and experimental optical power spectrum of QDSLD $S_e(\nu)$ (experimental data from [248, 249]; all spectra green, solid) versus frequency ν. All spectra are modelled with $M = 30$ individual Lorentzian-shaped Ornstein-Uhlenbeck spectra (yellow, dashed). For the experimental spectrum, the individual lineshapes contributing to the model are illustrated (blue, solid).

- Voigt profile ($z=(\nu+\mathrm{i}\gamma)/\sigma\sqrt{2}$, complementary error function $\mathrm{erfc}(z)$) [211]:

$$S_{\mathrm{v}}(\nu)=\frac{1}{\sqrt{\sigma^2}}\mathrm{Re}\left(e^{-z^2}\mathrm{erfc}(-\mathrm{i}z)\right) \tag{12.17}$$

- bandwidth limited box shape

$$S_{\mathrm{b}}(\nu)=\begin{cases}\sqrt{2\pi}/\gamma, & \text{for } |\nu|\leq\gamma/2m\\ 0, & \text{else.}\end{cases} \tag{12.18}$$

Using this approach, also the optical power spectrum of the QDSLD can be modelled as the sum of individual Lorentzian lineshapes. The linewidths, mean intensities and central frequencies of $M=30$ individual Ornstein-Uhlenbeck spectra are determined using a least square fit. The results of modelling arbitrary spectra as well as the optical power spectrum of a QDSLD can be seen in Fig. 12.2.

12.3 EMISSION PROPERTIES OF QUANTUM DOT SUPERLUMINESCENT DIODES

Using the linewidths γ_m, mean intensities I_m and central frequencies ν_m gathered from modelling the optical power spectrum of a QDSLD [248, 249] as a sum of Lorentzian lineshapes, the temporal correlation properties and the statistics of the emission can be investigated using stochastic simulation. The individual fields ε_m are calculated numerically according to (12.8). The electric field emitted by the diode ε_{d} results as a superposition of the individual fields as described by (12.1). All statistical and temporal correlation properties are calculated using $N=10{,}000$ realizations.

12.3.1 Intensity Distribution

A first quantity of interest is the intensity distribution of the QDSLD emission. For each realization, the intensity $I(t)=|\varepsilon(t)|^2$ is calculated and the distribution after an initial transient period is determined. As can be seen in Fig. 12.3, numerical results agree very well with an exponential distribution

$$p(I)=\frac{1}{\bar{I}}e^{-I/\bar{I}}, \tag{12.19}$$

where $\bar{I}=\sum_{m=1}^{M}I_m$ is the sum of all individual mean intensities. This distribution follows consequently from the fact, that the complex Gaussian field, as the name suggests, shows a Gaussian distribution for the real and the imaginary part of the field, implying an exponential distribution of the intensity [116].

12.3.2 Optical Power Spectrum

The spectral power density of the emission $S_{\mathrm{d}}(\nu)$ is determined using the Fourier transform of the stationary first-order temporal correlation function $G_{\mathrm{d}}^{(1)}(\tau)$ calculated according to (12.12). The results of the simulation (see Fig. 12.4) show good agreement with the experimental optical power spectrum [248, 249].

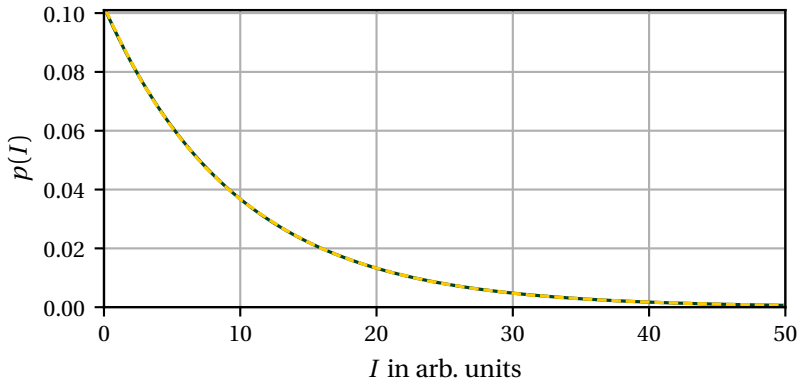

Figure 12.3: Probability density $p(I)$ versus intensity I of the QDSLD emission according to (12.19) (green, solid) and calculated numerically using stochastic simulation (yellow, dashed).

The width of an optical power spectrum can be defined via the Süssmann measure [250, 251]

$$b = \frac{1}{\int_0^\infty \mathrm{d}\nu\, s^2(\nu)}, \tag{12.20}$$

where $s(\nu) = S(\nu)/\int_0^\infty \mathrm{d}\nu\, S(\nu)$ is the normed spectral power density. This yields $b_\mathrm{d} = 4.85\,\mathrm{THz}$, implying a coherence time of $\tau_{\mathrm{c,d}} = 1/b_\mathrm{d} = 206.34\,\mathrm{fs}$, which matches experimental results of $b_\mathrm{e} = 4.29\,\mathrm{THz}$ and $\tau_\mathrm{e} = 233\,\mathrm{fs}$ [248] very well. The method of modelling emission spectra as a superposition of individual oscillators is therefore suitable to describe the first-order temporal correlation properties of QDSLDs.

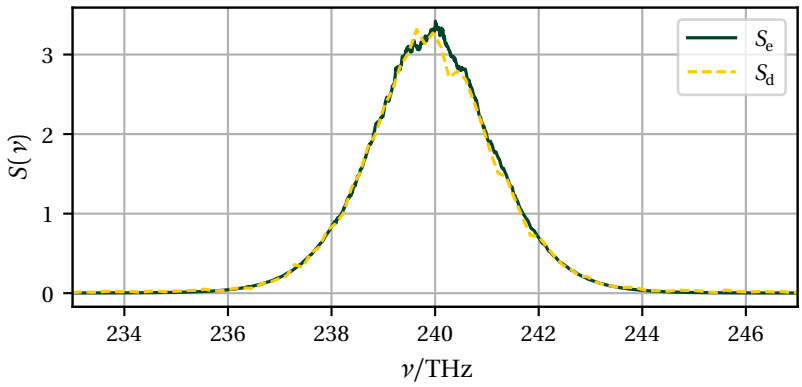

Figure 12.4: Experimental optical power spectrum $S_\mathrm{e}(\nu)$ (green, solid) versus frequency ν and simulation results $S_\mathrm{d}(\nu)$ (yellow, dashed) for $M = 30$ individual oscillators resulting as the Fourier transform of the stationary first-order temporal correlation function $G_\mathrm{d}^{(1)}(\tau)$ calculated using $N = 10,000$ realizations.

12.3.3 Second-Order Temporal Correlation Function

In addition to the investigation of the optical power spectrum, the classical photon statistics of the QDSLD emission can be investigated using the stochastic model. For this, the stationary normalized second-order temporal correlation function $g_{\mathrm{d}}^{(2)}(\tau)$ is calculated according to (12.11).

As can be seen in Fig. 12.5, the simulation shows qualitative agreement with experimental data [248, 249]. The central degree of second-order temporal coherence $g_{\mathrm{d}}^{(2)}(\tau = 0) \simeq 2$ implies a Gaussian photon distribution, which was also shown experimentally. Therefore, the described formalism of modelling QDSLD emission as a superposition of individual stochastic oscillators is also suited for the investigation of classical photon statistics.

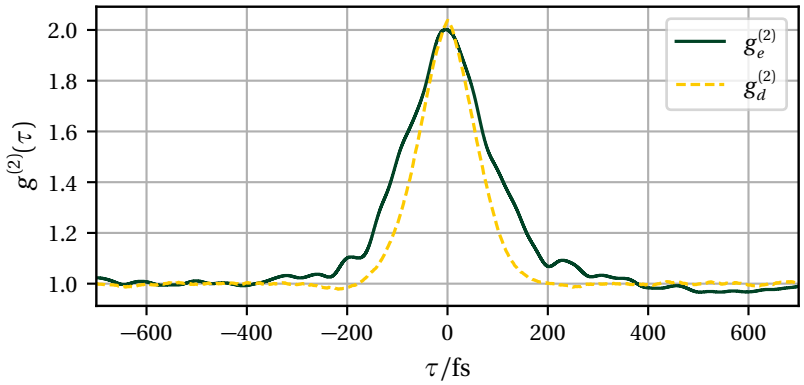

Figure 12.5: Experimental stationary normalized second-order temporal correlation function $g_{\mathrm{e}}^{(2)}(\tau)$ (green, solid) versus time τ [248, 249] and simulation results $g_{\mathrm{d}}^{(2)}(\tau)$ (yellow, dashed) using $M = 30$ individual oscillators and $N = 10{,}000$ realizations of the electric field.

13

TEMPERATURE-DEPENDENT INTENSITY NOISE SUPPRESSION

In this chapter, based on [38], we introduce a theoretical investigation of the temperature-dependent intensity noise suppression in QDSLDs as reported in [24] by *Blazek et al.*. The results from Chapter 12 regarding the statistical properties and temporal coherence characteristics are taken as reference values at room temperature. By introducing an interaction between the complex electric field propagating through the diode and a system of pumped three-level atoms in Section 13.1, the statistical properties of the emission can be manipulated via saturation inside the medium as described in Section 13.2. This includes the classical photon statistics as described by the central degree of second-order coherence

$$g^{(2)}(0) = \frac{\langle I^2 \rangle}{\langle I \rangle^2}. \tag{13.1}$$

In the sense of $g^{(2)}(0)$, varying degrees of coherence can be defined [252]. While $g^{(2)}(0) = 2.0$ indicates incoherent (thermal) light, the corresponding value indicating coherent emission is given by $g^{(2)}(0) = 1.0$. In the range between these two values, one speaks of partially coherent light. Exceeding the two boundaries the terms superbunched light for $g^{(2)}(0) > 2.0$ and antibunched light for $g^{(2)}(0) < 1.0$ can be introduced.

In Section 13.3, we show that by extracting the temperature-dependent features of the system from experimental measurements of the output intensity, we are able to reproduce the intensity noise suppression. Hence, we can interpret the reduction of the central degree of second-order coherence $g^{(2)}(0)$ from 2.0 to around 1.33 as a temperature driven saturation effect.

13.1 PUMPED THREE-LEVEL ATOM INTERACTING WITH COHERENT LIGHT FIELD

13.1.1 Diode Model

According to the Maxwell-Bloch equations [128, 253, 254], the complex electric field amplitude after travelling through a polarizable medium for a distance L along the z-direction is given by

$$\mathcal{E}(L,t) = \mathcal{E}(0, t - L/c) + \frac{\mathrm{i}k}{2\varepsilon_0} \int_0^L \mathrm{d}z \, \mathcal{P}^{(+)}(z, t - (L+z)/c), \tag{13.2}$$

where k is the wave number of the electric field, $\boldsymbol{\mathcal{P}}(\boldsymbol{r},t) = \boldsymbol{\mathcal{P}}^{(+)}(\boldsymbol{r},t) + \boldsymbol{\mathcal{P}}^{(-)}(\boldsymbol{r},t)$ is the complex polarization density and $\mathcal{P}^{(+)}(\boldsymbol{r},t) = |\boldsymbol{\mathcal{P}}^{(+)}(\boldsymbol{r},t)|$ is the positive frequency part of the polarization

density. Here, we assume, that $\mathcal{P}^{(+)}(\boldsymbol{r},t)$ is orientated parallel to the polarization vector of the incident electric field. Assuming the interaction takes place in an infinitesimally thin layer and taking into account a loss rate $0<\gamma<1$, which arises due to scattering and coupling out of the diode, an input-output-relation for the complex electric field amplitude can be formulated as

$$\varepsilon_{\text{out}}(t)=\eta\varepsilon_{\text{in}}(t)+\frac{\mathrm{i}kL}{2\varepsilon_0}P^{(+)}(t), \tag{13.3}$$

where $\eta=1-\gamma$ and $P^{(+)}(t)=|\boldsymbol{P}^{(+)}(t)|$ is the positive frequency part of the polarization. Due to the specialized waveguide geometry and anti-reflection coated facets of the QDSLD, no feedback mechanism has to be considered [248].

The QDs inside the active medium of the diode are modelled as pumped three-level systems (see Fig. 13.1). The incident electric field coherently drives the transition between levels $|1\rangle$ and $|2\rangle$ with the Rabi frequency

$$\Omega(t)=\frac{d_{21}\varepsilon_{\text{in}}(t)}{\hbar}, \tag{13.4}$$

where $d_{21}=|\boldsymbol{d}_{21}|=\langle 2|\hat{\boldsymbol{d}}|1\rangle$ is the dipole matrix element of this transition and $\hat{\boldsymbol{d}}$ is the dipole operator. Again, $\hat{\boldsymbol{d}}$ is assumed to be orientated parallel to the incident light field.

To determine the polarization $P(t)$, consider a volume V inside the diode, which is much smaller than the wavelength of the incident light ($V<\lambda^3$) and contains N QDs. The polarization density inside of V is given by [190]

$$\hat{\mathcal{P}}(\boldsymbol{r},t)=\frac{1}{V}\sum_{j=1}^{N}\hat{d}_j\delta(\boldsymbol{r}-\boldsymbol{r}_j), \tag{13.5}$$

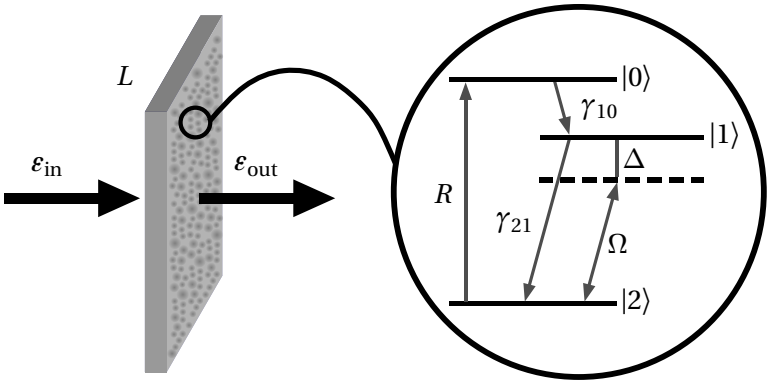

Figure 13.1: Schematic depiction of an incident complex electric field ε_{in} interacting with a layer of diode material of thickness L producing ε_{out}. The QDs inside the diode are modelled as pumped three-level systems with levels $|0\rangle$, $|1\rangle$ and $|2\rangle$. The transition between levels $|0\rangle$ and $|2\rangle$ is incoherently pumped with rate R. The transition between $|1\rangle$ and $|2\rangle$ is driven coherently with Rabi frequency $\Omega(t)$ and detuning $\Delta=\omega-\omega_{12}$, where ω and ω_{12} are the angular frequencies of the driving source and the transitions respectively. γ_{10} and γ_{21} are the decay rates of the respective transitions.

where $\hat{d}_j$ and $\boldsymbol{r}_j$ are the dipole operator and position of the j-th QD respectively. As $V < \lambda^3$, all QDs react to the same incident field and the density operator $\hat{\rho}$, describing the populations and coherences of the QDs, factorizes as $\hat{\rho} = \prod_{j=1}^{N} \hat{\rho}_j$. Therefore, the polarization yields

$$\begin{aligned} P(t) &= \frac{1}{V}\int_V \mathrm{d}^3 r \, \left\langle \hat{\mathcal{P}}(\boldsymbol{r},t)\right\rangle = \frac{N}{V}\int_V \mathrm{d}^3 r \, \mathrm{Tr}\left\{\hat{\rho}(t)\hat{d}\right\} \\ &= n(d_{12}\rho_{21}(t) + d_{21}\rho_{12}(t)), \end{aligned} \tag{13.6}$$

where n is the density of interacting QDs and $\rho_{ij} = \langle i|\hat{\rho}|j\rangle$ are the elements of the individual density matrix. Accordingly, the positive frequency part of the polarization can be identified as

$$P^{(+)}(t) = n d_{12}\rho_{21}(t). \tag{13.7}$$

13.1.2 Optical Bloch Equations

To calculate $\rho_{21}(t)$, the dynamics of the QDs have to be determined. For the given pumped three-level system, these are described by the optical Bloch equations [128]

$$\begin{aligned} \dot{\rho}_{00} &= -(R+\gamma_{10})\rho_{00} + R\rho_{22}, \\ \dot{\rho}_{11} &= \gamma_{10}\rho_{00} - \gamma_{21}\rho_{11} - \frac{\mathrm{i}}{2}\Omega^*\rho_{21} + \frac{\mathrm{i}}{2}\Omega\rho_{21}^*, \\ \dot{\rho}_{22} &= R\rho_{00} + \gamma_{21}\rho_{11} - R\rho_{22} + \frac{\mathrm{i}}{2}\Omega^*\rho_{21} - \frac{\mathrm{i}}{2}\Omega\rho_{21}^*, \\ \dot{\rho}_{21} &= -\left(\mathrm{i}\Delta + \frac{\Gamma}{2}\right)\rho_{21} - \frac{\mathrm{i}}{2}\Omega(\rho_{11}-\rho_{22}), \end{aligned} \tag{13.8}$$

where $\Gamma = \gamma_{21} + R$. As the coherence ρ_{21} decays much faster than the populations [128], it can be adiabatically eliminated. In this limit, the coherence ρ_{21} is connected to the inversion $w = \rho_{11} - \rho_{22}$ via

$$\rho_{21} = -\frac{\mathrm{i}}{2}\Omega\mathcal{D}w, \tag{13.9}$$

where $\mathcal{D} = 1/(\mathrm{i}\Delta + \Gamma/2)$. Inserting this into (13.8) yields the rate equations

$$\begin{aligned} \dot{\rho}_{00} &= -(R+\gamma_{10})\rho_{00} + R\rho_{22}, \\ \dot{\rho}_{11} &= \gamma_{10}\rho_{00} - (\gamma_{21}+\zeta)\rho_{11} + \zeta\rho_{22}, \\ \dot{\rho}_{22} &= R\rho_{00} + (\gamma_{21}+\zeta)\rho_{11} - (R+\zeta)\rho_{22}, \end{aligned} \tag{13.10}$$

where $\zeta = |\Omega|^2\mathcal{L}/\Gamma$ and $\mathcal{L} = (\Gamma/2)^2/(\Delta^2 + (\Gamma/2)^2)$.

Due to the stochastic nature of $\varepsilon_{\mathrm{in}}(t)$, this characteristic is also imposed on the elements of the density matrix ρ_{ij}. Consequently, the equations (13.10) are Ito stochastic differential equations. However, this fact can be neglected in the presented investigation. By assuming that the atomic dynamics are much faster than the stochastically fluctuating electric field amplitude $\varepsilon(t)$, the latter can be kept constant for every calculation of the atomic parameters. Therefore, the populations of the QD-levels are assumed to be in instantaneous equilibrium and the stationary solution of the rate equations (13.10) can be used to determine the instantaneous atomic coherence as

$$\rho_{21} = -\frac{\mathrm{i}}{2}\Omega\frac{w_0\mathcal{D}}{1+s}. \tag{13.11}$$

There, $w_0 = (R(\gamma_{10}-\gamma_{21})-\gamma_{10}\gamma_{21})/(\Gamma\gamma_{10}+2R\gamma_{21})$ is the unsaturated inversion and $s = \mathcal{L}I_{\text{in}}/I_{\text{s}}$ is the saturation parameter, where $I_{\text{in}} = |\varepsilon_{\text{in}}|^2$ and $I_{\text{s}} = \hbar^2\Gamma(\Gamma\gamma_{10}+2R\gamma_{21})/|d_{21}|^2(3R+2\gamma_{10})$ is the saturation intensity.

13.1.3 Intensity Input-Output-Relation

In this case, the input-output-relation (13.3) loses its time-dependence and can be expressed using atomic parameters as [254]

$$\varepsilon_{\text{out}} = (\eta+\alpha)\varepsilon_{\text{in}}, \tag{13.12}$$

with the complex nonlinear absorption coefficient $\alpha = \kappa\Gamma\mathcal{D}/(1+s)$, the nonlinear saturation parameter $\kappa = \alpha_0 kL$ and the linear, resonant absorption coefficient $\alpha_0 = n|d_{21}|^2 w_0/4\hbar\varepsilon_0\Gamma$. In the limit of a thin interaction layer, all terms exceeding linear order in κ can be neglected. Additionally, as a consequence of the broad THz-bandwidth of the emission, the detuning $\Delta = 0$ can be chosen to vanish. The input-output-relation for the intensity $I_{\text{out}} = |\varepsilon_{\text{out}}|^2$ then reads

$$I_{\text{out}}(I_{\text{in}}) = \left(\eta^2 + \frac{4\eta\kappa}{1+s}\right) I_{\text{in}} + \mathcal{O}(\kappa^2). \tag{13.13}$$

Consequently, the input-output-relation only has three independent parameters η, κ and I_{s}. As can be seen in Fig. 13.2, all parameter combinations produce nonlinear behaviour. While for some combinations, the input-output-relation is superlinear for small input intensities, they all show saturation and therefore sublinear behaviour for increasing input intensities.

13.2 PHOTON STATISTICS TRANSFORMATION

Due to the nonlinear connection between input and output intensity, implying saturation, the statistical properties of the output will differ from those of the input. As described in Section 12.3, the intensity distribution of the QDSLD emission is given by an exponential probability distribution according to (12.19), where the mean is $\langle I_{\text{in}}\rangle = \bar{I}$. The output photon

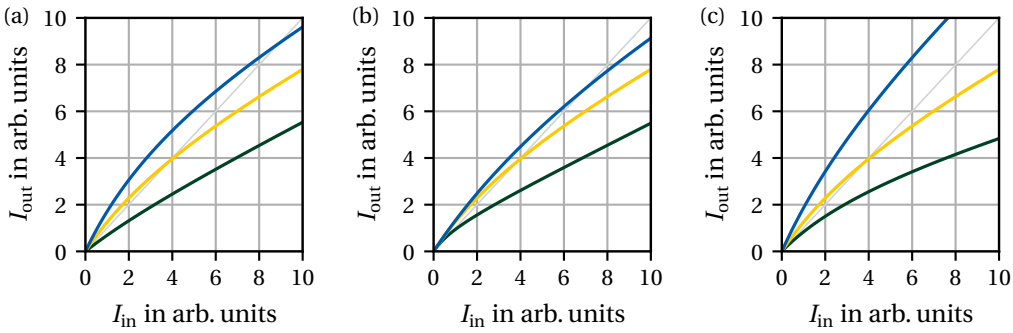

Figure 13.2: Output intensity I_{out} versus input intensity I_{in} according to (13.13) for varying (a) saturation parameter κ ($I_{\text{s}} = 5$, $\eta = 0.68$; green: $\kappa = 0.1$, yellow: $\kappa = 0.35$, blue: $\kappa = 0.55$), (b) saturation intensity I_{s} ($\kappa = 0.35$, $\eta = 0.68$; green: $I_{\text{s}} = 1$, yellow: $I_{\text{s}} = 5$, blue: $I_{\text{s}} = 9$) and (c) loss coefficient η ($I_{\text{s}} = 5$, $\kappa = 0.35$; green: $\eta = 0.5$, yellow: $\eta = 0.68$, blue: $\eta = 0.9$).

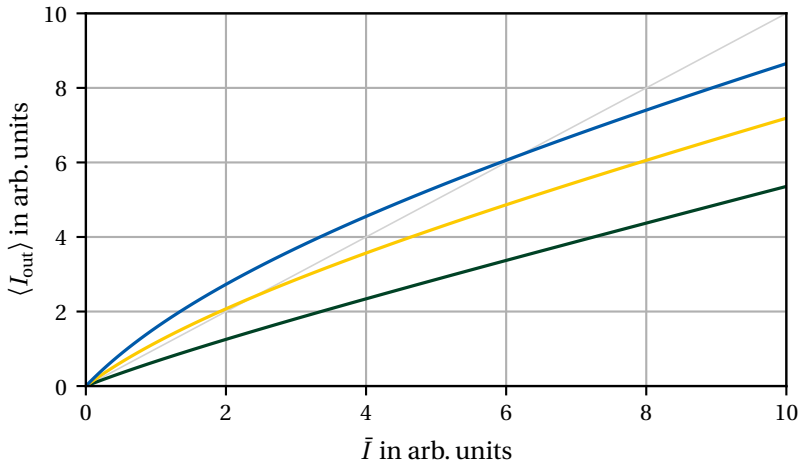

Figure 13.3: Mean output intensity $\langle I_{\text{out}} \rangle$ versus mean input intensity $\bar{I}$ for $\eta = 0.68$, $I_{\text{s}} = 5$ and $\kappa = 0.1$ (green), $\kappa = 0.35$ (yellow) and $\kappa = 0.55$ (blue).

statistics are determined by the moments of the transformed intensity distribution. The n-th moment, where $n \in \mathbb{N}$, is given by

$$\left\langle I_{\text{out}}^n(I_{\text{in}}) \right\rangle = \int_0^\infty \mathrm{d}I_{\text{in}}\, I_{\text{out}}^n(I_{\text{in}}) p(I_{\text{in}}). \tag{13.14}$$

Accordingly, the mean output intensity (see Fig. 13.3) can be calculated up to linear order in κ as

$$\langle I_{\text{out}} \rangle = \bar{I}\eta^2 + 4 I_{\text{s}} \eta \kappa [1 - \mathcal{I} u(\mathcal{I})], \tag{13.15}$$

where $\mathcal{I} = I_{\text{s}}/\bar{I}$, $u(\mathcal{I}) = e^{\mathcal{I}} \Gamma(0, \mathcal{I})$ and $\Gamma(a, z) = \int_z^\infty \mathrm{d}t\ t^{a-1} e^{-t}$ is the incomplete gamma function [211]. In the given situation, the interaction between light field and diode material takes place in a thin layer. Furthermore we concentrate the investigation on a spatial part of the diode near the output facet, where all transient behaviour can be neglected. Hence, gain through interaction and loss by means of scattering compensate each other and the mean intensity of the emission remains unchanged. This implies the self-consistent equilibrium condition

$$\langle I_{\text{out}} \rangle = \langle I_{\text{in}} \rangle = \bar{I}(\eta). \tag{13.16}$$

Using (13.15), the inaccessible loss rate $\eta(\bar{I})$ can be determined in favor of the equilibrium intensity $\bar{I}$. The central degree of second-order coherence $g_{\text{out}}^{(2)}(0)$ can be calculated analytically according to (13.1) in dependence of κ and I_{s}. Considering only contributions up to $\mathcal{O}(\kappa^2)$, one finds

$$g_{\text{out}}^{(2)}(0) = 2 - 8\kappa\mathcal{I}\left[1 + \mathcal{I}(1 - 2u(\mathcal{I})) - \mathcal{I}^2 u(\mathcal{I})\right]. \tag{13.17}$$

As can be seen in Fig. 13.4 the central degree of second-order coherence is reduced from an initial value of $g_{\text{out}}^{(2)}(0) = 2.0$ with increasing $\bar{I}$ for a set saturation intensity I_{s} and different positive values of κ. This implies, that intensity noise associated with $g_{\text{out}}^{(2)}(0)$ is suppressed while increasing $\bar{I}$, thereby entering a more saturated regime of the pumped three-level system. Hence, increasing saturation inside the diode suppresses intensity noise.

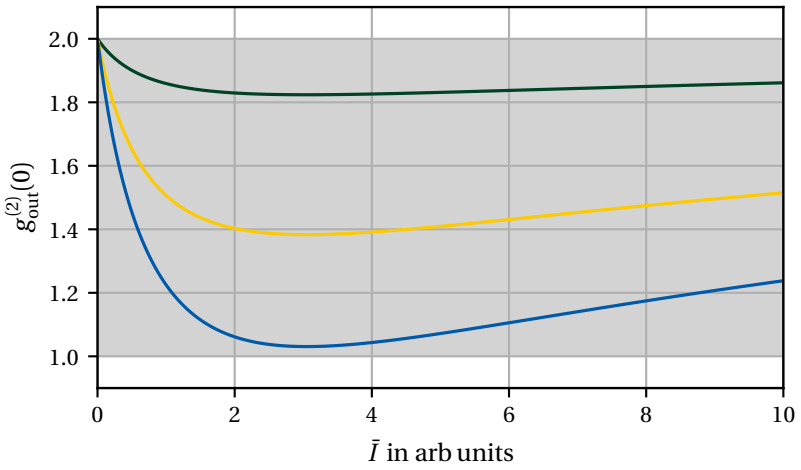

Figure 13.4: Central degree of second-order coherence $g^{(2)}_{\mathrm{out}}(0)$ versus mean input intensity $\bar{I}$ for $I_{\mathrm{s}} = 5$ and $\kappa = 0.1$ (green), $\kappa = 0.35$ (yellow) and $\kappa = 0.55$ (blue).

It has to be noted, that this effect is not present for negative κ, which would be the case if there is no inversion present in the medium. In fact, the central degree of second-order coherence surpasses $g^{(2)}_{\mathrm{out}}(0) = 2.0$ in this case, which implies that intensity fluctuations are actually enhanced if there is no inversion in the gain medium.

13.3 TEMPERATURE-DEPENDENT INTENSITY NOISE SUPPRESSION

The nature of the QDSLD opens up a lot of possibilities to introduce temperature dependence into this system via the semiconductor physics of the diode and the gain medium consisting of inhomogeneously broadened QDs. In essence, at around 190 K a maximum in radiative recombination occurs in the QD systems. This redistribution of carriers modifies the optical gain properties of the diode, which can be quantified via the measurement of the temperature-dependent mean intensity $\bar{I}$. At the temperatures of increased recombination, a peak in output intensity occurs, implying larger amplification and in turn affecting the photon emission properties. For further details on the semiconductor physics responsible for the temperature dependence of the QDSLD system refer to [38]. In this thesis, we focus on the quantum optical aspects of the observed intensity noise suppression.

13.3.1 Temperature-Dependent Mean Intensity

The aim of the presented discussion is the description of a temperature-dependent central degree of second-order coherence $g^{(2)}_{\mathrm{out}}(0, T)$. As discussed in the previous section, quantum optical methods enable a description of $g^{(2)}_{\mathrm{out}}(0)$ in dependence of the mean emission intensity $\bar{I}$. It was shown experimentally, that the mean intensity of the emission has a distinct dependence on the temperature of the diode [231]. This behaviour of the emitted power can be modelled as a function of temperature by a Gaussian function

$$\bar{I}(T) = \bar{I}\, e^{-(T-T_0)^2/\sigma_T^2} + \delta I. \tag{13.18}$$

Approximating the experimental data using a least square fit, there is good agreement for peak mean intensity $\bar{I} = 1.51 \pm 0.13$, central temperature $T_0 = (197.1 \pm 0.9)$ K, temperature standard

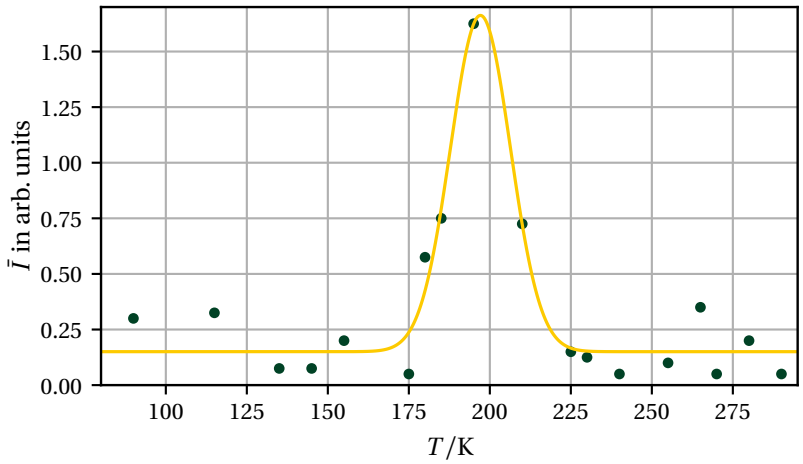

Figure 13.5: Mean intensity $\bar{I}$ versus temperature T. Both experimental data from [231] (green) and the Gaussian model (13.18) (yellow) attain a maximum in intensity at around 190 K.

deviation $\sigma_T = (13.1 \pm 1.0)$ K and offset intensity $\delta I = 0.15 \pm 0.03$ (see Fig. 13.5). The peak of emission intensity at T_0 implies an increase in diode efficiency at this temperature, which in turn signals the transition to a more saturated regime.

13.3.2 Temperature-Dependent Intensity Noise

The data gathered from fitting the experimental data for $\bar{I}(T)$ with (13.18) and describing $g^{(2)}_{\text{out}}(0)$ according to (13.17) for varying saturation parameters can be used to construct the temperature-dependent behaviour $g^{(2)}_{\text{out}}(0, T)$. As can be seen in Fig. 13.6, all saturation parameters combinations show a suppression of intensity fluctuations around 190 K. With parameters set to $I_s = 5$ and $\kappa = 0.35$, we reach good agreement with experimental data from [24]. The calculations using the fitted data from Fig. 13.5 do not reach a plateau of $g^{(2)}_{\text{out}}(0) = 2.0$ for high and low temperatures. This is due to the finite offset $\delta I = 0.15 \pm 0.03$ of $\bar{I}(T)$.

13.3.3 Discussion

Having developed a suitable description of the experimentally observed $g^{(2)}_{\text{out}}(0)$ reduction of hybrid light enables the search towards even more reduction of intensity fluctuations while keeping the broadband character of the emission. Adjusting the model parameters, for example to $I_s = 5$ and $\kappa = 0.55$, a suppression of intensity noise to about $g^{(2)}_{\text{out}}(0) = 1.09$ can be observed. This is very close to Poissonian statistics, however in this case not for a laser but for a broadband hybrid superluminescent diode. While the experimental and technological challenges in realizing such a light source, like finding an appropriate QD level system and sustaining the delicate balance of non-lasing, i.e. preventing a collapse of the spectral linewidth [255], are high, the potential in a broad area of applications is appealing.

In conclusion, we have demonstrated that a reduction of the central degree of second-order coherence $g^{(2)}_{\text{out}}(0)$ of the emitted light of a QDSLD towards the Poissonian limit can be described by statistics manipulation due to the joint mechanism of the nonlinear gain inside the medium and the QD emission properties. Using experimental findings for the temperature

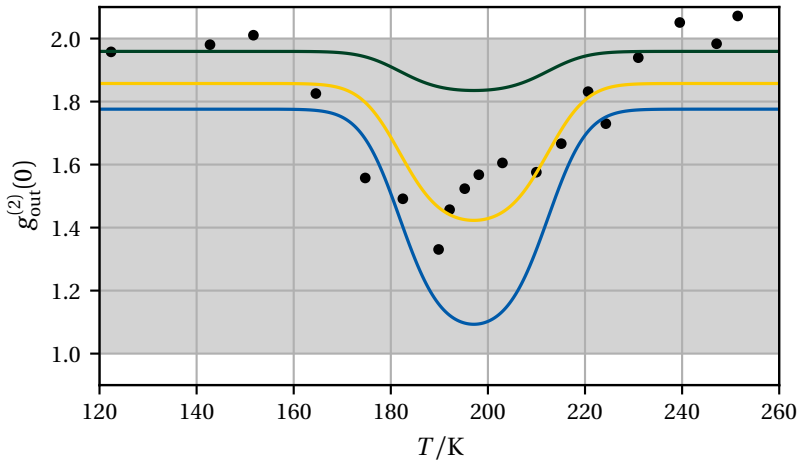

Figure 13.6: Central degree of second-order coherence $g^{(2)}_{\mathrm{out}}(0)$ versus temperature T. For $I_{\mathrm{s}} = 5$ and $\kappa = 0.35$ (yellow), there is good agreement between the presented model and experimental data from [24] (black). This agreement deteriorates for $\kappa = 0.1$ (green). Intensity noise suppression beyond $g^{(2)}_{\mathrm{out}}(0) = 1.33$ to approximately $g^{(2)}_{\mathrm{out}}(0) = 1.09$ can be observed for $\kappa = 0.55$ (blue).

dependence of the emitted power, observed intensity noise suppression at around 190 K can be reproduced. As the diode enters a more saturated regime for higher mean intensities, our investigation allows to interpret the effect as a temperature driven saturation effect inside the diode.

The characteristics of silent white light, high output power, a broadband optical spectrum, spatially directed emission and low intensity noise, are very favorable in a plethora of applications [256–259]. In particular, in the realm of artificial neural networks such light source could be used to realize efficient setups and improve the performance levels of implementations based on similar emitters [21, 260].

Part V

CONCLUSION

14

SUMMARY AND OUTLOOK

In this thesis, we study atomic and optical approaches to deep learning. These include using the movement of thermal atoms in optically shaped potential landscapes, harnessing the nonlinear dynamics in coherent matter waves and describing the temperature-dependent intensity noise suppression in quantum dot superluminescent diodes via statistics manipulation due to interaction with the pumped diode material. With these diverse approaches, we aim to contribute to the ongoing search for physical implementations of artificial neural networks with a special focus on optical and atomic systems.

THERMAL CLOUD NEURAL NETWORK

In a first approach to atomic implementations of artificial neural networks, we develop an algorithmic procedure using thermal atoms in optical dipole potentials to implement a neuron. Considering thermal atoms trapped in an external box potential, far-detuned lasers offer the possibility to shape the potential landscape in which the atoms move, constantly colliding with each other and eventually reaching thermal equilibrium. This dependence of the particle movement on the externally applied optical field enables the identification of inputs and output of a neuron.

Assuming a thermal cloud equilibrated in a box potential, a neuron input can be implemented using a single optical potential produced by a laser, yielding, for example, a Gaussian shaped intensity distribution according to (6.14). Depending on the sign of the input, a red-detuned laser for positive inputs or a blue-detuned laser for negative inputs has to be used. The absolute value of the input is implemented via the peak intensity of the laser. After equilibration of the thermal atoms in the new potential landscape, now dominantly determined by the input laser, the external box trap is turned off. The number of atoms still trapped by the input laser can be identified as output of the neuron. This algorithm yields a sigmoidal-like nonlinear activation function for the thermal cloud neuron, enabling training via steepest descent methods and the backpropagation algorithm.

As shown in Chapter 7, the thermal cloud neural network is able to solve the XOR problem after a reasonable training period, being able to identify every test sample correctly within set error margins. Additionally, the network can be used in the recognition of hand-written digits based on the MNIST dataset. For this real-world problem, requiring a larger network architecture, the thermal cloud neural network reaches an average error rate of 16.45% at the end of the training. This value exceeds the one using an identical network architecture with artificial neurons showing sigmoidal activation functions (average error rate 7.13%, see Section 3.3.2). This hints at the fact, that the activation function produced by the described implementation is not very well suited for such problems.

The thermal cloud neural network shows two mayor drawbacks. Firstly, it is not completely integrated. For communication between layers, output given by measurements of particle numbers has to be converted into inputs for the next layer, which are laser parameters. This requires to constantly interfere with the system, analyzing measurements and adjusting the laser setup. Secondly, the whole implementation is rather slow. As two equilibration processes are performed in the described algorithm, the duration of a neuron implementation for reasonable experimental parameters can be estimated as $t_{\mathrm{N}} \approx 688\,\mathrm{ms}$. Comparing this to calculation speeds of modern computers, this is far too slow for reasonable implementation. Therefore, while being a good platform to understand the intricacies of implementing an artificial neural network, the drawbacks of the thermal cloud neural network weight too heavy to pursue an implementation, unless improvements to the algorithm can be found.

FOUR-WAVE MIXING NEURAL NETWORK

In another approach to atomic implementations of artificial neural networks, we focus on the four-wave mixing process of coherent matter waves in Bose-Einstein condensates. Describing the condensate in mean-field theory using the Gross-Pitaevskii equation, we perform a detailed investigation of a superposition of four plane waves in a homogeneous condensate. By introducing appropriate coordinates and choosing a suitable dimensionless description, we show that the dynamics of the relative particle numbers and phases of the momentum components participating in the process are given by Josephson-like oscillations. Choosing certain parameters for two constants of motion, we are able to derive analytical expressions for those dynamics in terms of Jacobian elliptic functions. These results can be equally applied to two-dimensional four-wave mixing with four distinguished momentum states or one-dimensional four-wave mixing with two distinguished momentum states and two internal hyperfine states of the condensate atoms.

In addition to deriving analytical expressions of the four-wave mixing dynamics, we also investigate numerical solutions using simulations of the Gross-Pitaevskii equation on a discrete periodic grid. An instability in the simulation, arising due to numerical noise from Fast Fourier Transforms producing population outside of the four-wave mixing states, changes the frequency of the observed oscillations. Yet, their main characteristics persist. Hence, the analytical solution can be recovered using filter masks in momentum space.

With these results and the in-depth understanding of the four-wave mixing process in matter waves, we implement a complex-valued neuron, the four-wave mixing neuron. Identifying the complex amplitudes of three momentum components of the process as input and the remaining component as output, a nonlinear activation function can be established by setting a fixed duration τ_{F} of the neuron implementation. Due to the experimental accessibility of relative particle numbers and phases, we analyze the nonlinear response of the neuron to varying inputs in terms of those variables. This yields numerical results for the nonlinear activation function and its partial derivatives. Those can be used to train the neuron using steepest descent methods. To test the learning capabilities of the four-wave mixing neuron, we train it to solve the XOR problem. As the four-wave mixing neuron is a complex-valued neuron, a single realization is indeed able to solve this problem [222]. By developing an encoding scheme of the XOR problem to make good use of the value range of its nonlinear activation function, the four-wave mixing neuron is able to learn to solve the XOR problem, showing similar performance compared to conventional implementations and being able to identify every test sample correctly at the end of the training procedure.

In pursuit of setting up a four-wave mixing neural network, two critical problems have to be solved: multiple neurons have to be run in parallel without influencing each other and output information of one layer has to be converted to input for the next layer. While spatially separating the neurons provides a perfect parallelization opportunity, this also prevents an efficient transport of the output information without using a measurement and re-implementation.

A promising way to implement parallel neurons is to stack four-wave mixing configurations in momentum space. In this scenario, the neurons overlap perfectly and a communication of output to the next layer is feasible. However, as shown in Section 11.1.2, the dynamics of the stacked four-wave mixing configurations are not independent of each other. While the particle numbers in the individual configurations are conserved, coherent coupling between them leads to changes in the dynamics compared to the isolated case.

While this behaviour is disruptive in the realm of neural networks and prevents an efficient implementation, the mutual interactions are actually desirable in another approach to machine learning, namely reservoir computing. With the ability to stack multiple four-wave mixing configurations, one could realize a reservoir computing setup with precisely defined input and output channels and highly nonlinear dynamics. This kind of implementation seems to be well suited for coherent matter waves in Bose-Einstein condensates and offers a research direction worth pursuing.

Yet, we are able to find an alternative way to realize parallelizable four-wave mixing neurons and set up a neural network. Making use of the extended internal structure of the condensate atoms, four-wave mixing neurons in multiple internal components can be perfectly parallelized if all intra-component scattering lengths are equal and the inter-component scattering lengths vanish. In this case, two-dimensional four-wave mixing configurations in different components do not interact and show equal dynamics. With that, tree-like networks can be set up, ensuring that outputs of one layer can be used as input for the next layer by cyclically permuting the roles of the momentum components as input and output. We show that such a four-wave mixing neural network is able to learn to distinguish patterns in a logical 6-bit sequence via training using the backpropagation algorithm.

An advantage of the four-wave mixing neural network in comparison to the thermal atom implementation is a dramatic increase in speed. Using experimental parameters from [139], the duration of one run through a two-layer network can be estimated to be in the realm of $100\,\mu$s. Also, the complex-valued nature of the four-wave mixing neuron provides potential to extract more information from a single run through the network. Yet, there are parts of the proposed implementation demanding delicate attention. While the four-wave mixing neuron shows solid robustness against noise in the training process, the initialization sequence has to be performed very precisely to warrant good performance of the neuron. Additionally, searching for appropriate hyperfine levels or Feshbach resonances yielding the idealized scenario from Section 11.1.3 is of high importance. In the case that the four-wave mixing neural network is deemed reasonable for experimental implementation, an investigation of the power consumption would be of high interest.

SILENT WHITE LIGHT

Based on experimental measurements and with the potential of being applied as light sources in optical neural networks, we investigate the temperature-dependent intensity noise suppression of quantum dot superluminscent diodes. The first- and second-order temporal correlation properties of the emission of such diodes are modelled by a superposition of independent stochastic emitters. We show that such an approach is able to reconstruct the optical

power spectrum of the diode using stochastic simulations, enabling access to other quantities such as the intensity probability distribution and second-order temporal correlation function.

By considering the interaction between such emission and the pumped diode material, we show that photon statistics manipulation occurs in dependence of the system parameters. Using experimental data for the temperature dependence of the mean intensity emitted from the diode, we are able to reconstruct a temperature-dependent suppression of intensity noise in terms of the central degree of second-order coherence from thermal noise $g^{(2)}(0) = 2.0$ at room temperature to around $g^{(2)}(0) = 1.33$ at $T = 190\,\mathrm{K}$. Our model allows us to interpret the observations as a temperature induced saturation effect in the pumped diode material.

Being able to tune intensity noise can prove to be beneficial in various applications such as optical coherence tomography [256] and rotation sensing [257]. Also in the implementation in optical neural networks, light sources with spatially directed, high-powered, broadband emission with low intensity noise are desirable. Our model of intensity noise suppression might be used in determining optimal operation conditions of such diodes or be of usage in the design of new light sources.

Part VI

APPENDIX

A

TIME-DISCRETE APPROXIMATION OF STOCHASTIC DIFFERENTIAL EQUATIONS

A.1 PICARD-ITERATION OF STOCHASTIC DIFFERENTIAL EQUATIONS

The integral form of the stochastic differential equation [116]

$$\mathrm{d}x(t) = a(x(t))\mathrm{d}t + b(x(t))\mathrm{d}W(t) \tag{A.1}$$

is given by

$$x(t) = x_0 + \int_{t_0}^{t} \mathrm{d}t'\, a(x(t')) + \int_{t_0}^{t} \mathrm{d}W(t')\, b(x(t')). \tag{A.2}$$

Approximation schemes for this general solution can be developed using a Picard-type iteration

$$\begin{aligned} x^{[0]} &= x_0, \\ x^{[l+1]} &= x_0 + \int_{t_0}^{t} \mathrm{d}t'\, a\left(x^{[l]}(t')\right) + \int_{t_0}^{t} \mathrm{d}W(t')\, b\left(x^{[l]}(t')\right). \end{aligned} \tag{A.3}$$

This yields for the first iterations of (A.1)

$$x^{[0]} = x_0, \tag{A.4}$$

$$x^{[1]} = x_0 + a(x_0)\int_{t_0}^{t} \mathrm{d}t' + b(x_0)\int_{t_0}^{t} \mathrm{d}W(t'), \tag{A.5}$$

$$\begin{aligned} x^{[2]} = x_0 &+ \int_{t_0}^{t} \mathrm{d}t'\, a\left(x_0 + a(x_0)\int_{t_0}^{t'} \mathrm{d}t'' + b(x_0)\int_{t_0}^{t'} \mathrm{d}W(t'')\right) \\ &+ \int_{t_0}^{t} \mathrm{d}W(t')\, b\left(x_0 + a(x_0)\int_{t_0}^{t'} \mathrm{d}t'' + b(x_0)\int_{t_0}^{t'} \mathrm{d}W(t'')\right). \end{aligned} \tag{A.6}$$

From the second iteration onwards the terms inside the integrals have to be Taylor expanded to second-order to take all necessary integrals into account. The relevant terms from this expression yield, for example,

$$a\left(x_0 + a(x_0)\int_{t_0}^{t'} \mathrm{d}t'' + b(x_0)\int_{t_0}^{t'} \mathrm{d}W(t'')\right) = L^0 a(x_0)\int_{t_0}^{t'} \mathrm{d}t'' + L^1 a(x_0)\int_{t_0}^{t'} \mathrm{d}W(t''), \tag{A.7}$$

with $L^0 = a(x_0)\partial_x + \frac{1}{2}b^2(x_0)\partial_x^2$ and $L^1 = b(x_0)\partial_x$. Therefore, (A.6) reads

$$\begin{aligned} x^{[2]} = x_0 &+ b(x_0)\int_{t_0}^{t} \mathrm{d}W(t') + a(x_0)\int_{t_0}^{t} \mathrm{d}t' + L^1 b(x_0)\int_{t_0}^{t}\int_{t_0}^{t'} \mathrm{d}W(t')\mathrm{d}W(t'') \\ &+ L^1 a(x_0)\int_{t_0}^{t}\int_{t_0}^{t'} \mathrm{d}t'\mathrm{d}W(t'') + L^0 b(x_0)\int_{t_0}^{t}\int_{t_0}^{t'} \mathrm{d}W(t')\mathrm{d}t'' + L^0 a(x_0)\int_{t_0}^{t}\int_{t_0}^{t'} \mathrm{d}t'\mathrm{d}t''. \end{aligned} \tag{A.8}$$

For higher-order iterations, increasingly more encapsulated expressions have to be expanded.

A.2 TIME-DISCRETE APPROXIMATION

A time-discrete approximation scheme $x(t_i)$ is said to converge strongly with order $\gamma > 0$, if for the final time instant T and $N = T/\Delta$, there is a finite ϵ and $\Delta_0 > 0$ such that [127]

$$\langle |x(T) - x(t_N)| \rangle \leq \epsilon \Delta^\gamma \tag{A.9}$$

for any time-discretization $0 < \Delta < \Delta_0$. A so-called strong Taylor scheme can be constructed from the iteration above.

A criterium [127] for the terms of the iteration series required for the associated strong Taylor scheme to achieve a desired order of strong convergence γ can be given as follows. The scheme converges to order of an integer γ if it includes all combinations of integrals up to this order, with time differentials $\mathrm{d}t$ being of order 1 and Wiener noise increments $\mathrm{d}W(t)$ being of order 1/2. Simulation schemes of half-integer γ additionally require the inclusion of the pure time integral of order $\gamma + 1/2$. The well-known lowest-order simulation schemes, the Euler- and Milstein-schemes, can be constructed in this manner. The Euler scheme, being of order 1/2 of strong convergence, results by discretizing the time step in (A.5), yielding

$$x_{i+1} = x_i + a(x_i)\Delta t + b(x_i)\Delta W, \tag{A.10}$$

with $x_i = x(t_i)$, $\Delta t = t_{i+1} - t_i$ and ΔW a Gaussian random process with $\langle \Delta W \rangle = 0$ and $\langle \Delta W^2 \rangle = \Delta t$.

The order 1.0 Milstein scheme [261] can be constructed by considering all integrals from (A.6) up to the order of 1.0, yielding

$$x(t) = x_0 + a(x_0)\int_{t_0}^{t} \mathrm{d}t' + b(x_0)\int_{t_0}^{t} \mathrm{d}W(t') + L^1 b(x_0)\int_{t_0}^{t}\int_{t_0}^{t'} \mathrm{d}W(t')\mathrm{d}W(t''). \tag{A.11}$$

The double stochastic integral can be calculated as [116]

$$\int_{t_0}^{t}\int_{t_0}^{t'} \mathrm{d}W(t')\mathrm{d}W(t'') = \frac{1}{2}\left[\Delta W^2 - \Delta t\right], \tag{A.12}$$

yielding the iteration rule

$$x_{i+1} = x_i + a(x_i)\Delta t + b(x_i)\Delta W + \frac{1}{2}L^1 b(x_i)\left[\Delta W^2 - \Delta t\right]. \tag{A.13}$$

B

ATOMIC PARAMETERS OF RUBIDIUM-87

The atomic parameters used in stochastic simulations of ^{87}Rb gases are listed in Table B.1

Table B.1: Atomic parameters for stochastic simulations of ^{87}Rb atoms.

Quantity	Symbol	Value	Ref.
atomic mass	m	$1.443 \cdot 10^{-25}$ kg	[121]
Doppler temperature	T_D	$145.57\,\mu$K	[122]
particle diameter	a	600 pm	[125]
friction coefficient	β	$5.65 \cdot 10^{-17}$ kg/s	

C

TWO-DIMENSIONAL FOUR-WAVE MIXING STATE PREPARATION

The desired state after initialization for FWM is a superposition of plane waves with wave vectors $\boldsymbol{k}_1$, $\boldsymbol{k}_2$, $\boldsymbol{k}_3$ and $\boldsymbol{k}_4$, fulfilling the conditions (9.19). There, all possible combinations for relative populations n_1, n_2, n_3 and n_4 with $\sum_{j=1}^{4} n_j = 1$ should be realizable.

We suggest to use atomic beamsplitters based on Bragg diffraction to populate the momentum states. This method is based on the interaction between the BEC in its internal ground state and two counterpropagating laser beams. In this scenario, energy and momentum conservation have to hold [202],

$$\hbar\omega_1 + \frac{\hbar^2 k_\mathrm{i}^2}{2m} = \hbar\omega_2 + \frac{\hbar^2 k_\mathrm{f}^2}{2m}, \qquad \boldsymbol{k}_\mathrm{i} + \boldsymbol{k}_1 = \boldsymbol{k}_\mathrm{f} + \boldsymbol{k}_2, \tag{C.1}$$

with the initial $\boldsymbol{k}_\mathrm{i}$ and the final $\boldsymbol{k}_\mathrm{f}$ wave vector of the BEC and the frequencies ω_1 and ω_2 and wave vectors $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ of the laser beams, respectively. If the two laser beams are perfectly anti-collinear, the momentum transfer in the BEC can be characterized as

$$\boldsymbol{k}_\mathrm{f} - \boldsymbol{k}_\mathrm{i} = \boldsymbol{k}_1 - \boldsymbol{k}_2 = 2\boldsymbol{k}_\mathrm{L}, \tag{C.2}$$

where $\boldsymbol{k}_\mathrm{L} = (\boldsymbol{k}_1 + \boldsymbol{k}_2)/2$.

For a shallow lattice $U(\boldsymbol{r}) = 0$, the ground state energy $\hbar\omega_g$ of the BEC scales quadratically with the wave number according to (9.20). Hence, the laser frequencies have to be chosen

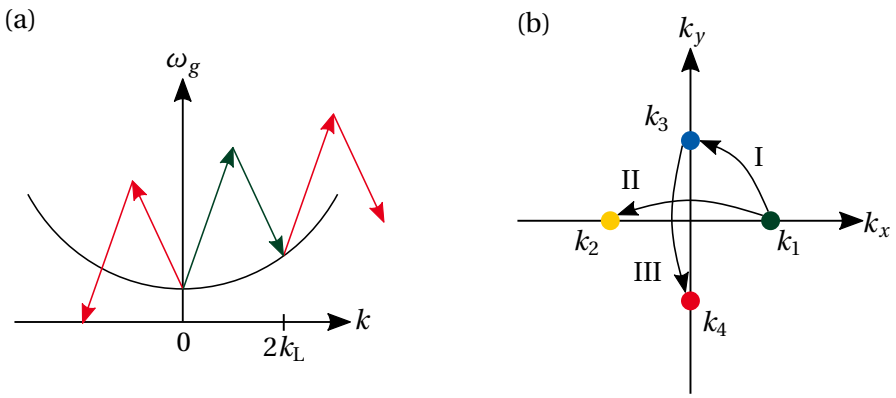

Figure C.1: (a) Energy diagram for Bragg-diffraction versus wave number k. A ground-state BEC initially at rest experiences population transfer to a state with $2\boldsymbol{k}_\mathrm{L}$ (green). Population transfer to other momentum states does not appear (red), as the conditions (C.1) are not fulfilled. (b) Proposed initialization sequence for FWM setup fulfilling (9.19). Three Bragg pulses are used to set up any combination of populations between the FWM states, while ensuring that no population is transferred outside of the FWM states.

carefully, such that population transfer between momentum states is energetically permitted (see Fig. C.1 (a)).

With that, a controlled initialization of momentum states can be performed, as initial states can be targeted individually and final states are given by the momentum and energy conditions (C.1). In Bragg diffraction, the portion of the population $0 \leq p_j \leq 1$ transferred between the momentum states can be controlled via the interaction duration between the BEC and the laser beams [202]. To avoid unwanted transitions outside of the FWM states, the preparation sequence shown in Fig. C.1 (b) and Table C.1 is developed. For visualization, we choose $\boldsymbol{k}_1 = \hat{\boldsymbol{k}}_x$, $\boldsymbol{k}_2 = -\hat{\boldsymbol{k}}_x$, $\boldsymbol{k}_3 = \hat{\boldsymbol{k}}_y$ and $\boldsymbol{k}_4 = -\hat{\boldsymbol{k}}_y$. However, all combinations fulfilling (9.19) can be prepared by the described procedure.

After the pulse sequence, the total particle number is transferred into the FWM states,

$$(1-p_3)(1-p_2)+p_3(1-p_2)+(1-p_4)p_2+p_4p_2=1, \tag{C.3}$$

and all desired combinations can be realized (see Fig. C.2).

Table C.1: Relative particle numbers in the FWM states after each Bragg pulse.

pulse	n_0	n_1	n_2	n_3	n_4
0	0	1	0	0	0
I	0	$(1-p_2)$	0	p_2	0
II	0	$(1-p_3)(1-p_2)$	$p_3(1-p_2)$	p_2	0
III	0	$(1-p_3)(1-p_2)$	$p_3(1-p_2)$	$(1-p_4)p_2$	p_4p_2

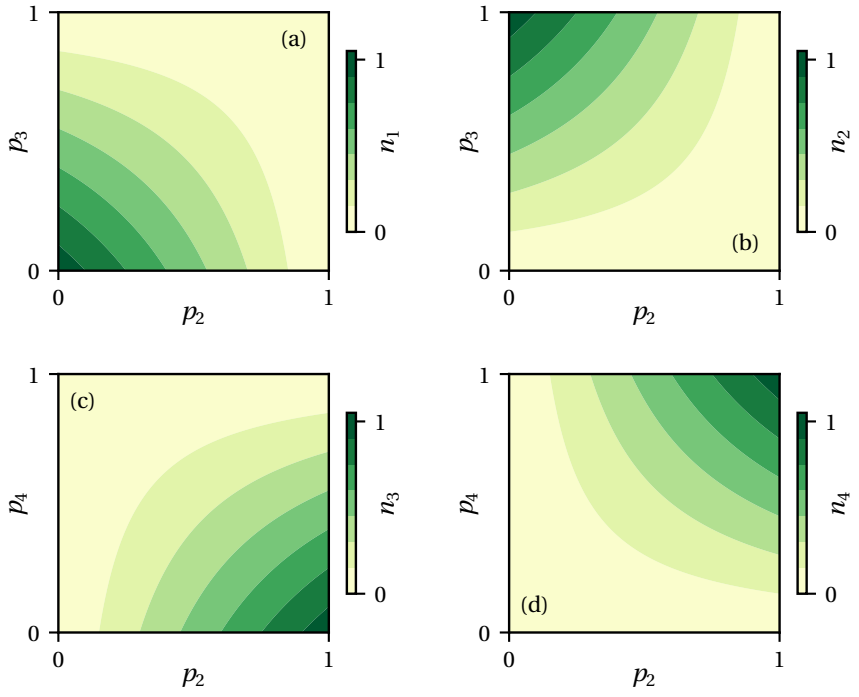

Figure C.2: Resulting relative particle numbers (a) n_1, (b) n_2, (c) n_3 and (d) n_4 for the initialization sequence described by Fig. C.1 and Table C.1. All probabilities range from 0 to 1, conserving the total probability (particle number) (C.3).

D

ENCODING SCHEME FOR FOUR-WAVE MIXING NEURAL NETWORK

The example presented in Section 11.2.3 to train and test a FWM neural network is given by a logical 6-bit sequence. To make use of the full nonlinear activation functions of the FWM neurons embedded in the network, an encoding scheme for the inputs and outputs has to be developed. While logical 0 is represented by a particle number input $n = 0.35$, a logical 1 is given by $n = 0.45$. The desired outputs used for training and testing are chosen to enable the network to distinguish different input patterns. The complete encoding scheme can be seen in Table D.1. As the input channels are chosen such that the output particle numbers of the FWM neurons in the hidden layer shown an exchange symmetry regarding those input particle numbers, the permutations of logical inputs of an individual FWM neuron are omitted.

Table D.1: Input and output encoding of a logical 6-bit sequence for training and testing on the FWM neural network described in Section 11.2.3.

Logical sequence	$n_{1,1}$	$n_{1,2}$	$n_{2,3}$	$n_{2,4}$	$n_{3,3}$	$n_{3,4}$	$\tilde{n}_{3,4}$
000000	0.35	0.35	0.35	0.35	0.35	0.35	0.106
100000	0.45	0.35	0.35	0.35	0.35	0.35	0.121
001000	0.35	0.35	0.45	0.35	0.35	0.35	0.131
000010	0.35	0.35	0.35	0.35	0.45	0.35	0.131
110000	0.45	0.45	0.35	0.35	0.35	0.35	0.147
101000	0.45	0.35	0.45	0.35	0.35	0.35	0.149
100010	0.45	0.35	0.35	0.35	0.45	0.35	0.149
001010	0.35	0.35	0.45	0.35	0.45	0.35	0.164
111000	0.45	0.45	0.45	0.35	0.35	0.35	0.174
110010	0.45	0.45	0.35	0.35	0.45	0.35	0.174
001100	0.35	0.35	0.45	0.45	0.35	0.35	0.182
000011	0.35	0.35	0.35	0.35	0.45	0.45	0.182
101010	0.45	0.35	0.45	0.35	0.45	0.35	0.184
101100	0.45	0.35	0.45	0.45	0.35	0.35	0.197
100011	0.45	0.35	0.35	0.35	0.45	0.45	0.197
111100	0.45	0.45	0.45	0.45	0.35	0.35	0.206
110011	0.45	0.45	0.35	0.35	0.45	0.45	0.206
111010	0.45	0.45	0.45	0.35	0.45	0.35	0.210
001110	0.35	0.35	0.45	0.45	0.45	0.35	0.232
001011	0.35	0.35	0.45	0.35	0.45	0.45	0.232
101110	0.45	0.35	0.45	0.45	0.45	0.35	0.250
101011	0.45	0.35	0.45	0.35	0.45	0.45	0.250
111110	0.45	0.45	0.45	0.45	0.45	0.35	0.261
111011	0.45	0.45	0.45	0.35	0.45	0.45	0.261
001111	0.35	0.35	0.45	0.45	0.45	0.45	0.348
101111	0.45	0.35	0.45	0.45	0.45	0.45	0.368
111111	0.45	0.45	0.45	0.45	0.45	0.45	0.376

BIBLIOGRAPHY

[1] Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, Nature **521**, 436–444 (2015) (cit. on p. 1).

[2] T. Davenport and R. Kalakota, *The potential for artificial intelligence in healthcare*, Future Healthc. J. **6**, 94–98 (2019) (cit. on p. 1).

[3] M. May, *Eight ways machine learning is assisting medicine*, Nat. Med. **27**, 2–3 (2021) (cit. on p. 1).

[4] H. A. Elmarakeby, J. Hwang, R. Arafeh, J. Crowdis, S. Gang, D. Liu, S. H. AlDubayan, K. Salari, S. Kregel, C. Richter, et al., *Biologically informed deep neural network for prostate cancer discovery*, Nature **598**, 348–352 (2021) (cit. on p. 1).

[5] M. F. Dixon, I. Halperin, and P. Bilokon, *Machine Learning in Finance* (Springer International Publishing, Cham, 2020) (cit. on p. 1).

[6] H. Ghoddusi, G. G. Creamer, and N. Rafizadeh, *Machine learning in energy economics and finance: A review*, Energy Econ. **81**, 709–727 (2019) (cit. on p. 1).

[7] R. Rai, M. K. Tiwari, D. Ivanov, and A. Dolgui, *Machine learning in manufacturing and industry 4.0 applications*, Int. J. Prod. Res. **59**, 4773–4778 (2021) (cit. on p. 1).

[8] L. Meng, B. McWilliams, W. Jarosinski, H.-Y. Park, Y.-G. Jung, J. Lee, and J. Zhang, *Machine Learning in Additive Manufacturing: A Review*, JOM **72**, 2363–2377 (2020) (cit. on p. 1).

[9] G. H. Gu, J. Noh, I. Kim, and Y. Jung, *Machine learning for renewable energy materials*, J. Mater. Chem. A **7**, 17096–17117 (2019) (cit. on p. 1).

[10] N. Shabbir, R. AhmadiAhangar, L. Kutt, M. N. Iqbal, and A. Rosin, *Forecasting Short Term Wind Energy Generation using Machine Learning*, in 2019 IEEE 60th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON) (2019), pp. 1–4 (cit. on p. 1).

[11] A. Sohani, H. Sayyaadi, C. Cornaro, M. H. Shahverdian, M. Pierro, D. Moser, N. Karimi, M. H. Doranehgard, and L. K. Li, *Using machine learning in photovoltaics to create smarter and cleaner energy generation systems: A comprehensive review*, J. Clean. Prod. **364**, 132701 (2022) (cit. on p. 1).

[12] M. R. Bachute and J. M. Subhedar, *Autonomous Driving Architectures: Insights of Machine Learning and Deep Learning Algorithms*, Mach. Learn. with Applications **6**, 100164 (2021) (cit. on p. 1).

[13] U. Dogan, J. Edelbrunner, and I. Iossifidis, *Autonomous driving: A comparison of machine learning techniques by means of the prediction of lane change behavior*, in 2011 IEEE International Conference on Robotics and Biomimetics (2011), pp. 1837–1843 (cit. on p. 1).

[14] H. Fujiyoshi, T. Hirakawa, and T. Yamashita, *Deep learning-based image recognition for autonomous driving*, IATSS Res. **43**, 244–252 (2019) (cit. on p. 1).

[15] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, *Machine learning and the physical sciences*, Rev. Mod. Phys. **91**, 045002 (2019) (cit. on p. 1).

[16] K. J. Wu, *Google's New AI Is a Master of Games, but How Does It Compare to the Human Mind?*, Accessed: Nov. 28 2023 (cit. on p. 1).

[17] A. Mehonic and A. J. Kenyon, *Brain-inspired computing needs a master plan*, Nature **604**, 255–260 (2022) (cit. on p. 1).

[18] X. Sui, Q. Wu, J. Liu, Q. Chen, and G. Gu, *A Review of Optical Neural Networks*, IEEE Access **8**, 70773–70783 (2020) (cit. on pp. 1, 21).

[19] D. Brunner, B. Penkovsky, B. A. Marquez, M. Jacquot, I. Fischer, and L. Larger, *Tutorial: Photonic neural networks in delay systems*, J. Appl. Phys. **124** (2018) (cit. on pp. 1, 103).

[20] M. Calzavara, Y. Kuriatnikov, A. Deutschmann-Olek, F. Motzoi, S. Erne, A. Kugi, T. Calarco, J. Schmiedmayer, and M. Prüfer, *Optimizing Optical Potentials With Physics-Inspired Learning Algorithms*, Phys. Rev. Appl. **19**, 044090 (2023) (cit. on pp. 1, 103).

[21] F. Hu, J. A. Holguin-Lerma, Y. Mao, P. Zou, C. Shen, T. Khee Ng, B. S. Ooi, and N. Chi, *Demonstration of a low-complexity memory-polynomial-aided neural network equalizer for CAP visible-light communication with superluminescent diode*, Opto-Electron. Adv. **3**, 200009 (2020) (cit. on pp. 1, 103, 120).

[22] P. Wei, C. Cheng, and T. Liu, *A Photonic Transducer-Based Optical Current Sensor Using Back-Propagation Neural Network*, IEEE Photonics Technology Letters **28**, 1513–1516 (2016) (cit. on pp. 1, 103).

[23] M. Hermans, P. Antonik, M. Haelterman, and S. Massar, *Embodiment of Learning in Electro-Optical Signal Processors*, Phys. Rev. Lett. **117**, 128301 (2016) (cit. on pp. 1, 103).

[24] M. Blazek and W. Elsäßer, *Coherent and thermal light: Tunable hybrid states with second-order coherence without first-order coherence*, Phys. Rev. A **84**, 063840 (2011) (cit. on pp. 1, 103–104, 113, 119–120).

[25] Y. Zuo, B. Li, Y. Zhao, Y. Jiang, Y.-C. Chen, P. Chen, G.-B. Jo, J. Liu, and S. Du, *All-optical neural network with nonlinear activation functions*, Optica **6**, 1132 (2019) (cit. on pp. 2, 22).

[26] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, *Observation of Bose-Einstein Condensation in a Dilute Atomic Vapor*, Science **269**, 198–201 (1995) (cit. on p. 2).

[27] M. R. Andrews, C. G. Townsend, H.-J. Miesner, D. S. Durfee, D. M. Kurn, and W. Ketterle, *Observation of Interference Between Two Bose Condensates*, Science **275**, 637–641 (1997) (cit. on pp. 2, 38).

[28] R. W. Boyd, *Nonlinear Optics*, 3rd ed. (Academic Press, Burlington, 2008) (cit. on pp. 2, 57–58).

[29] M. Trippenbach, Y. Band, and P. Julienne, *Four wave mixing in the scattering of Bose-Einstein condensates*, Opt. Express **3**, 530 (1998) (cit. on pp. 2, 59).

[30] M. Trippenbach, Y. B. Band, and P. S. Julienne, *Theory of four-wave mixing of matter waves from a Bose-Einstein condensate*, Phys. Rev. A **62**, 023608 (2000) (cit. on pp. 2, 57, 59, 62).

[31] L. Deng, E. W. Hagley, J. Wen, M. Trippenbach, Y. Band, P. S. Julienne, J. E. Simsarian, K. Helmerson, S. L. Rolston, and W. D. Phillips, *Four-wave mixing with matter waves*, Nature **398**, 218–220 (1999) (cit. on pp. 2, 59, 61).

[32] J. M. Vogels, K. Xu, and W. Ketterle, *Generation of Macroscopic Pair-Correlated Atomic Beams by Four-Wave Mixing in Bose-Einstein Condensates*, Phys. Rev. Lett. **89**, 020401 (2002) (cit. on pp. 2, 59).

[33] A. J. Leggett, *Bose-Einstein condensation in the alkali gases: Some fundamental concepts*, Rev. Mod. Phys. **73**, 307–356 (2001) (cit. on pp. 2, 65).

[34] A. Smerzi, S. Fantoni, S. Giovanazzi, and S. R. Shenoy, *Quantum Coherent Atomic Tunneling between Two Trapped Bose-Einstein Condensates*, Phys. Rev. Lett. **79**, 4950–4953 (1997) (cit. on pp. 2, 65).

[35] J. Williams, R. Walser, J. Cooper, E. Cornell, and M. Holland, *Nonlinear Josephson-type oscillations of a driven, two-component Bose-Einstein condensate*, Phys. Rev. A **59**, R31–R34 (1999) (cit. on pp. 2, 65).

[36] K. N. Hansmann and R. Walser, *Forming complex neurons by four-wave mixing in a Bose-Einstein condensate*, Phys. Rev. A **109**, 013302 (2024) (cit. on pp. 4, 57, 77).

[37] K. N. Hansmann and R. Walser, *Stochastic Simulation of Emission Spectra and Classical Photon Statistics of Quantum Dot Superluminescent Diodes*, J. Mod. Phys. **12**, 22–34 (2021) (cit. on pp. 4, 104).

[38] K. N. Hansmann, F. Dommermuth, W. Elsäßer, and R. Walser, *Silent White Light*, (2023) (cit. on pp. 4, 113, 118).

[39] A. M. Turing, *On Computable Numbers, with an Application to the Entscheidungsproblem*, Proc. London Math. Soc. **s2-42**, 230–265 (1936) (cit. on p. 7).

[40] W. S. McCulloch and W. Pitts, *A Logical Calculus of the Ideas Immanent in Nervous Activity*, Bull. Math. Biol. **5**, 115–133 (1943) (cit. on pp. 7, 9, 11).

[41] F. Rosenblatt, *The perceptron: A probabilistic model for information storage and organization in the brain*, Psycol. Rev. **65**, 386–408 (1958) (cit. on pp. 7, 9, 11).

[42] M. Minsky and S. A. Papert, *Perceptrons* (MIT Press, Cambridge, 1969) (cit. on pp. 7, 16, 81).

[43] P. Werbos, *Beyond regression: new tools for prediction and analysis in the behavioral sciences*, Ph.D. thesis (Harvard University, 1974) (cit. on pp. 7, 15).

[44] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, 2016) (cit. on pp. 7–8, 11).

[45] S. Haykin, *Neural Networks and Learning Machines*, 3rd (Pearson Prentice Hall, New York, 2009) (cit. on pp. 7, 9, 12–15, 17).

[46] L. A. Urry, M. L. Cain, S. A. Wasserman, P. V. Minorsky, and J. B. Reece, *Campbell Biology*, 11th ed. (Pearson, Hallbergmoos, 2019) (cit. on pp. 8–9).

[47] M. Trepel, *Neuroanatomie*, 5th ed. (Urban & Fischer in Elsevier, München, 2011) (cit. on p. 8).

[48] A. Graves, *Supervised Sequence Labelling with Recurrent Neural Networks*, Vol. 385, Studies in Computational Intelligence (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012) (cit. on p. 11).

[49] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning representations by back-propagating errors*, Nature **323**, 533–536 (1986) (cit. on p. 11).

[50] K. J. Lang, A. H. Waibel, and G. E. Hinton, *A time-delay neural network architecture for isolated word recognition*, Neural Netw. **3**, 23–43 (1990) (cit. on p. 11).

[51] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, *Backpropagation Applied to Handwritten Zip Code Recognition*, Neural Comput. **1**, 541–551 (1989) (cit. on p. 11).

[52] D. O. Hebb, *The Organization of Behavior* (Wiley, New York, 1949) (cit. on p. 11).

[53] B. Widrow and M. E. Hoff, *Adaptive switching circuits*, IRE WESCON Convention Record **4**, 96–104 (1960) (cit. on p. 11).

[54] D. E. Rumelhart and J. L. McClelland, *Parallel Distributed Processing* (The MIT Press, Cambridge, 1986) (cit. on p. 11).

[55] J. McClelland, D. Rumelhart, and G. Hinton, *The appeal of parallel distributed processing*, in *Computing and intelligence* (American Association for Artificial Intelligence, 1995), pp. 305–341 (cit. on p. 11).

[56] G. E. Hinton, S. Osindero, and Y.-W. Teh, *A Fast Learning Algorithm for Deep Belief Nets*, Neural Comput. **18**, 1527–1554 (2006) (cit. on p. 11).

[57] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, *Greedy layer-wise training of deep networks*, NIPS'2006 (2007) (cit. on p. 11).

[58] Y. Bengio and Y. LeCun, *Scaling learning algorithms towards AI*, Large Scale Kernel Machines (2007) (cit. on p. 11).

[59] M. Ranzato, C. Poultney, S. Chopra, and Y. LeCun, *Efficient learning of sparse representations with an energy-based model*, NIPS'2006 (2007) (cit. on p. 11).

[60] Y. LeCun, C. Cortes, and C. J. C. Burges, *The MNIST database of handwritten digits*, Accessed: Nov. 28 2023 (cit. on pp. 11, 16–18).

[61] A. Krizhevsky and G. Hinton, *Learning multiple layers of features from tiny images*, tech. rep. (University of Toronto, 2009) (cit. on p. 11).

[62] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, *ImageNet: A large-scale hierarchical image database*, in 2009 IEEE Conference on Computer Vision and Pattern Recognition (2009), pp. 248–255 (cit. on p. 11).

[63] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., *ImageNet Large Scale Visual Recognition Challenge*, International Journal of Computer Vision **115**, 211–252 (2015) (cit. on p. 11).

[64] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, and L. Fei-Fei, *Large-Scale Video Classification with Convolutional Neural Networks*, in 2014 IEEE conference on computer vision and pattern recognition (2014), pp. 1725–1732 (cit. on p. 11).

[65] P. F. Brown, J. Cocke, S. A. D. Pietra, V. J. D. Pietra, F. Jelinek, J. D. Lafferty, R. L. Mercer, and P. S. Roossin, *A statistical approach to machine translation*, Computational Linguistics **16**, 79–85 (1990) (cit. on p. 11).

[66] A. Krizhevsky, I. Sutsekver, and G. Hinton, *ImageNet classification with deep convolutional neural networks*, NIPS'2012 (2012) (cit. on pp. 11, 21).

[67] L. Deng, M. L. Seltzer, D. Yu, A. Acero, A.-R. Mohamed, and G. Hinton, *Binary coding of speech spectrograms using a deep auto-encoder*, in Interspeech 2010 (2010), pp. 1692–1695 (cit. on p. 11).

[68] G. Hinton, L. Deng, D. Yu, G. Dahl, A.-R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. Sainath, et al., *Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups*, IEEE Signal Processing Magazine **29**, 82–97 (2012) (cit. on p. 11).

[69] P. Sermanet, K. Kavukcuoglu, S. Chintala, and Y. LeCun, *Pedestrian Detection with Unsupervised Multi-Stage Feature Learning*, (2012) (cit. on p. 11).

[70] D. Cireşan, U. Meier, J. Masci, and J. Schmidhuber, *Multi-column deep neural network for traffic sign classification*, Neural Netw. **32**, 333–338 (2012) (cit. on p. 11).

[71] R. Collobert, K. Kavukcuglo, and C. Farabet, *Torch7: A Matlab-like environment for machine learning*, BigLearn, NIPS Workshop (2011) (cit. on p. 11).

[72] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems*, (2016) (cit. on p. 11).

[73] J. Mendel and R. McLaren, *Reinforcement-Learning Control and Pattern Recognition Systems*, in *Mathematics in science and engineering*, Vol. 66, C (Academic Press, Inc., 1970), pp. 287–318 (cit. on p. 12).

[74] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*, Vol. 12, Texts in Applied Mathematics (Springer New York, New York, NY, 2002) (cit. on p. 13).

[75] R. Germundsson and E. W. Weisstein, *XOR*, Accessed: Nov. 28 2023 (cit. on p. 16).

[76] D. S. Touretzky and D. A. Pomerleau, *What is hidden in hidden layers*, Byte **14**, 227–233 (1989) (cit. on p. 16).

[77] C. C. Aggarwal, *Neural Networks and Deep Learning* (Springer International Publishing, Cham, 2018) (cit. on p. 18).

[78] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, *Gradient-based learning applied to document recognition*, Proceedings of the IEEE **86**, 2278–2324 (1998) (cit. on pp. 18–19).

[79] S. Mittal and S. Vaishay, *A survey of techniques for optimizing deep learning on GPUs*, J. Syst. Archit. **99**, 101635 (2019) (cit. on p. 21).

[80] S. Mittal, *A Survey on optimized implementation of deep learning models on the NVIDIA Jetson platform*, J. Syst. Archit. **97**, 428–442 (2019) (cit. on p. 21).

[81] D. Patterson, J. Gonzalez, Q. Le, C. Liang, L.-M. Munguia, D. Rothchild, D. So, M. Texier, and J. Dean, *Carbon Emissions and Large Neural Network Training*, (2021) (cit. on p. 21).

[82] G. Wetzstein, A. Ozcan, S. Gigan, S. Fan, D. Englund, M. Soljačić, C. Denz, D. A. B. Miller, and D. Psaltis, *Inference in artificial intelligence with deep optics and photonics*, Nature **588**, 39–47 (2020) (cit. on pp. 21–22).

[83] J. W. Goodman, A. R. Dias, and L. M. Woody, *Fully parallel, high-speed incoherent optical method for performing discrete Fourier transforms*, Opt. Lett. **2**, 1 (1978) (cit. on p. 21).

[84] H. J. Caulfield and S. Dolev, *Why future supercomputing requires optics*, Nat. Photon. **4**, 261–263 (2010) (cit. on p. 21).

[85] D. A. B. Miller, *Attojoule Optoelectronics for Low-Energy Information Processing and Communications*, J. Lightwave Technol. **35**, 346–396 (2017) (cit. on p. 21).

[86] N. H. Farhat, D. Psaltis, A. Prata, and E. Paek, *Optical implementation of the Hopfield model*, Appl. Opt. **24**, 1469 (1985) (cit. on p. 21).

[87] Y. Kuratomi, A. Takimoto, K. Akiyama, and H. Ogawa, *Optical neural network using vector-feature extraction*, Appl. Opt. **32**, 5750 (1993) (cit. on p. 21).

[88] S. Gao, J. Yang, Z. Feng, and Y. Zhang, *Implementation of a large-scale optical neural network by use of a coaxial lenslet array for interconnection*, Appl. Opt. **36**, 4779 (1997) (cit. on p. 21).

[89] D. Psaltis, D. Brady, X.-G. Gu, and S. Lin, *Holography in artificial neural networks*, Nature **343**, 325–330 (1990) (cit. on p. 21).

[90] T.-Y. Cheng, D.-Y. Chou, C.-C. Liu, Y.-J. Chang, and C.-C. Chen, *Optical neural networks based on optical fiber-communication system*, Neurocomputing **364**, 239–244 (2019) (cit. on p. 21).

[91] X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, *All-optical machine learning using diffractive deep neural networks*, Science **361**, 1004–1008 (2018) (cit. on p. 21).

[92] A. Ryou, J. Whitehead, M. Zhelyeznyakov, P. Anderson, C. Keskin, M. Bajcsy, and A. Majumdar, *Free-space optical neural network based on thermal atomic nonlinearity*, Photonics Res. **9**, B128 (2021) (cit. on p. 22).

[93] S. Skinner, J. Steck, and E. Behrman, *Optical neural network using Kerr-type nonlinear materials*, in Proceedings of the Fourth International Conference on Microelectronics for Neural Networks and Fuzzy Systems (1994), pp. 12–15 (cit. on p. 22).

[94] A. Dejonckheere, F. Duport, A. Smerieri, L. Fang, J.-L. Oudar, M. Haelterman, and S. Massar, *All-optical reservoir computer based on saturation of absorption*, Opt. Express **22**, 10868 (2014) (cit. on p. 22).

[95] M. A. Nahmias, T. F. de Lima, A. N. Tait, H.-T. Peng, B. J. Shastri, and P. R. Prucnal, *Photonic Multiply-Accumulate Operations for Neural Networks*, IEEE J. Sel. Top. Quantum Electron. **26**, 1–18 (2020) (cit. on p. 22).

[96] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, et al., *Deep learning with coherent nanophotonic circuits*, Nat. Photon. **11**, 441–446 (2017) (cit. on p. 22).

[97] A. N. Tait, T. F. de Lima, E. Zhou, A. X. Wu, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, *Neuromorphic photonic networks using silicon photonic weight banks*, Sci. Rep. **7**, 7430 (2017) (cit. on p. 22).

[98] J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, *All-optical spiking neurosynaptic networks with self-learning capabilities*, Nature **569**, 208–214 (2019) (cit. on p. 22).

[99] C. Kaspar, B. J. Ravoo, W. G. van der Wiel, S. V. Wegner, and W. H. P. Pernice, *The rise of intelligent matter*, Nature **594**, 345–355 (2021) (cit. on p. 22).

[100] H. P. Nautrup, T. Metger, R. Iten, S. Jerbi, L. M. Trenkwalder, H. Wilming, H. J. Briegel, and R. Renner, *Operationally meaningful representations of physical systems in neural networks*, Mach. Learn.: Sci. Technol. **3**, 045025 (2022) (cit. on p. 22).

[101] L. G. Wright, T. Onodera, M. M. Stein, T. Wang, D. T. Schachter, Z. Hu, and P. L. McMahon, *Deep physical neural networks trained with backpropagation*, Nature **601**, 549–555 (2022) (cit. on p. 22).

[102] T. W. Hughes, I. A. D. Williamson, M. Minkov, and S. Fan, *Wave physics as an analog recurrent neural network*, Sci. Adv. **5** (2019) (cit. on p. 22).

[103] M. Romera, P. Talatchian, S. Tsunegi, F. Abreu Araujo, V. Cros, P. Bortolotti, J. Trastoy, K. Yakushiji, A. Fukushima, H. Kubota, et al., *Vowel recognition with four coupled spin-torque nano-oscillators*, Nature **563**, 230–234 (2018) (cit. on p. 22).

[104] L. Mennel, J. Symonowicz, S. Wachter, D. K. Polyushkin, A. J. Molina-Mendoza, and T. Mueller, *Ultrafast machine vision with 2D material neural network image sensors*, Nature **579**, 62–66 (2020) (cit. on p. 22).

[105] J. T. M. Walraven, *Quantum Gases*, tech. rep. (University of Amsterdam, Amsterdam, 2019) (cit. on pp. 25, 54).

[106] C. Cohen-Tannoudji, *Atomic motion in laser light*, 1992 (cit. on pp. 25, 35–36).

[107] J. Dalibard and C. Cohen-Tannoudji, *Atomic motion in laser light: connection between semiclassical and quantum descriptions*, J. Phys. B **18**, 1661–1683 (1985) (cit. on pp. 25, 36).

[108] R. Brown, *A brief account of microscopical observations made in the months of June, July and August 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies*, The Philosophical Magazine **4**, 161–173 (1828) (cit. on p. 25).

[109] A. Einstein, *Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen*, Ann. Phys. **322**, 549–560 (1905) (cit. on pp. 25, 29).

[110] M. von Smoluchowski, *Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen*, Ann. Phys. **326**, 756–780 (1906) (cit. on p. 25).

[111] P. Langevin, *Sur la théorie du mouvement brownien*, C. R. Acad. Sci. **146**, 530–533 (1908) (cit. on pp. 26, 31).

[112] D. S. Lemons and A. Gythiel, *Paul Langevin's 1908 paper "On the Theory of Brownian Motion" ["Sur la théorie du mouvement brownien," C. R. Acad. Sci. (Paris) 146 , 530–533 (1908)]*, Am. J. Phys. **65**, 1079–1081 (1997) (cit. on pp. 26, 31).

[113] H. Kramers, *Brownian motion in a field of force and the diffusion model of chemical reactions*, Physica **7**, 284–304 (1940) (cit. on p. 26).

[114] A. W. Knapp, *Connection between Brownian motion and potential theory*, J. Math. Anal. Appl. **12**, 328–349 (1965) (cit. on p. 26).

[115] W. Nolting, *Grundkurs Theoretische Physik 2*, 9th ed., Springer-Lehrbuch (Springer Berlin Heidelberg, Berlin, Heidelberg, 2014) (cit. on pp. 26–27).

[116] C. Gardiner, *Stochastic Methods*, 4th ed. (Springer Berlin Heidelberg, Berlin, Heidelberg, 2009) (cit. on pp. 27–28, 33, 104–105, 109, 129–130).

[117] W. Nolting, *Grundkurs Theoretische Physik 6*, Springer-Lehrbuch (Springer Berlin Heidelberg, Berlin, Heidelberg, 2014) (cit. on p. 28).

[118] A. Deriglazov, *Classical Mechanics*, 2nd ed. (Springer Cham, Cham, 2017) (cit. on p. 29).

[119] H. Yoshida, *Construction of higher order symplectic integrators*, Phys. Lett. A **150**, 262–268 (1990) (cit. on p. 30).

[120] E. Hairer, G. Wanner, and C. Lubich, *Geometric Numerical Integration*, Vol. 31, Springer Series in Computational Mathematics (Springer-Verlag, Berlin/Heidelberg, 2006), p. 644 (cit. on p. 30).

[121] M. P. Bradley, J. V. Porto, S. Rainville, J. K. Thompson, and D. E. Pritchard, *Penning Trap Measurements of the Masses of ^{133}Cs, $^{87,85}Rb$ and ^{23}Na with Uncertainties ≤ 0.2 ppb*, Phys. Rev. Lett. **83**, 4510–4513 (1999) (cit. on pp. 31, 131).

[122] D. A. Steck, *Rubidium 87 D Line Data, Version 2.2.2, last revised 9 July 2021*, Accessed: Nov. 28 2023 (cit. on pp. 31, 131).

[123] N. Navon, R. P. Smith, and Z. Hadzibabic, *Quantum gases in optical boxes*, Nat. Phys. **17**, 1334–1341 (2021) (cit. on pp. 31, 38).

[124] F. W. Sears and L. Salinger, Gerhard, *Thermodynamics, kinetic theory, and statistical thermodynamics*, 3rd ed. (Addison-Wesley Publishing Company, Reading, Massachusetts, 1978) (cit. on p. 31).

[125] M. Mantina, A. C. Chamberlin, R. Valero, C. J. Cramer, and D. G. Truhlar, *Consistent van der Waals Radii for the Whole Main Group*, J. Phys. Chem. A **113**, 5806–5812 (2009) (cit. on pp. 32, 131).

[126] C. J. Pethick and H. Smith, *Bose–Einstein Condensation in Dilute Gases* (Cambridge University Press, 2008) (cit. on pp. 32, 51–55).

[127] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1992) (cit. on pp. 33, 105, 130).

[128] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Atom—Photon Interactions* (Wiley, 1998) (cit. on pp. 36–37, 95, 113, 115).

[129] J. P. Gordon and A Ashkin, *Motion of atoms in a radiation trap*, Phys. Rev. A **21**, 1606–1617 (1980) (cit. on p. 36).

[130] D. Walls and G. J. Milburn, *Quantum Optics* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008) (cit. on p. 37).

[131] M. Orszag, *Quantum Optics* (Springer International Publishing, Cham, 2016) (cit. on p. 37).

[132] H. J. Metcalf and P. van der Straten, *Laser Cooling and Trapping*, Graduate Texts in Contemporary Physics (Springer New York, New York, NY, 1999) (cit. on pp. 37, 52).

[133] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari, *Theory of Bose-Einstein condensation in trapped gases*, Rev. Mod. Phys. **71**, 463–512 (1999) (cit. on p. 38).

[134] I. Bloch, J. Dalibard, and W. Zwerger, *Many-body physics with ultracold gases*, Rev. Mod. Phys. **80**, 885–964 (2008) (cit. on pp. 38, 53).

[135] A. Görlitz, J. M. Vogels, A. E. Leanhardt, C. Raman, T. L. Gustavson, J. R. Abo-Shaeer, A. P. Chikkatur, S. Gupta, S. Inouye, T. Rosenband, et al., *Realization of Bose-Einstein Condensates in Lower Dimensions*, Phys. Rev. Lett. **87**, 130402 (2001) (cit. on pp. 38, 53).

[136] M. Greiner, I. Bloch, O. Mandel, T. W. Hänsch, and T. Esslinger, *Exploring Phase Coherence in a 2D Lattice of Bose-Einstein Condensates*, Phys. Rev. Lett. **87**, 160405 (2001) (cit. on p. 38).

[137] I. Bloch, *Ultracold quantum gases in optical lattices*, Nat. Phys. **1**, 23–30 (2005) (cit. on p. 38).

[138] A. L. Gaunt, T. F. Schmidutz, I. Gotlibovych, R. P. Smith, and Z. Hadzibabic, *Bose-Einstein Condensation of Atoms in a Uniform Potential*, Phys. Rev. Lett. **110**, 200406 (2013) (cit. on pp. 38, 53, 57, 60).

[139] L. Chomaz, L. Corman, T. Bienaimé, R. Desbuquois, C. Weitenberg, S. Nascimbène, J. Beugnon, and J. Dalibard, *Emergence of coherence via transverse condensation in a uniform quasi-two-dimensional Bose gas*, Nat. Commun. **6**, 6162 (2015) (cit. on pp. 38, 57, 60, 78, 125).

[140] M. Tajik, B. Rauer, T. Schweigler, F. Cataldini, J. Sabino, F. S. Møller, S.-C. Ji, I. E. Mazets, and J. Schmiedmayer, *Designing arbitrary one-dimensional potentials on an atom chip*, Opt. Express **27**, 33474 (2019) (cit. on p. 38).

[141] N. Navon, A. L. Gaunt, R. P. Smith, and Z. Hadzibabic, *Emergence of a turbulent cascade in a quantum gas*, Nature **539**, 72–75 (2016) (cit. on p. 38).

[142] J. L. Ville, R. Saint-Jalm, É. Le Cerf, M. Aidelsburger, S. Nascimbène, J. Dalibard, and J. Beugnon, *Sound Propagation in a Uniform Superfluid Two-Dimensional Bose Gas*, Phys. Rev. Lett. **121**, 145301 (2018) (cit. on p. 38).

[143] J. Zhang, C. Eigen, W. Zheng, J. A. P. Glidden, T. A. Hilker, S. J. Garratt, R. Lopes, N. R. Cooper, Z. Hadzibabic, and N. Navon, *Many-Body Decay of the Gapped Lowest Excitation of a Bose-Einstein Condensate*, Phys. Rev. Lett. **126**, 060402 (2021) (cit. on p. 38).

[144] R. Lopes, C. Eigen, N. Navon, D. Clément, R. P. Smith, and Z. Hadzibabic, *Quantum Depletion of a Homogeneous Bose-Einstein Condensate*, Phys. Rev. Lett. **119**, 190404 (2017) (cit. on p. 38).

[145] B. Rauer, S. Erne, T. Schweigler, F. Cataldini, M. Tajik, and J. Schmiedmayer, *Recurrences in an isolated quantum many-body system*, Science **360**, 307–310 (2018) (cit. on p. 38).

[146] N. Davidson, H. Jin Lee, C. S. Adams, M. Kasevich, and S. Chu, *Long Atomic Coherence Times in an Optical Dipole Trap*, Phys. Rev. Lett. **74**, 1311–1314 (1995) (cit. on p. 38).

[147] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, Wiley Series in Pure and Applied Optics (John Wiley & Sons, Inc., New York, USA, 1991) (cit. on pp. 38, 103).

[148] E. Hecht, *Optik* (De Gruyter, 2018) (cit. on p. 38).

[149] G. Gauthier, I. Lenton, N. McKay Parry, M. Baker, M. J. Davis, H. Rubinsztein-Dunlop, and T. W. Neely, *Direct imaging of a digital-micromirror device for configurable microscopic optical potentials*, Optica **3**, 1136 (2016) (cit. on p. 38).

[150] G. Gauthier, T. A. Bell, A. B. Stilgoe, M. Baker, H. Rubinsztein-Dunlop, and T. W. Neely, *Dynamic high-resolution optical trapping of ultracold atoms*, in *Advances in atomic, molecoular, and optcial physics 70* (2021), pp. 1–101 (cit. on p. 38).

[151] K. B. Davis, M.-O. Mewes, M. A. Joffe, M. R. Andrews, and W. Ketterle, *Evaporative Cooling of Sodium Atoms*, Phys. Rev. Lett. **74**, 5202–5205 (1995) (cit. on p. 41).

[152] L. Vestergaard Hau, B. D. Busch, C. Liu, Z. Dutton, M. M. Burns, and J. A. Golovchenko, *Near-resonant spatial images of confined Bose-Einstein condensates in a 4-Dee magnetic bottle*, Phys. Rev. A **58**, R54–R57 (1998) (cit. on p. 41).

[153] W. Ketterle, D. S. Durfee, and D. M. Stamper-Kurn, *Making, probing and understanding Bose-Einstein condensates*, (1999) (cit. on p. 41).

[154] S. Bose, *Plancks Gesetz und Lichtquantenhypothese*, Z. Phys. **26**, 178–181 (1924) (cit. on p. 51).

[155] A. Einstein, *Quantentheorie des einatomigen idealen Gases*, Sitzungsberichte der Preussischen Akademie der Wissenschaften, 261 (1924) (cit. on p. 51).

[156] F. London, *The λ-Phenomenon of Liquid Helium and the Bose-Einstein Degeneracy*, Nature **141**, 643–644 (1938) (cit. on p. 51).

[157] D. G. Fried, T. C. Killian, L. Willmann, D. Landhuis, S. C. Moss, D. Kleppner, and T. J. Greytak, *Bose-Einstein Condensation of Atomic Hydrogen*, Phys. Rev. Lett. **81**, 3811–3814 (1998) (cit. on p. 51).

[158] S. Chu, C. Cohen-Tannoudji, and W. D. Phillips, *The Nobel Prize in Physics 1997*, Accessed: Nov. 28 2023 (cit. on p. 52).

[159] E. Cornell, W. Ketterle, and C. E. Wieman, *The Nobel Prize in Physics 2001*, Accessed: Nov. 28 2023 (cit. on p. 52).

[160] O. Penrose and L. Onsager, *Bose-Einstein Condensation and Liquid Helium*, Phys. Rev. **104**, 576–584 (1956) (cit. on p. 52).

[161] C. N. Yang, *Concept of Off-Diagonal Long-Range Order and the Quantum Phases of Liquid He and of Superconductors*, Rev. Mod. Phys. **34**, 694–704 (1962) (cit. on p. 52).

[162] N. P. Proukakis, D. W. Snoke, and P. B. Littlewood, eds., *Universal Themes of Bose-Einstein Condensation* (Cambridge University Press, 2017) (cit. on p. 53).

[163] S. Inouye, M. R. Andrews, J. Stenger, H.-J. Miesner, D. M. Stamper-Kurn, and W. Ketterle, *Observation of Feshbach resonances in a Bose–Einstein condensate*, Nature **392**, 151–154 (1998) (cit. on p. 53).

[164] J. Reichel and V. Vuletic, *Atom Chips* (Wiley-VCH, Berlin, 2011) (cit. on p. 53).

[165] C. Ryu, M. Andersen, P. Cladé, V. Natarajan, K. Helmerson, and W. Phillips, *Observation of Persistent Flow of a Bose-Einstein Condensate in a Toroidal Trap*, Phys. Rev. Lett. **99**, 260401 (2007) (cit. on pp. 53, 70).

[166] D. M. Stamper-Kurn and M. Ueda, *Spinor Bose gases: Symmetries, magnetism, and quantum dynamics*, Rev. Mod. Phys. **85**, 1191–1244 (2013) (cit. on p. 53).

[167] P. G. Kevrekidis, D. J. Frantzeskakis, and R. Carretero-González, eds., *Emergent Nonlinear Phenomena in Bose-Einstein Condensates*, Vol. 45, Atomic, Optical, and Plasma Physics (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008) (cit. on p. 53).

[168] M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch, and I. Bloch, *Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms*, Nature **415**, 39–44 (2002) (cit. on p. 53).

[169] M. Albiez, R. Gati, J. Fölling, S. Hunsmann, M. Cristiani, and M. K. Oberthaler, *Direct Observation of Tunneling and Nonlinear Self-Trapping in a Single Bosonic Josephson Junction*, Phys. Rev. Lett. **95**, 010402 (2005) (cit. on p. 53).

[170] S. Levy, E. Lahoud, I. Shomroni, and J. Steinhauer, *The a.c. and d.c. Josephson effects in a Bose–Einstein condensate*, Nature **449**, 579–583 (2007) (cit. on p. 53).

[171] M. Lewenstein, A. Sanpera, and V. Ahufinger, *Ultracold Atoms in Optical Lattices* (Oxford University Press, 2012) (cit. on p. 53).

[172] A. D. Cronin, J. Schmiedmayer, and D. E. Pritchard, *Optics and interferometry with atoms and molecules*, Rev. Mod. Phys. **81**, 1051–1129 (2009) (cit. on p. 53).

[173] N. P. Proukakis and B. Jackson, *Finite-temperature models of Bose–Einstein condensation*, J. Phys. B: At. Mol. Opt. **41**, 203002 (2008) (cit. on p. 53).

[174] M. P. A. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher, *Boson localization and the superfluid-insulator transition*, Phys. Rev. B **40**, 546–570 (1989) (cit. on p. 53).

[175] C. D'Errico, E. Lucioni, L. Tanzi, L. Gori, G. Roux, I. P. McCulloch, T. Giamarchi, M. Inguscio, and G. Modugno, *Observation of a Disordered Bosonic Insulator from Weak to Strong Interactions*, Phys. Rev. Lett. **113**, 095301 (2014) (cit. on p. 53).

[176] A. D. K. Finck, J. P. Eisenstein, L. N. Pfeiffer, and K. W. West, *Exciton Transport and Andreev Reflection in a Bilayer Quantum Hall System*, Phys. Rev. Lett. **106**, 236807 (2011) (cit. on p. 53).

[177] H. Yokoyama, *Physics and Device Applications of Optical Microcavities*, Science **256**, 66–70 (1992) (cit. on p. 53).

[178] J. Kasprzak, M. Richard, S. Kundermann, A. Baas, P. Jeambrun, J. M. J. Keeling, F. M. Marchetti, M. H. Szymańska, R. André, J. L. Staehli, et al., *Bose–Einstein condensation of exciton polaritons*, Nature **443**, 409–414 (2006) (cit. on p. 53).

[179] R. Balili, V. Hartwell, D. Snoke, L. Pfeiffer, and K. West, *Bose-Einstein Condensation of Microcavity Polaritons in a Trap*, Science **316**, 1007–1010 (2007) (cit. on p. 53).

[180] J. Klaers, J. Schmitt, F. Vewinger, and M. Weitz, *Bose–Einstein condensation of photons in an optical microcavity*, Nature **468**, 545–548 (2010) (cit. on p. 53).

[181] S. O. Demokritov, V. E. Demidov, O. Dzyapko, G. A. Melkov, A. A. Serga, B. Hillebrands, and A. N. Slavin, *Bose–Einstein condensation of quasi-equilibrium magnons at room temperature under pumping*, Nature **443**, 430–433 (2006) (cit. on p. 53).

[182] K. Hebeler, J. M. Lattimer, C. J. Pethick, and A. Schwenk, *Equation of State and Neutron Start Properties Constrained by Nuclear Physics and Observation*, Ap. J. **773**, 11 (2013) (cit. on p. 53).

[183] A. Griffin, D. W. Snoke, and S. Stringari, eds., *Bose-Einstein Condensation* (Cambridge University Press, 1995) (cit. on p. 53).

[184] P. Sikivie and Q. Yang, *Bose-Einstein Condensation of Dark Matter Axions*, Phys. Rev. Lett. **103**, 111301 (2009) (cit. on p. 53).

[185] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Pseudoparticles*, Phys. Rev. Lett. **38**, 1440–1443 (1977) (cit. on p. 53).

[186] C. Rovelli, *Notes for a brief history of quantum gravity*, (2000) (cit. on p. 53).

[187] L. E. Ballentine, *Quantum Mechanics* (Wolrd Scientific, 2014) (cit. on pp. 54–55).

[188] E. Fermi, *Motion of neutrons in hydrogeous substances*, Ricerca Scientifica **7**, 13–52 (1936) (cit. on p. 55).

[189] E. L. Hill, *Hamilton's principle and the conservation theorems of mathematical physics*, Rev. Mod. Phys. **23**, 253–260 (1951) (cit. on pp. 55, 60).

[190] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Photons and Atoms* (Wiley, Weinheim, 1997) (cit. on pp. 55, 60, 114).

[191] T. Lancaster and S. J. Blundell, *Quantum Field Theory for the Gifted Amateur* (Oxford University Press, 2014) (cit. on p. 56).

[192] E. P. Gross, *Structure of a quantized vortex in boson systems*, Nuovo Cimento **20**, 454–477 (1961) (cit. on p. 56).

[193] L. P. Pitaevskii, *Vortex Lines in an imperfect Bose gas*, Sov. Phys. JETP **13**, 451–454 (1961) (cit. on p. 56).

[194] Y. Wu, X. Yang, C. P. Sun, X. J. Zhou, and Y. Q. Wang, *Theory of four-wave mixing with matter waves without the undepleted pump approximation*, Phys. Rev. A **61**, 043604 (2000) (cit. on pp. 57, 62).

[195] C. Sun, C. Hang, G. Huang, and B. Hu, *Investigation of Four-Wave Mixing of Matter Waves in Bose-Einstein Condensates*, Mod. Phys. Lett. B **18**, 375–383 (2004) (cit. on pp. 59, 62).

[196] E. V. Goldstein and P. Meystre, *Quantum theory of atomic four-wave mixing in Bose-Einstein condensates*, Phys. Rev. A **59**, 3896–3901 (1999) (cit. on p. 60).

[197] J. P. Burke, P. S. Julienne, C. J. Williams, Y. B. Band, and M. Trippenbach, *Four-wave mixing in Bose-Einstein condensate systems with multiple spin states*, Phys. Rev. A **70**, 033606 (2004) (cit. on p. 60).

[198] D. Pertot, B. Gadway, and D. Schneble, *Collinear Four-Wave Mixing of Two-Component Matter Waves*, Phys. Rev. Lett. **104**, 200402 (2010) (cit. on pp. 60, 70).

[199] K. M. Hilligsøe and K. Mølmer, *Phase-matched four wave mixing and quantum beam splitting of matter waves in a periodic potential*, Phys. Rev. A **71**, 041602 (2005) (cit. on p. 60).

[200] P. Martin, B. Oldaker, A. Miklich, and D. Pritchard, *Bragg scattering of atoms from a standing light wave*, Phys. Rev. Lett. **60**, 515–518 (1988) (cit. on p. 61).

[201] M. Kozuma, L. Deng, E. W. Hagley, J. Wen, R. Lutwak, K. Helmerson, S. L. Rolston, and W. D. Phillips, *Coherent Splitting of Bose-Einstein Condensed Atoms with Optically Induced Bragg Diffraction*, Phys. Rev. Lett. **82**, 871–875 (1999) (cit. on p. 61).

[202] A. Neumann, M. Gebbe, and R. Walser, *Aberrations in (3+1)-dimensional Bragg diffraction using pulsed Laguerre-Gaussian laser beams*, Phys. Rev. A **103**, 43306 (2021) (cit. on pp. 61, 133–134).

[203] P. Reineker, M. Schulz, B. M. Schulz, and R. Walser, *Theoretische Physik*, 2nd ed. (Wiley-VCH, Weinheim, 2021) (cit. on p. 65).

[204] E. Noether, *Invariante Variationsprobleme*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, 235–257 (1918) (cit. on p. 65).

[205] B. Josephson, *Possible new effects in superconductive tunnelling*, Phys. Lett. **1**, 251–253 (1962) (cit. on p. 66).

[206] F. Tafuri, ed., *Fundamentals and Frontiers of the Josephson Effect*, Vol. 286, Springer Series in Materials Science (Springer International Publishing, Cham, 2019) (cit. on p. 66).

[207] S. Shapiro, *Josephson Currents in Superconducting Tunneling: The Effect of Microwaves and Other Observations*, Phys. Rev. Lett. **11**, 80–82 (1963) (cit. on p. 66).

[208] P. A. Braun, *Discrete semiclassical methods in the theory of Rydberg atoms in external fields*, Rev. Mod. Phys. **65**, 115–161 (1993) (cit. on p. 66).

[209] S. Raghavan, A. Smerzi, S. Fantoni, and S. R. Shenoy, *Coherent oscillations between two weakly coupled Bose-Einstein condensates: Josephson effects, π oscillations, and macroscopic quantum self-trapping*, Phys. Rev. A **59**, 620–633 (1999) (cit. on p. 66).

[210] C. T. Weiß, *Reversible Dynamics of Cold Quantum Gases and Irreversible Coupling to Carbon Nanotubes*, Ph.D. Thesis (Eberhard-Karls-Universität Tübingen, 2010) (cit. on p. 66).

[211] F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, *NIST Digital Library of Mathematical Functions*, Accessed: Nov. 28 2023 (cit. on pp. 68, 107, 109, 117).

[212] L. Amico, A. Osterloh, and F. Cataliotti, *Quantum Many Particle Systems in Ring-Shaped Optical Lattices*, Phys. Rev. Lett. **95**, 063201 (2005) (cit. on p. 70).

[213] A. L. Fetter, *Rotating trapped Bose-Einstein condensates*, Rev. Mod. Phys. **81**, 647–691 (2009) (cit. on p. 70).

[214] H. M. Cataldo and D. M. Jezek, *Bose-Hubbard model in a ring-shaped optical lattice with high filling factors*, Phys. Rev. A **84**, 013602 (2011) (cit. on p. 70).

[215] A. M. Rey, K. Burnett, I. I. Satija, and C. W. Clark, *Entanglement and the Mott transition in a rotating bosonic ring lattice*, Phys. Rev. A **75**, 063616 (2007) (cit. on p. 70).

[216] P. L. Halkyard, M. P. A. Jones, and S. A. Gardiner, *Rotational response of two-component Bose-Einstein condensates in ring traps*, Phys. Rev. A **81**, 061602 (2010) (cit. on p. 70).

[217] J. Williams, R. Walser, J. Cooper, E. A. Cornell, and M. Holland, *Excitation of a dipole topological state in a strongly coupled two-component Bose-Einstein condensate*, Phys. Rev. A **61**, 033612 (2000) (cit. on p. 70).

[218] W. Demtröder, *Laser Spectroscopy* (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008) (cit. on pp. 70–71).

[219] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Springer Berlin Heidelberg, Berlin, Heidelberg, 1957) (cit. on p. 71).

[220] D. S. Hall, M. R. Matthews, J. R. Ensher, C. E. Wieman, and E. A. Cornell, *Dynamics of Component Separation in a Binary Mixture of Bose-Einstein Condensates*, Phys. Rev. Lett. **81**, 1539–1542 (1998) (cit. on p. 71).

[221] A. Hirose, *Complex-Valued Neural Networks*, 2nd ed., Vol. 400, Studies in Computational Intelligence (Springer Berlin Heidelberg, Berlin, Heidelberg, 2012) (cit. on pp. 77, 81).

[222] T. Nitta, *Solving the XOR problem and the detection of symmetry using a single complex-valued neuron*, Neural Networks **16**, 1101–1105 (2003) (cit. on pp. 81, 124).

[223] M. Lukoševičius and H. Jaeger, *Reservoir computing approaches to recurrent neural network training*, Comput. Sci. Rev. **3**, 127–149 (2009) (cit. on p. 92).

[224] K. Nakajima and I. Fischer, eds., *Reservoir Computing*, Natural Computing Series (Springer Singapore, Singapore, 2021) (cit. on p. 92).

[225] S. Ortín, M. C. Soriano, L. Pesquera, D. Brunner, D. San-Martín, I. Fischer, C. R. Mirasso, and J. M. Gutiérrez, *A Unified Framework for Reservoir Computing and Extreme Learning Machines based on a Single Time-delayed Neuron*, Sci. Rep. **5**, 14945 (2015) (cit. on p. 92).

[226] G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, and A. Hirose, *Recent advances in physical reservoir computing: A review*, Neural Netw. **115**, 100–123 (2019) (cit. on p. 92).

[227] G. Marcucci, D. Pierangeli, and C. Conti, *Theory of Neuromorphic Computing by Waves: Machine Learning by Rogue Waves, Dispersive Shocks, and Solitons*, Phys. Rev. Lett. **125**, 093901 (2020) (cit. on p. 92).

[228] G. Milano, G. Pedretti, K. Montano, S. Ricci, S. Hashemkhani, L. Boarino, D. Ielmini, and C. Ricciardi, *In materia reservoir computing with a fully memristive architecture based on self-organizing nanowire networks*, Nat. Mater. **21**, 195–202 (2022) (cit. on p. 92).

[229] B. Fischer, M. Chemnitz, Y. Zhu, N. Perron, P. Roztocki, B. MacLellan, L. Di Lauro, A. Aadhi, C. Rimoldi, T. H. Falk, et al., *Neuromorphic Computing via Fission-based Broadband Frequency Generation*, Adv. Sci. (2023) (cit. on p. 92).

[230] T.-P. Lee, C. Burrus, and B. Miller, *A stripe-geometry double-heterostructure amplified-spontaneous-emission (superluminescent) diode*, IEEE J. Quantum Electron. **9**, 820–828 (1973) (cit. on pp. 103–104).

[231] M. Blazek, *Thermisch und kohärent: Erzeugung neuartiger Lichtzustände mittels Quantenpunkt-Superlumineszenzdioden*, PhD thesis (Technische Universität Darmstadt, 2011) (cit. on pp. 103, 118–119).

[232] F. Boitier, A. Godard, E. Rosencher, and C. Fabre, *Measuring photon bunching at ultrashort timescale by two-photon absorption in semiconductors*, Nat. Phys. **5**, 267–270 (2009) (cit. on p. 104).

[233] A. Uskov, T. Berg, and J. Mork, *Theory of Pulse-Train Amplification Without Patterning Effects in Quantum-Dot Semiconductor Optical Amplifiers*, IEEE J. Quantum Electron. **40**, 306–320 (2004) (cit. on p. 104).

[234] P. Bardella, M. Rossetti, and I. Montrosset, *Modeling of Broadband Chirped Quantum-Dot Super-Luminescent Diodes*, IEEE J. Sel. Top. Quantum Electron. **15**, 785–791 (2009) (cit. on p. 104).

[235] A. F. Forrest, M. Krakowski, P. Bardella, and M. A. Cataluna, *Wide and tunable spectral asymmetry between narrow and wide facet outputs in a tapered quantum-dot superluminescent diode*, Opt. Express **28**, 846 (2020) (cit. on p. 104).

[236] M. Rossetti, P. Bardella, and I. Montrosset, *Time-Domain Travelling-Wave Model for Quantum Dot Passively Mode-Locked Lasers*, IEEE J. Quantum Electron. **47**, 139–150 (2011) (cit. on p. 104).

[237] Z. Q. Li and Z. M. S. Li, *Comprehensive Modeling of Superluminescent Light-Emitting Diodes*, IEEE J. Quantum Electron. **46**, 454–461 (2010) (cit. on p. 104).

[238] F. Friedrich, W. Elsäßer, and R. Walser, *Emission spectrum of broadband quantum dot superluminescent diodes*, Opt. Commun. **458**, 124449 (2020) (cit. on p. 104).

[239] S. Hartmann, F. Friedrich, A. Molitor, M. Reichert, W. Elsäßer, and R. Walser, *Tailored quantum statistics from broadband states of light*, New J. Phys. **17**, 043039 (2015) (cit. on p. 104).

[240] A. Jechow, M. Seefeldt, H. Kurzke, A. Heuer, and R. Menzel, *Enhanced two-photon excited fluorescence from imaging agents using true thermal light*, Nat. Photonics **7**, 973–976 (2013) (cit. on p. 104).

[241] J. Kiethe, A. Heuer, and A. Jechow, *Second-order coherence properties of amplified spontaneous emission from a high-power tapered superluminescent diode*, Laser Phys. Lett. **14**, 086201 (2017) (cit. on p. 104).

[242] F. Friedrich, *Hybrid Coherent Light: Modeling light-emitting quantum dot superluminescent diodes*, PhD thesis (Technische Universität Darmstadt, 2019) (cit. on p. 104).

[243] G. E. Uhlenbeck and L. S. Ornstein, *On the Theory of the Brownian Motion*, Phys. Rev. **36**, 823–841 (1930) (cit. on p. 104).

[244] R. J. Glauber, *The Quantum Theory of Optical Coherence*, Phys. Rev. **130**, 2529–2539 (1963) (cit. on p. 105).

[245] N. Wiener, *Generalized harmonic analysis*, Acta Math. **55**, 117–258 (1930) (cit. on p. 105).

[246] A. Khintchine, *Korrelationstheorie der stationären stochastischen Prozesse*, Math. Ann. **109**, 604–615 (1934) (cit. on p. 105).

[247] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, 1995) (cit. on p. 105).

[248] S. Hartmann and W. Elsäßer, *A novel semiconductor-based, fully incoherent amplified spontaneous emission light source for ghost imaging*, Sci. Rep. **7**, 41866 (2017) (cit. on pp. 108–111, 114).

[249] S. Blumenstein, *Classical ghost imaging with opto-electronic light sources*, PhD Thesis (Technische Universität Darmstadt, 2017) (cit. on pp. 108–109, 111).

[250] G. Süssmann, *Uncertainty Relation: From Inequality to Equality*, Z. Naturforsch. **52**, 49–52 (1997) (cit. on p. 110).

[251] W. P. Schleich, *Quantum Optics in Phase Space* (Wiley, Berlin, 2001) (cit. on p. 110).

[252] R. Loudon, *The quantum theory of light*, 3rd ed. (Oxford University Press, Oxford, 2000) (cit. on p. 113).

[253] R. Walser and P. Zoller, *Laser-noise-induced polarization fluctuations as a spectroscopic tool*, Phys. Rev. A **49**, 5067–5077 (1994) (cit. on p. 113).

[254] P. Meystre and M. Sargent, *Elements of Quantum Optics*, edited by P. Meystre and M. Sargent (Springer Berlin Heidelberg, Berlin, Heidelberg, 2007) (cit. on pp. 113, 116).

[255] S. Hartmann, A. Molitor, M. Blazek, and W. Elsäßer, *Tailored first- and second-order coherence properties of quantum dot superluminescent diodes via optical feedback*, Opt. Lett. **38**, 1334 (2013) (cit. on p. 119).

[256] D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, et al., *Optical Coherence Tomography*, Science **254**, 1178–1181 (1991) (cit. on pp. 120, 126).

[257] K. Böhm, R. Ulrich, P. Russer, and E. Weidel, *Low-noise fiber-optic rotation sensing*, Opt. Lett. **6**, 64 (1981) (cit. on pp. 120, 126).

[258] T. B. Pittman, Y. H. Shih, D. V. Strekalov, and A. V. Sergienko, *Optical imaging by means of two-photon quantum entanglement*, Phys. Rev. A **52**, R3429–R3432 (1995) (cit. on p. 120).

[259] A. Gatti, M. Bache, D. Magatti, E. Brambilla, F. Ferri, and L. A. Lugiato, *Coherent imaging with pseudo-thermal incoherent light*, J. Mod. Opt. **53**, 739–760 (2006) (cit. on p. 120).

[260] B. Paroli, G. Martini, M. Potenza, M. Siano, M. Mirigliano, and P. Milani, *Solving classification tasks by a receptron based on nonlinear optical speckle fields*, Neural Netw. **166**, 634–644 (2023) (cit. on p. 120).

[261] G. N. Milstein, *Approximate Integration of Stochastic Differential Equations*, Theory Probab. Its Appl. **19**, 557–562 (1975) (cit. on p. 130).

LIST OF PUBLICATIONS

- K. N. Hansmann, R. Walser
  Forming complex neurons by four-wave mixing in a Bose-Einstein condensate
  Physical Review A **12**, 013302 (2024)
  https://doi.org/10.1103/PhysRevA.109.013302

- K. N. Hansmann, F. Dommermuth, W. Elsäßer, R. Walser
  Silent White Light
  Submitted to Physical Review Research (2023)
  https://doi.org/10.48550/arXiv.2310.06834

- K. N. Hansmann, R. Walser
  Stochastic Simulation of Emission Spectra and Classical Photon Statistics of Quantum Dot Superluminescent Diodes
  Journal of Modern Physics **12**, 22-34 (2021)
  https://doi.org/10.4236/jmp.2021.121003

CONFERENCE CONTRIBUTIONS, WORKSHOPS & SUMMER SCHOOLS

- **Poster**
  Four-Wave Mixing Neuron
  Frontiers of Quantum Metrology 2023, Santa Barbara, California, USA

- **Talk (Q 2.4)**
  Four-Wave Mixing Neuron
  DPG Frühjahrstagung 2023, Hannover, Germany

- **International School of Physics "Enrico Fermi"**
  Quantum Fluids of Light and Matter
  Societa Italiana di Fisica, Varenna, Italy, 01. - 07.07.2022

- **Lake Como School**
  Statistical Physics of Deep Learning
  Fondazione Alessandro Volta, Como, Italy, 13. - 17.06.2022

- **Bad Honnef Physics School**
  Ultracold Quantum Gases
  Deutsche Physikalische Gesellschaft, Bad Honnef, Germany, 08. - 14.08.2021

- **Talk (EJ-3.5)** (online due to COVID-19 pandemic)
  Silent White Light: Intensity Noise Suppression in Superluminescent Diodes
  2021 Conference on Lasers and Electro-Optics/Europe - European Quantum Electronics Virtual Conferences (CLEO/Europe-EQEC 2021)

- **Talk (Q 27.8)** (canceled due to COVID-19 pandemic)
  Stochastic Coherence Analysis of Superluminescent Diodes
  DPG Frühjahrstagung 2020, Hannover, Germany

- **Talk (Q 48.6)**
  Ghost Imaging with Broad-Area Superluminescent Diodes
  DPG Frühjahrstagung 2019, Rostock, Germany

- **Poster**
  Hybrid Coherent Light: Modeling Quantum-Dot Superluminescent Diodes
  Greenhorn-Meeting 2018, Heidelberg, Germany

- **Talk**
  Linewidth determination at Second Harmonic Generation
  Greenhorn-Meeting 2018, Heidelberg, Germany

DANKSAGUNG

Zum Abschluss dieser Arbeit - und meiner Promotion an der Technischen Universität Darmstadt - möchte ich die Gelegenheit nutzen und mich bei all denen bedanken, die in irgendeiner Art und Weise dazu beigetragen haben, dass die letzten etwas mehr als vier Jahre nicht nur eine lehr- und erfolgreiche, sondern auch eine sehr schöne Zeit waren.

Zuallererst gilt mein Dank meinem Doktorvater Prof. Dr. Reinhold Walser. Vielen Dank, dass Sie mir die Promotion in ihrer Arbeitsgruppe ermöglicht haben, für die vielen reichhaltigen Diskussionen, Ihr Vertrauen in mich, eine Vielzahl von Aufgaben gewissenhaft zu erledigen, Ihr so umfassendes physikalisches Wissen, das Sie so gerne mit mir geteilt haben, und nicht zuletzt die Möglichkeit die Welt zu bereisen und interessante Menschen und Orte kennenzulernen.

Darüber hinaus bedanke ich mich recht herzlich bei Herrn Prof. Enno Giese für die Bereitschaft als Zweitgutachter meiner Dissertation zu fungieren, sowie bei Herrn Prof. Gerhard Birkl und Herrn Prof. Bernhard Urbaszek für den Einsatz als weitere Prüfer in meiner Disputation.

Mein nächster Dank gilt allen ehemaligen und aktuellen Mitgliedern der AG „Theoretische Quantendynamik" mit denen ich seit meinem Einstieg in die Gruppe im Juni 2018 zusammenarbeiten durfte. Ein recht herzlicher Dank gilt Frau Schütze und Frau Kutschera, dafür dass man sich bei organisatorischen Fragen immer an sie wenden konnte, für ihre großen Kenntnisse über die innersten Abläufe der TU, sowie für die Organisation der ein oder anderen Feier. Ein weiterer Dank gebührt allen Mitarbeitern der AG, von denen ich vor allem zu Beginn meiner Zeit in der Gruppe wahnsinnig viel lernen durfte. Vielen Dank an Franzi, die mich noch während meiner Masterarbeit betreute, an Antje, von der ich so viel über Formatierung und Präsentationen gelernt habe wie von niemand anderem, und an Jan, an den man sich immer mit Fragen wenden konnte, egal ob physikalischer, computer-bezogener oder sonst welcher Natur. Über die letzten Jahre hinweg gilt Oleks ein ganz besonderer Dank dafür, dass er immer Zeit gefunden hat sein so reichhaltiges Wissen zu teilen und auch neben dem Alltäglichen gerne Zeit zusammen verbracht hat. Auch an alle, die in sonst irgendeiner Form Zeit mit mir in der AG verbracht haben: Vielen Dank.

Ein ganz besonderer Dank gebührt all denen, die meine Zeit in Darmstadt nicht nur lehrreich und arbeitsam gestaltet haben, sondern auch spaßig, gesellig und einfach nur schön. Ohne die ein oder andere Stunde Vergnügen nach einem langen Tag im Fachschaftsraum oder sonstwo wäre die Studienzeit deutlich trister gewesen. Dafür vielen Dank an Matze, Kuchi, Basti, Janina, Till, Max und Philipp und darauf, dass wir auch weiterhin die Zeit finden, schöne Stunden in Hotzenplotz, Theke und Co. zu verbringen.

Auch abseits von Darmstadt möchte ich mich bei allen bedanken, die mir über die letzten zehn Jahre - und auch schon viel länger - gute Freunde waren und für die nötige Abwechslung gesorgt haben. Leider (oder zum Glück) sind das wohl zu viele, um hier einzeln aufgezählt zu werden. Aber es sei euch allen versichert, dass ich euch allen sehr dankbar bin, für die coolen Aktionen, den absurden Quatsch und die leckeren Kaltgetränke.

Nun gilt mein Dank auch noch meiner Freundin Emma. Auch wenn du mich nicht den kompletten Weg begleitest hast, hast du vor allem in den letzten Monaten des nicht-endenwollenden Schreibens ganz maßgeblich dazu beigetragen, dass mir der Spaß nie verloren ging und dass ich immer etwas hatte worauf ich mich freuen konnte. Mille Grazie.

Zu guter Letzt gilt mein Dank meinen Eltern, Monika und Werner Hansmann. In einer Zeit, in der ich gezweifelt habe, ob ich den richtigen Weg eingeschlagen habe, habt ihr mich stets unterstützt und seid mir mit guten Ratschlägen zur Seite gestanden. Auch als die ganze Welt verrückt wurde, habt ihr mir die Möglichkeit gegeben, meine Ruhe bei euch zu finden, und mir über die vier Jahre Promotion, die zehn Jahre Uni und mein ganzes Leben immer die Sicherheit gegeben meinen Weg zu gehen. Dafür bin ich euch unendlich dankbar und freue mich euch diese Arbeit als kleines Dankeschön zu widmen.

ERKLÄRUNG LAUT PROMOTIONSORDNUNG

§8 ABS. 1 LIT. C PROMO

Ich versichere hiermit, dass die elektronische Version meiner Dissertation mit der schriftlichen Version übereinstimmt

§8 ABS. 1 LIT. D PROMO

Ich versichere hiermit, dass zu einem vorherigen Zeitpunkt noch keine Promotion versucht wurde. In diesem Fall sind nähere Angaben über Zeitpunkt, Hochschule, Dissertationsthema und Ergebnis dieses Versuchs mitzuteilen.

§9 ABS. 1 PROMO

Ich versichere hiermit, dass die vorliegende Dissertation selbstständig und nur unter Verwendung der angegebenen Quellen verfasst wurde.

§9 ABS. 2 PROMO

Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Darmstadt, den 06.12.2023

Kai Niklas Hansmann