

Characterizing Photoelectron Momentum Distributions

Comparing Spherical Harmonic Reconstruction Methods

Nora Randsalu



LUND
UNIVERSITY

September 10, 2026

Main supervisor: Mathieu Gisselbrecht
Co-supervisor: Emilia My Kjærdsdam Telléus
Examiner: Francesca Curbis

Department of Physics
Faculty of Engineering LTH, Lund University

Abstract

This Master's thesis explores the reconstruction of photoelectron angular distributions using spherical harmonic decomposition with a view of characterize the polarization state of the ionizing light field. Photoelectron angular distributions describe the measured angular distribution of emitted electrons, which follows from the probability of detecting an electron at different angles. These distributions contain information about the polarization properties of the ionizing field.

Three different reconstruction approaches are investigated. Firstly, a direct fitting method is explored, where a coherent sum of spherical harmonics is directly fitted to the angular distribution. Secondly, a projective method is investigated, where the distribution is first projected onto an orthogonal basis of spherical harmonics before being fitted. Finally, a machine learning approach based on a convolutional neural network with a physics-informed decoder is considered. The performance of each method is examined by investigating the reconstruction performance for several angular distributions corresponding to different polarization states, decreasing signal intensities, and varying noise levels. Both additive white Gaussian noise and Poisson noise are used in order to evaluate the methods under experimental-like conditions.

The results show that all three methods are able to accurately reconstruct the parameters describing the polarization state for both high-intensity and moderate-noise distributions. The direct fitting method gives the highest accuracy for varying signal intensity, while the projective method is most robust against Poisson noise. The machine learning method achieves comparable reconstruction performance to the two other methods, and provides more consistent predictions for different distributions, whereas the numerical optimization-based methods show a stronger dependence on the specific polarization state. A limitation observed for all methods is the difficulty in accurately reconstructing polarization states where one of the spherical harmonic coefficients used to describe the angular distribution approach zero.

Overall, the findings give insight into the potentials and limitations of the different reconstruction methods and their performance under different conditions. Possible improvements of the current methods, as well as future applications to more complex systems, are also discussed.

Popular Science Summary

Light is an important tool when we want to study the dynamics of electrons in atoms and other small quantum systems. In particular, very short light pulses are useful because they are short enough to capture processes that happen on the timescale of electron motion. In attosecond science, attosecond pulses are used, which are one quintillionth of a second long.

One important property of light is its polarization, which describes the direction in which the electric field of the light oscillates. You might be familiar with the concept from 3D cinema. The two lenses in a pair of 3D glasses filter different polarizations so that each eye receives a differently polarized image, which when combined in our brains will make the image appear three dimensional. But polarization is not only useful in 3D cinema, it can also be used in scientific research. When a photon with sufficient energy is absorbed by the system, an electron can be emitted. These photoelectrons can be detected, and the probability of detecting an electron at different angles forms a photoelectron angular distribution. The shape of this distribution depends on the properties of the ionizing light, meaning that different polarization states produce different patterns. This Master's thesis investigates how these photoelectron angular distributions can be used to determine the polarization of the light that caused the electron emission.

To describe these patterns, spherical harmonics can be used, which are mathematical functions that can be combined in different ways to describe spherical shapes and patterns. For different polarizations of the light, there will be different coefficients connected to the spherical harmonics. In other words, these coefficients contain information about the light. Therefore, if these coefficients can be retrieved from a given angular distribution, they can be used to determine the polarization of the light that produced the distribution. This coefficient reconstruction is the main focus of this thesis.

Three different reconstruction approaches are investigated in this thesis. The first method directly compares a reconstructed angular distribution with the input distribution, and adjusts the coefficients in the reconstructed distribution until the best agreement has been found. For the second method, the angular distribution is first projected onto a new basis of spherical harmonics before being compared to the input distribution. Instead of comparing the full angular distribution pixel by pixel, this gives us a smaller set of coefficients that describe the main features of the distribution, which reduces the number of parameters that are needed to be optimized. The third and final method uses machine learning, more specifically a convolutional neural network. The network is trained to recognize the relationship between the angular distributions and their corresponding coefficients.

To evaluate the methods, a large variety of simulated angular distributions describing different polarization states is used. The methods are also tested for varying signal intensities and noise levels, to investigate how well they could perform under conditions similar to those in a real experiment. The types of noise considered are Gaussian noise and Poisson noise. Gaussian noise can be seen as background noise from false particle detections, while Poisson noise describes the statistical fluctuations that arise from having a limited number of detected electrons.

The results show that all three methods can accurately reconstruct the spherical harmonic coefficients for high intensities and moderate noise, especially for relatively high levels of Gaussian noise. The direct method generally maintains the best performance as the signal intensity is reduced, whereas the projective method is most robust, especially with respect to Poisson noise. The machine learning method achieves comparable performance to the other two methods, and gives more consistent results for different polarization states. The two other methods appear to be more dependent on the specific polarization state being reconstructed. They perform much better for some polarization states than for others.

Overall, the results demonstrate that photoelectron angular distributions can be used to retrieve information about the polarization of the light used for electron emission. The comparison of the three methods highlights their different strengths and limitations, and provides a basis for further development of current reconstruction methods. Although only relatively simple systems are investigated in this work, the results provide insights which could be useful for further investigation of more complex processes and the use of experimental data. It could especially provide guidance on what signal quality is needed when collecting experimental data.

Acknowledgments

I would like to thank my supervisor Emilia My Kjærdsdam Telléus for her guidance and support throughout the project, which helped me to finally bring this thesis to completion. I am also very grateful to Mathieu Gisselbrecht and Samuel Eriksson for their valuable feedback and input.

Nora Randsalu

Contents

1	Introduction	9
1.1	Motivation and Purpose	9
1.2	Problem Statement	9
1.3	Research Questions	10
1.4	Contributions	10
2	Background	10
2.1	Simulation Framework	10
2.1.1	Dipole Approximation and Single-Active-Electron Model	10
2.1.2	Strong Field Approximation (SFA)	11
2.2	Experimental Context	12
2.2.1	High-order Harmonic Generation (HHG)	12
2.2.2	Photoionization Dynamics	13
2.2.3	Selection Rules for One-Photon Photoionization	14
2.2.4	Angular Distributions	15
2.2.5	Polarization Parametrization	16
3	Methodology	17
3.1	Objectives of Computational Work	18
3.2	Data Simulation	18
3.2.1	Assumptions and Simplifications	18
3.2.2	Generation of Angular Distribution	18
3.2.3	Additive Gaussian Noise	20
3.2.4	Poisson Noise	21
3.3	Reconstructional methods	22
3.3.1	Spherical Harmonic Decomposition	22
3.3.2	Physical Constraints	22
3.3.3	Numerical Optimization Procedure	23
3.3.4	Direct Fitting Method	23
3.3.5	Projective Method	24
3.3.6	Machine Learning Approach	25
3.4	Evaluation Metrics	28

3.4.1	Reconstruction Error	28
3.4.2	Coefficient Comparison	29
4	Results	29
4.1	Dependence on Signal Intensity	29
4.2	Noise Robustness Analysis	33
4.3	Comparison Reconstructed Coefficients	39
4.4	Limitations in Reconstructual Methods	44
5	Conclusions	45
5.1	Summary of Results	45
5.2	Improvements and Future Work	46
5.2.1	Improvement of Current Methods	46
5.2.2	Experimental Application of Results	47
5.2.3	Extension of the Physical Scope	47
A	Basic Quantum Mechanics	51
A.1	Commutation Relations	51
A.2	Parity Operator	51
A.3	Orbital Angular Momentum	51
A.4	Spin Angular Momentum	52
A.5	Total Angular Momentum	52
B	Quantum Mechanical Representation of Atoms	52
B.1	Hydrogenic Atoms	52
B.2	Solution of the Angular Part of the Schrödinger Equation	52
C	MATLAB Scripts	54
C.1	Direct Fitting Method	54
C.2	Projective Method	54
C.3	Machine Learning Method	56

List of Figures

2.1	Illustration of High-order Harmonic Generation	13
2.2	Angular probability density for p-orbital electrons	16
2.3	Illustration of inclination angle α	17
2.4	Polarization scan for helium ionization from s -subshell	17
3.1	Radial distribution along with energy ranges	20
3.2	Examples of angular distributions	20
3.3	Angular distribution under additive Gaussian noise	21
3.4	Angular distribution under Poisson noise	22
3.5	Illustration of the projective method	25
4.1	Comparison of R^2 for reconstructional methods at varying signal intensity	30
4.2	Comparison of R^2 for full reconstruction and the Gaunt basis projection at varying signal intensity	31
4.3	Comparison of R^2 for different cases at varying signal intensities	31
4.4	Comparison of angular distributions for different cases at varying signal intensities	32
4.5	Comparison of different initial guesses	33
4.6	Comparison of rescaling factors for the projective method	33
4.7	Comparison of R^2 for reconstructional methods at varying noise levels	36
4.8	Comparison of R^2 for full reconstruction and the Gaunt basis projection at varying noise levels	37
4.9	Comparison of R^2 for different cases at varying Gaussian noise	38
4.10	Comparison of R^2 for different cases at varying Poisson noise	39
4.11	Comparison of reconstructed physical coefficients for varying inclination angle	41
4.12	Comparison of reconstructed physical coefficients for varying relative phase	42
4.13	Comparison of reconstructed physical coefficients in a zoomed-in view	42
4.14	Comparison of reconstructed physical coefficients under Gaussian noise	43
4.15	Comparison of reconstructed physical coefficients under Poisson noise	44

List of Tables

3.1	SFA simulation parameters	19
3.2	Optimization parameters	23
3.3	CNN architecture	27
3.4	CNN hyperparameters	27
3.5	Coefficient of determination interpretation	28
4.1	Result signal intensity dependence	30
4.2	Result additive Gaussian noise analysis	34
4.3	Result Poisson noise analysis	35

1 Introduction

Understanding and controlling light-matter interactions requires knowledge of the properties of the ionizing light field, which can be accessed through photoelectron angular distributions obtained from photoionization experiments. This chapter introduces the motivation and purpose of the project, as well as outlines the project's goals and contributions to the field of research.

1.1 Motivation and Purpose

When photons of sufficient energy are absorbed by a small quantum system, electrons will be emitted into well defined continuum states, with well defined momentum probability distributions [1]. By using ionizing light pulses with pulse durations on the attosecond timescale, light-matter interactions and electron dynamics can be studied in real time, as the timescale of energy splitting of electronic states, $\hbar/\Delta E$, is within this regime. Ongoing research on attosecond electron dynamics, has for example demonstrated fundamental phenomena such as wave-particle duality [2].

Through high-harmonic generation, attosecond pulses can be generated and used in photoionization experiments that produce photoelectrons [3]. The angular distributions of the photoelectron momentum contain information about the properties of the ionizing light field. Since the angular distribution can be described using spherical harmonics, the corresponding coefficients provide a way to characterize properties of the ionizing field, such as its polarization state [4].

The characterization of attosecond light fields is important for understanding and controlling light-matter interactions. However, extracting this information from photoelectron angular distributions is challenging, which motivates the development and evaluation of methods to reconstruct the physical parameters used to describe the ionizing field.

1.2 Problem Statement

The angular distribution gives the electron emission probability as a function of emission angle, and can be described using a coherent sum of spherical harmonics representing the emitted electron wavefunction [4]. The polarization governs allowed transitions through selection rules, which in turn determines the contributing spherical harmonic components [1]. The corresponding coefficients therefore contain information about the polarization state of the ionizing light field. Hence, by extracting these coefficients from the measured angular distribution, the polarization properties can be characterized.

However, extracting the spherical harmonic coefficients from experimentally obtained distributions is challenging. The measured momentum distribution contains only the absolute square of the wavefunction, meaning that the information is not directly accessible. Also, different types of noise introduced during experiments, such as background noise or fluctuations due to limited measurement time, can further complicate the reconstruction. Developing robust reconstruction methods capable of recovering spherical harmonic components from such distributions is therefore an important problem in attosecond science.

This Master thesis project seeks to investigate photoelectron momentum distributions, as well as evaluate methods used for characterizing attosecond light-matter interactions by reconstructing the spherical harmonic decompositions used to describe photoelectron angular distributions. The work includes both the simulation of probability distributions of photoelectron momenta for specified light fields, and the development of reconstruction methods for decomposing these distributions into coherent sums of spherical harmonics. The goal is to evaluate how well the different approaches can recover the spherical harmonic coefficients and characterize the polarization state of the ionizing light pulses.

1.3 Research Questions

The main objectives to be investigated in this project are:

- To what extent can photoelectron momentum distributions be decomposed into coherent sums of spherical harmonics to characterize the polarization state of the ionizing light field?
- How do different reconstruction methods compare in reconstructing spherical harmonic coefficients from simulated momentum distributions?
- How robust are the different reconstruction methods when applied to noisy data intended to simulate experimental conditions?

1.4 Contributions

This project aims to contribute to the field of attosecond science, more specifically to the study of simple systems where selection rules apply and it is possible to obtain meaningful spherical harmonic decompositions from photoelectron angular distributions. By investigating the robustness of the methods with respect to increasing levels of noise, insights are hopefully obtained into the limits of reliable reconstruction from experimental data and a framework given for further development of more robust and optimized algorithm designs. In particular, the results can provide practical guidance with regard to minimum measurement time and tolerable level of background noise needed to obtain reliable reconstruction results.

2 Background

This chapter presents the theoretical background that is the foundation for the computational work. First, the framework used for the simulations is introduced, explaining the interaction between quantum mechanical systems and classical electromagnetic fields, and how Strong Field Approximation (SFA) can be used to model these processes. Thereafter, the experimental background is presented, including high-harmonic generation (HHG), photoionization processes, and selection rules governing transitions. The connection between the polarization state of the ionizing light field and the resulting photoelectron angular distribution is also discussed, showing how the angular distribution can be described using spherical harmonics.

2.1 Simulation Framework

The interaction between an electromagnetic field and an atomic system can lead to ionization. These processes can be described using the Strong-Field Approximation (SFA), which provides a simplified model of electron motion after ionization [5]. In this work, photoelectron momentum distributions are generated using SFA simulations, providing the simulated data used as input for the reconstruction methods.

2.1.1 Dipole Approximation and Single-Active-Electron Model

In semi-classical theory, light is treated as a classical electromagnetic field using Maxwell's equations, while the matter is described by quantum mechanics formalism. The electromagnetic field can affect quantum states, for example through photoionization induced by extreme-ultraviolet light (XUV). Quantum states can also affect the electromagnetic field, one example being the generation of XUV pulses from High-order Harmonic Generation (HHG) [5].

For interactions between a gas of atoms and a laser pulse described by a classical electromagnetic field, the problem can be simplified by assuming only one electron participates in the interaction at a given time, and

that this electron is initially in the ground state $|g\rangle$. The motion of the parent ion, to which the electron is bound, can also be neglected assuming the parent ion to be localized at $\mathbf{R}$. The electromagnetic field at the position of the electron can therefore be expressed using dipole approximation [6]:

$$\mathbf{E}(t) = \mathbf{E}(\mathbf{R}, t) \quad (2.1)$$

The parent ion and the electromagnetic field are assumed to not interact, apart from the contribution from static Coulomb potential. Hence, neglecting the parent ion gives Hamiltonian

$$H(t) = \frac{\mathbf{p}^2}{2m} + V_{\text{Coulomb}} - \mathbf{d} \cdot \mathbf{E}(t) \quad (2.2)$$

where m is the electron mass, $\mathbf{d} = q\mathbf{r} = -e\mathbf{r}$ is the dipole operator with electron charge $q = -e$, and the length gauge $\mathbf{d} \cdot \mathbf{E}(t)$ is used to express the interaction between the electron and the electromagnetic field [6].

Using the velocity gauge, the dynamics can instead be expressed using the vector potential $\mathbf{A}(t)$ [5]:

$$H(t) = \frac{1}{2} (\mathbf{p} + \mathbf{A}(t))^2 + V_{\text{Coulomb}} \quad (2.3)$$

The static and field free Hamiltonian can be written as

$$H_0 = \frac{\mathbf{p}^2}{2m} + V_{\text{Coulomb}} \quad (2.4)$$

and the time dependent interaction expressed in the length gauge as

$$V(t) = -\mathbf{d} \cdot \mathbf{E}_e(t) \quad (2.5)$$

the following time-dependent Schrödinger equation (TDSE) is given for states $|\psi(t)\rangle$ [6]:

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = [H_0 + V(t)] |\psi(t)\rangle \quad (2.6)$$

A full solution to the problem requires further approximation, since the numerical solution to the time-dependent Schrödinger equation is generally computational demanding. One such approach is the application of Strong Field Approximation (SFA) [7] [6].

2.1.2 Strong Field Approximation (SFA)

Strong Field Approximation (SFA) is a mathematical framework that allows the solution of equation 2.6 without having to use perturbation theory to describe the interaction term $V(t)$ [8] [9]. For SFA, the following assumptions are made:

1. The external field does not couple with any excited bound state, hence all bound states can be neglected except the ground state. It is therefore sufficient to consider a set of electronic states $\{|g\rangle, |\mathbf{p}_k\rangle\}$ where $|g\rangle$ is the ground state and $|\mathbf{p}_k\rangle$ is the electron's continuum state corresponding to kinetic momentum $\mathbf{p}_k = m\mathbf{v}$ [8].
2. The continuum states $|\mathbf{p}_k\rangle$ are eigenstates to Hamiltonian H_0 , with eigenvalues $\frac{\mathbf{p}_k^2}{2m}$ [8]. In other words, the Coulomb potential V_{Coulomb} is neglected [8] [7].
3. The amplitude $a(t)$ of the ground state must be known and is often approximated as $a(1) = 1$ for sufficiently short or moderate laser pulses. This approximation is valid if the ponderomotive energy U_p , defined as the average energy of an oscillating charged particle in a high frequency, spatially varying electromagnetic field, is lower than the saturation energy of system [8].

With this approximation, the state $|\psi(t)\rangle$ can be written as

$$|\psi(t)\rangle = e^{iI_p t/\hbar} (a(t)|g\rangle + \int d^3p_k b(\mathbf{p}_k, t)|\mathbf{p}_k\rangle), \quad H_0|g\rangle = E_0|g\rangle \quad (2.7)$$

where $a(t)$ describes the amplitude of the ground state and $b(\mathbf{p}_k, t)$ describes the amplitude of the field-free continuum states $|\mathbf{p}_k\rangle$. The exponential term represents the global phase oscillating with the ground state energy given by the ionization potential $E_0 = -I_p$, which for hydrogenic atoms is $I_p \approx \frac{Z^2}{2}$ [8] [5].

The field-free continuum states $|\mathbf{p}_k\rangle$ in equation 2.7 can be written as Volkov states, describing the interaction between the laser field in the velocity gauge [5]:

$$|\psi_V(\mathbf{p}_k, t)\rangle = \exp\left[-\frac{i}{\hbar}S(\mathbf{p}_k, t, t_0)\right]|\mathbf{p}_k\rangle, \quad S(\mathbf{p}_k, t, t_0) = \int_{t_0}^t \frac{1}{2}(\mathbf{p}_k + \mathbf{A}(t'))^2 dt' \quad (2.8)$$

The semiclassical action $S(\mathbf{p}_k, t, t_0)$ describes the relation between the classical motion of the electron and the phase of its quantum mechanical wave function, under the influence of vector potential $\mathbf{A}(t)$ [5].

2.2 Experimental Context

Attosecond pulses are commonly generated through high-order harmonic generation (HHG) and used to study electron dynamics through photoionization experiments, such as ionizing simple atomic systems [3]. The emitted electrons can then be detected in coincidence with the remaining ions, using three-dimensional momentum spectroscopy, producing photoelectron momentum distributions that contain information about the ionization process and the properties of the ionizing light field [10].

In this work, the simulated attosecond pulses are based on such experimental HHG sources, while the HHG process itself can be described within the Strong-Field Approximation (SFA) [11].

2.2.1 High-order Harmonic Generation (HHG)

In attosecond physics, electron dynamics occur on the time scale of a few hundred attoseconds. To observe these processes, light pulses of at least a similar time duration are required. Such short pulses contain a broad range of frequencies, since a shorter pulse duration requires a wider frequency range according to the uncertainty principle. One approach to reach the attosecond regime is High-order Harmonic Generation (HHG), where low frequency infrared (IR) laser pulses can be converted into high frequency extreme-ultraviolet (XUV) attosecond pulses [11] [7]. HHG process can be modelled using the SFA framework, and the generated XUV pulses can be used to photoionize atomic systems [11].

The HHG process is a non-linear optical process that can be understood in three steps, illustrated in figure 2.1. The first step of the HHG process is ionization induced by a strong laser field [12]. At sufficiently high laser field intensities, the potential barrier binding the electrons to the nucleus is reduced, making quantum tunneling possible. An electron can therefore penetrate the lowered potential barrier and escape the atom into the continuum as a free electron [5] [9].

The second step of HHG is electron propagation in the continuum, driven by the laser field [12]. After ionization, the electron is accelerated by the driving laser field, resulting in high kinetic energy [5] [9]. As the oscillating electric field of the electromagnetic wave evolves, the direction is reversed, driving the electron back toward its parent ion and enabling recollision [5].

The third step of the process is electron recombination with the parent ion, resulting in emission of short burst high-energy photons [12] [5]. During recombination, both the kinetic energy E_k of the electron and the ionization potential I_p are emitted, producing photons with energy given by [5]:

$$\hbar\omega = I_p + E_k \quad (2.9)$$

There are two important limitations that the three-step model does not account for. If the intensity of the driving laser field is too high, the atom becomes completely ionized, leaving no bound electrons and therefore no oscillating dipole moment, which is essential for the HHG process. Furthermore, if the wavelength of the driving field is too large, the spatial spreading of the free electron wave function during propagation will increase considerably. This will result in the recombination probability decreasing and becoming negligible. In order to satisfy above limitations, near infrared laser fields are typically used as the driving field in HHG [5].

HHG typically produce an attosecond pulse train, where one pulse is emitted every half optical cycle of the driving IR field [3] [13]. This periodic emission in time leads to a spectrum containing high-order harmonics of the IR frequency, which gives the process its name. Since the emission occurs every half cycle, the harmonics are separated by $2\omega_{\text{IR}}$ [5]. In the experiments relevant for this project, the driving IR pulses consist of only a few optical cycles, resulting in typically 2-3 XUV pulses [3]. However, for studying ultrafast electron dynamics, it is often desirable to have an isolated Single Attosecond Pulse (SAP) for each IR pulse rather than a pulse train. This can be achieved using grating techniques, such as controlling the Carrier Envelope Phase (CEP) of the driving IR pulse and polarization grating [5].

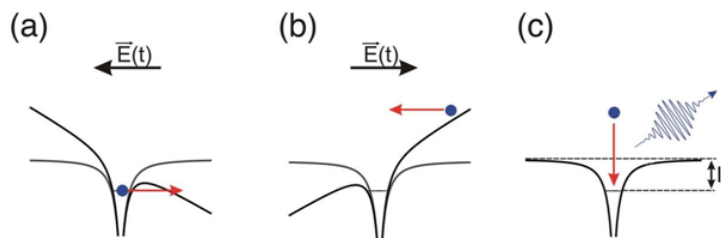


Figure 2.1: Illustration of the three-step model of HHG: (a) tunneling (b) acceleration (c) recombination [12]

2.2.2 Photoionization Dynamics

The development of HHG sources capable of generating attosecond pulses has enabled the study of photoionization dynamics with very high temporal resolution, providing insight into fundamental quantum phenomena and ultrafast electron dynamics in atoms and molecules [14] [2].

The ionization energy of a system is defined as lowest energy required to remove an electron from the system, and is equal to the lowest binding energy of the electron [15]. During photoionization, the electron absorbs energy from the electromagnetic field and transitions from a bound state into a continuum state with a well-defined momentum [3] [1]. The first ionization energy corresponds to the removal of one electron, such as the single ionization of helium considered in this work [15].

The photoionization process can be described by separating the electron dynamics into a bound part and a continuum part. Before ionization, the electron remains in the initial bound state with an amplitude $a(t)$ representing the probability that it still belongs to the atom. After ionization, the electron is described by a continuum amplitude $b(\mathbf{p}, t)$, that gives the probability of finding the electron with momentum $\mathbf{p}$ at time t . If only the initial momentum at time of ionization is considered, the time dependence can be omitted, and the probability of measuring an electron with specified momentum is given by the absolute square of the amplitude [8]:

$$P(\mathbf{p}) = |b(\mathbf{p})|^2 \quad (2.10)$$

The dynamics of $b(\mathbf{p}, t)$ can be derived using the time-dependent Schrödinger in equation 2.6, along with the expansion of the wave function presented in equation 2.7 and the projection onto continuum state $\langle \mathbf{p}_k |$. Furthermore, within the SFA framework, the interaction between the ionized electron and the remaining atomic potential is neglected, meaning that rescattering effects do not need to be included. Assuming that the ground state amplitude $a(t)$ is known, the solution for an initially empty continuum state, $b(\mathbf{p}_k, t_0) = 0$, can be written as [8]:

$$b(\mathbf{p}, t) = \frac{i}{\hbar} \int_{t_0}^t dt \mathbf{E}_e(s) \cdot \mathbf{d}_{\mathbf{p}+e\mathbf{A}_e(s),g} \exp \left\{ -\frac{i}{\hbar} S(\mathbf{p}, t, s) \right\} a(s) \quad (2.11)$$

The continuum amplitude can be interpreted as a sum of contributions for different ionization times. The term $\mathbf{E}_e(s) \cdot \mathbf{d}_{\mathbf{p}+e\mathbf{A}_e(s),g}$ describes the transition from the bound state $|g\rangle$ to the continuum, where $\mathbf{E}_e(s)$ is the ionizing electric field and $\mathbf{d}_{\mathbf{p}+e\mathbf{A}_e(s),g}$ is the dipole matrix element describing the bound-continuum state coupling. The exponential term contains the semiclassical action $S(\mathbf{p}_k, t, s)$, and describes the phase accumulated by the electron while it propagates in the continuum from ionization time s to the observation time t . The final photoelectron momentum distribution is obtained by integrating over all possible ionization times [8].

2.2.3 Selection Rules for One-Photon Photoionization

The transition amplitude for one-photon ionization from an initial bound ground state $|\phi_g\rangle$ to a final continuum state $|\phi_p\rangle$ is determined by the dipole matrix element $\langle\phi_p|T(t)|\phi_g\rangle$. By using first-order time-dependent perturbation theory, in which only the first-order contribution to the transition amplitude is retained, to describe the interaction between the atom and the electromagnetic field, the transition amplitude is given by [16] [6]:

$$a_{\mathbf{p},g} = -\frac{i}{\hbar} \int dt \langle\phi_p|T(t)|\phi_g\rangle e^{i\Omega_{\mathbf{p},g}t} \quad (2.12)$$

$\Omega_{\mathbf{p},g} = (E_{\mathbf{p}} - E_g)/\hbar$ is the Bohr angular frequency, with E_g and $E_{\mathbf{p}}$ denoting the energies of the initial and final states [16]. The operator $T(t)$ describes the interaction between the atom and the electromagnetic field which using electric dipole approximation it is given by

$$T(t) = e \mathbf{E}_e(t) \cdot \mathbf{r} \quad (2.13)$$

where e is the elementary charge, $\mathbf{r}$ is the electron position vector relative to the nucleus, and $\mathbf{E}_e(t)$ is the electric field of the ionizing pulse [16] [6]. For a field polarized in the (x, z) -plane, equation 2.12 can be written as:

$$a_{\mathbf{p},g} = -\frac{ie}{\hbar} \int dt e^{i\Omega_{\mathbf{p},g}t} [E_x(t) \langle\phi_p|x|\phi_g\rangle + E_z(t) \langle\phi_p|z|\phi_g\rangle] \quad (2.14)$$

In order to evaluate the transition amplitude, both the initial and final states can be expanded using spherical harmonics. Assuming the initial state is bound, and that the final state is described as a continuum state, the following expansions are obtained [17]:

$$|\phi_g\rangle = R_{n_i l_i}(r) Y_{l_i m_i}(\mathbf{r}), \quad |\phi_p\rangle = \sqrt{\frac{1}{k}} \sum_{l,m} i^l e^{-i\eta_l(k)} Y_{lm}^*(\mathbf{p}) R_{kl}(r) Y_{lm}(\mathbf{r}) \quad (2.15)$$

where $R_{n_i l_i}(r)$ and $R_{kl}(r)$ are the radial wavefunctions of the bound and continuum states respectively, $r = |\mathbf{r}|$ is the electron-nucleus distance, $k = |\mathbf{p}|/\hbar$ is the electron wave number, and $\eta_l(k)$ is the scattering phase shift. Note that the continuum wavefunction is expanded in spherical harmonics in momentum space.

The Cartesian coordinates x and z can be expressed in terms of spherical harmonics according to [18]:

$$z = \sqrt{\frac{4\pi}{3}} r Y_{10}(\mathbf{r}), \quad x = \sqrt{\frac{2\pi}{3}} r (Y_{1,-1}(\mathbf{r}) - Y_{11}(\mathbf{r})) \quad (2.16)$$

which allows for the transition amplitude to be separated into radial and angular contributions:

$$a_{\mathbf{p},g} = -\frac{ie}{\hbar} \sqrt{\frac{4\pi}{3}} \sum_{l,m} (-i)^l e^{-i\eta_l(k)} Y_{lm}(\mathbf{p}) \langle R_{kl}|r|R_{n_i l_i}\rangle \times \int dt e^{i\Omega_{\mathbf{p},g}t} \left[E_z(t) \langle lm|10|l_i m_i\rangle + \frac{E_x(t)}{\sqrt{2}} (\langle lm|1-1|l_i m_i\rangle - \langle lm|11|l_i m_i\rangle) \right] \quad (2.17)$$

Note that the shorthand notation $\langle lm|1m'|l_i m_i \rangle$ represents the angular part of the dipole matrix. It is given by the integral over the spherical harmonics of the initial bound state with angular quantum numbers l_i, m_i , the final continuum state with angular quantum numbers l, m , and the dipole operator. The index $m' = 0, \pm 1$ describes the different magnetic components of the dipole operator, which correspond to the different polarization components of the electric field [19].

The radial matrix element describes overlap between the initial and final states:

$$\langle R_{kl}|r|R_{n_i l_i} \rangle = \int_0^\infty R_{kl}^*(r) r R_{n_i l_i}(r) r^2 dr \quad (2.18)$$

whereas while the angular matrix element gives the angular momentum coupling between the states [19]:

$$\langle lm|1m'|l_i m_i \rangle = \int Y_{lm}^*(\mathbf{r}) Y_{1m'}(\mathbf{r}) Y_{l_i m_i}(\mathbf{r}) d\Omega \quad (2.19)$$

A transition is only possible when the angular matrix element is non-zero, which gives the dipole selection rules for transitions between bound and continuum states:

$$\langle lm|1m'|l_i m_i \rangle \neq 0 \iff \begin{cases} l = l_i \pm 1, \\ m = m_i + m' \end{cases} \quad (2.20)$$

These selection rules governs the allowed transitions between the initial bound state and the final continuum state during photoionization [13]. The first condition describes the change in orbital angular momentum caused by the dipole interaction, while the second condition describes the change in the magnetic quantum number depending on the polarization ionizing field. For example, for the ionization of helium from a s -subshell, where $l_i = 0$ and $m_i = 0$, only final states with $l = 1$ and $m = 0, \pm 1$ are allowed, corresponding to p -wave continuum states [1]. The magnetic quantum number m is determined by the polarization of the ionizing light. Similarly, ionization from a p -subshell with $l_i = 1$ can result in final continuum states with $l = 0$ or $l = 2$, corresponding to s - and d -states, respectively.

2.2.4 Angular Distributions

From equation 2.17, it can be seen that the transition amplitude can be expressed as a sum of terms, multiplied by spherical harmonics. The probability distribution of the emitted photoelectrons in momentum space can be obtained from the transition amplitude according to [4]:

$$P(\mathbf{p}) = |a(\mathbf{p})|^2 = \left| \sum_{l,m} c_{lm}(|\mathbf{p}|) Y_{lm}(\theta, \phi) \right|^2 \quad (2.21)$$

The coefficients $c_{l,m}$ determine the contribution of each spherical harmonic component to the total distribution. These complex coefficients contain information about the polarization state and relative phase of the ionizing light field, and therefore determine the shape of the resulting photoelectron angular distribution [4] [14]. This dependence can also be seen in equation 2.17, where the different polarization components of the electric field, contribute to different spherical harmonic terms.

The momentum vector $\mathbf{p}$ is described using spherical coordinates, where the direction of the emitted electron is given by angles (θ, ϕ) [4]. By considering a fixed momentum magnitude $|\mathbf{p}|$ corresponding to a specific energy, only the angular dependence of the momentum distribution remains and the photoelectron angular distribution can be written as:

$$P(\theta, \phi) = \left| \sum_{l,m} c_{l,m} Y_{l,m}(\theta, \phi) \right|^2 \quad (2.22)$$

For ionization from an s -subshell, the previously discussed dipole selection rules restrict the allowed final states to $l = 1$, and the angular distribution can be described by:

$$P(\theta, \phi) = |c_{1,-1}Y_{1,-1}(\theta, \phi) + c_{1,0}Y_{1,0}(\theta, \phi) + c_{1,1}Y_{1,1}(\theta, \phi)|^2 \quad (2.23)$$

The absolute values of the $m = \pm 1$ coefficients are assumed to be equal, $|c_{1,1}| = |c_{1,-1}|$, as follows from equation 2.17, where the $m = \pm 1$ components contribute with equal magnitude.

As mentioned, ionization from a helium s -subshell allows electrons in final continuum states with $l = 1$ and $m = 0, \pm 1$, where the polarization of the ionizing light governs the magnetic quantum number. z -polarized light gives $m = 0$, whereas x -polarized light produces a linear combination of the $m = \pm 1$ states. Electrons with $l = 1$ are referred to as p-orbitals, and for $m = 0$ the angular probability density represented by two two lobes along the chosen quantization axis, typically taken to be the z -axis. The linear combination of the $l = 1, m = \pm 1$ spherical harmonics have a similar two-lobed shaped probability, but oriented along a different axis [4] [20]. Illustrations of the two cases are presented in figure 2.2.

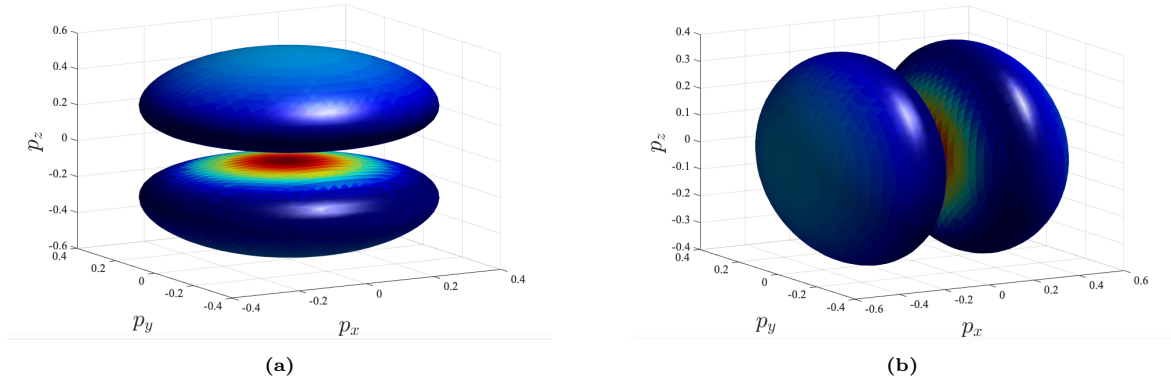


Figure 2.2: Angular probability density for p-orbital electrons. (a): Ionization from x -polarized light, corresponding to $l = 1, m = 0$ spherical harmonics. (b): Ionization from x -polarized light, corresponding to a linear combination of $l = 1, m = \pm 1$ spherical harmonics.

2.2.5 Polarization Parametrization

The polarization state of the driving ionizing field can be defined by the relative amplitudes and phase between the orthogonal components of the electric field. For light propagation along the y -axis, the electric field is confined to the (x, z) -plane with E_x and E_z . As illustrated in figure 2.3, the inclination angle α describes the relative contribution of the two components according to

$$E_x \propto \sin(\alpha), \quad E_z \propto \cos(\alpha) \quad (2.24)$$

Since the absolute phase of the electric field does not affect the physical polarization state, only the relative phase β between the z -component and x -component is considered:

$$E_x = |E_x|e^{i\beta}, \quad E_z = |E_z| \quad (2.25)$$

For $\beta = 0$ or $\beta = \pi$, the two field components oscillate in phase or out of phase, resulting in linear polarization. For $\beta = \pm\pi/2$, the components are phase shifted, resulting in circular polarization with opposite rotation directions.

The relation between the polarization state of the driving field and the spherical harmonic coefficients can be understood from the angular distribution. The presence of only $Y_{1,0}$ corresponds to a purely z -oriented contribution, while a combination of the $m = \pm 1$ spherical harmonics results in a dipole aligned along the x -direction, as seen in equation 2.17. Hence, the coefficients should follow the electric field components in accordance to:

$$c_{1,0} \propto E_z, \quad c_{1,1} \propto -\frac{E_x}{\sqrt{2}}, \quad c_{1,-1} \propto \frac{E_x}{\sqrt{2}} \quad (2.26)$$

Figure 2.4 shows how the spherical harmonic coefficients depend on inclination angle and relative phase for the case of helium ionization from an s -subshell.

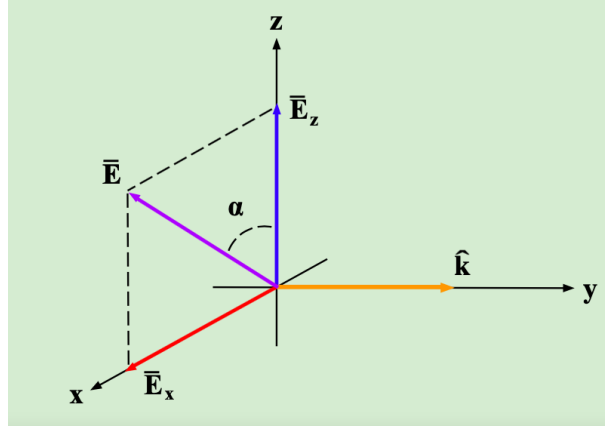


Figure 2.3: An illustration of the inclination angle α , which is given as the angle between the x - and z -components of the electric field. The light propagation direction is indicated by $\hat{k}$.

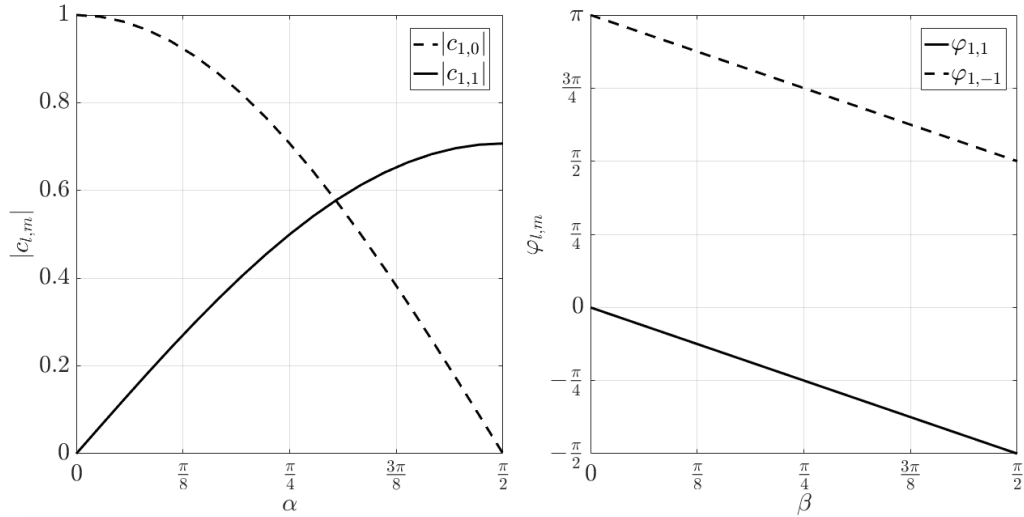


Figure 2.4: Polarization scan for inclination angle $\alpha = [0, \pi/2]$ and $\beta = [0, \pi/2]$, including expected absolute values of amplitude coefficients $c_{1,0}$, $c_{1,1}$ as well as phase coefficients $\varphi_{1,1}$, $\varphi_{1,-1}$.

3 Methodology

The previous chapter introduced the physical background of attosecond light-matter interactions, including the strong-field approximation, high-harmonic generation, and photoionization processes. The relation between the polarization state of the ionizing field and the resulting photoelectron angular distribution was also introduced, showing how the angular distribution can be described using spherical harmonic coefficients that contain information about the polarization state.

This chapter describes the computational framework used to reconstruct such parameters from simulated or experimental photoelectron angular distributions. The generation of simulated data, the modeling of experimental conditions, and the implementation of the different reconstruction methods are presented, as well the evaluation metrics used for comparing the performance of the different methods.

3.1 Objectives of Computational Work

The objectives of the computational work were:

- To take simulated angular distributions as input for the reconstruction methods.
- To retrieve the spherical harmonic decomposition from the angular distribution using equation 2.23.
- To evaluate and compare different methods doing the above, with respect to their robustness and ability to accurately reconstruct the angular distributions and underlying physical coefficients.

3.2 Data Simulation

In order to evaluate the reconstructional methods, data containing photoelectron momentum distributions for known polarization states were generated numerically. A MATLAB implementation was provided to generate ideal simulated data, where the simulation model was based on SFA for ionization governed by attosecond XUV pulses.

Experimental conditions were simulated by adding artificial noise to the generated data. In order to generate noisy data, a new model was constructed in MATLAB. Noise was added directly to the angular probability distributions prior to any decomposition or reconstruction step. The level of noise was controlled in order to evaluate the stability of the reconstruction under non-ideal conditions and identify possible limits for reliable parameter recovery. Both additive Gaussian noise and Poisson noise were considered.

3.2.1 Assumptions and Simplifications

The system was modeled using helium atoms, and the dipole transition was evaluated assuming ionization from an initial s -subshell with quantum numbers $l = 0$, $m = 0$. Consequently, allowed transitions had to satisfy selection rules $\Delta l = \pm 1$, and $\Delta m = 0, \pm 1$, which corresponds to excitation into p -type continuum state with quantum numbers $l = 1$, $m = 0, \pm 1$.

Only XUV pulses were included in the simulations, with no IR driving field. The pulse configuration used in the simulation code consisted of two orthogonal pulses with equal amplitude and no temporal delay, effectively behaving as a single attosecond pulse. This configuration enabled different polarization states to be simulated. Using only single attosecond pulses was sufficient, as the aim of the simulated data was to investigate photoelectron angular distributions rather than the temporal dynamics.

Inclination angle and the relative phase between pulses were varied when generating the simulated data. The simulations were performed on a discrete three-dimensional momentum grid, and only pulses with arbitrary orientation in the (x, z) plane were considered.

3.2.2 Generation of Angular Distribution

The electric field of the two XUV pulses, effectively acting as one pulse, was modeled as a sinusoidal wave with a temporal Gaussian envelope:

$$E_{\text{XUV}}^i(t) = \sqrt{\frac{2I_{\text{XUV}}}{c_0 \epsilon_0}} \cos(\omega_{\text{XUV}}(t - t_{\text{XUV}}(i)) + \beta(i)) \exp\left[-\frac{(t - t_{\text{XUV}}(i))^2}{\sigma^2}\right] \quad (3.1)$$

with the first term being the field amplitude, ω the carrier frequency, σ is related to the pulse duration, $t_{\text{XUV}}(i)$ is the center time of the i th pulse, and $\beta(i)$ the phase of the i th pulse. The total XUV pulse

was given by adding all pulses along with each inclination angle $\alpha(i)$:

$$E_{\text{XUV}}(t) = \sum_{i=1}^2 E_{\text{XUV}}^{\text{pulse}}(t)_i \begin{pmatrix} \cos(\alpha_{\text{XUV}}(i)) \\ 0 \\ \sin(\alpha_{\text{XUV}}(i)) \end{pmatrix} \quad (3.2)$$

Using the Volkov state description in equation 2.8, the expression for the transition amplitude in equation 2.17 and the relation between the transition amplitude and the photoelectron probability distribution in equation 2.21, the photoelectron probability distribution is given by the following relation

$$a(\mathbf{p}) = \int dt \mathbf{E}_{\text{XUV}}(t) \cdot \mathbf{d}(\mathbf{p} + \mathbf{A}(t)) \exp \left[\frac{i}{\hbar} (I_p t - S(t)) \right], \quad P(\mathbf{p}) = |a(\mathbf{p})|^2 \quad (3.3)$$

The parameter values used are presented in table 3.1.

Parameter	Value
Ionization potential	$I_p = 24.6 \text{ eV}$
XUV pulse intensity	$I_{\text{XUV}} = 1 \text{ (relative units)}$
XUV angular frequency	$\omega_{\text{XUV}} \approx 5 \cdot 10^{16} \text{ rad/s}$
XUV pulse duration	$\sigma_{\text{XUV}} = 200 \text{ as}$
Angular grid resolution	$N_\theta \times N_\phi = 80 \times 80$

Table 3.1: Parameters used in the SFA simulation.

The final momentum distribution was transformed into spherical coordinates, in order to convert momentum distributions into angular probability distributions $P(\theta_i, \phi_j)$:

$$\mathbf{p} = (p_x, p_y, p_z), \quad p_x = p \cos \phi \sin \theta, \quad p_y = p \sin \phi \sin \theta, \quad p_z = p \cos \theta \quad (3.4)$$

where $p = |\mathbf{p}|$ denotes the magnitude of the radial momentum vector. As discussed in the background chapter, the angular distribution can be obtained by considering a fixed momentum magnitude, such that only the angular dependence of the momentum distribution will remain. To study the angular structure, the momentum distributions were divided into eight radial momentum intervals, corresponding to ranges in energy as the radial momentum is related to the kinetic energy according to:

$$E_{\text{kin}} = \frac{p^2}{2m} \quad (3.5)$$

Each energy range had a width of approximately 1.5 eV and angular distributions were extracted within each selected range. These energy ranges were treated independently as:

$$P(\theta, \phi) = P(\theta, \phi, E_{i-1} \leq E < E_i), \quad i = 1, \dots, 8 \quad (3.6)$$

Because the signal intensity decreases with increasing energy, the highest intensity and clearest angular distribution is observed for the first energy range. An illustration of the energy dependence of the radial distribution of the emitted electron wavefunction is shown in figure 3.1.

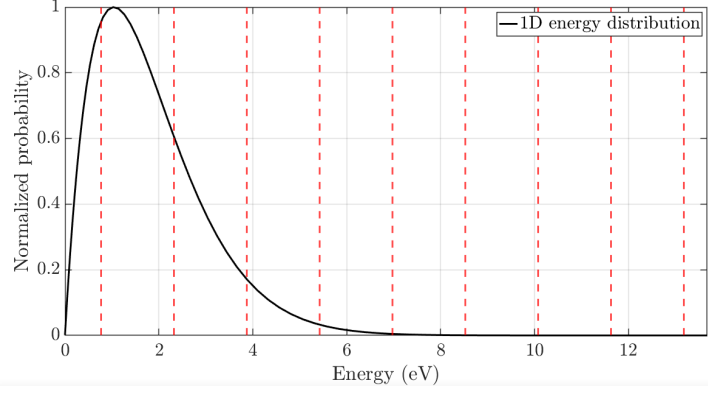


Figure 3.1: Radial distribution shown together with the energy ranges used, indicated by vertical lines. The first range starts at ~ 0.8 eV and ends at ~ 2.3 eV

20 samples were generated by varying the inclination angle from x - to z -polarization, while keeping the relative phase fixed at $\beta = 0^\circ$. 10 samples were generated by varying the relative phase from fully linear polarization to fully circular polarization, keeping the inclination angle fixed at $\alpha = \pi/4$. Additionally, 16 samples were generated by varying the inclination angle around x -polarized light by steps of $\sim 1^\circ$. Figure 3.2 shows four examples of the reconstructed angular distributions for the first four energy ranges, clearly showing the characteristics of different initial polarization and the decline in intensity with energy.

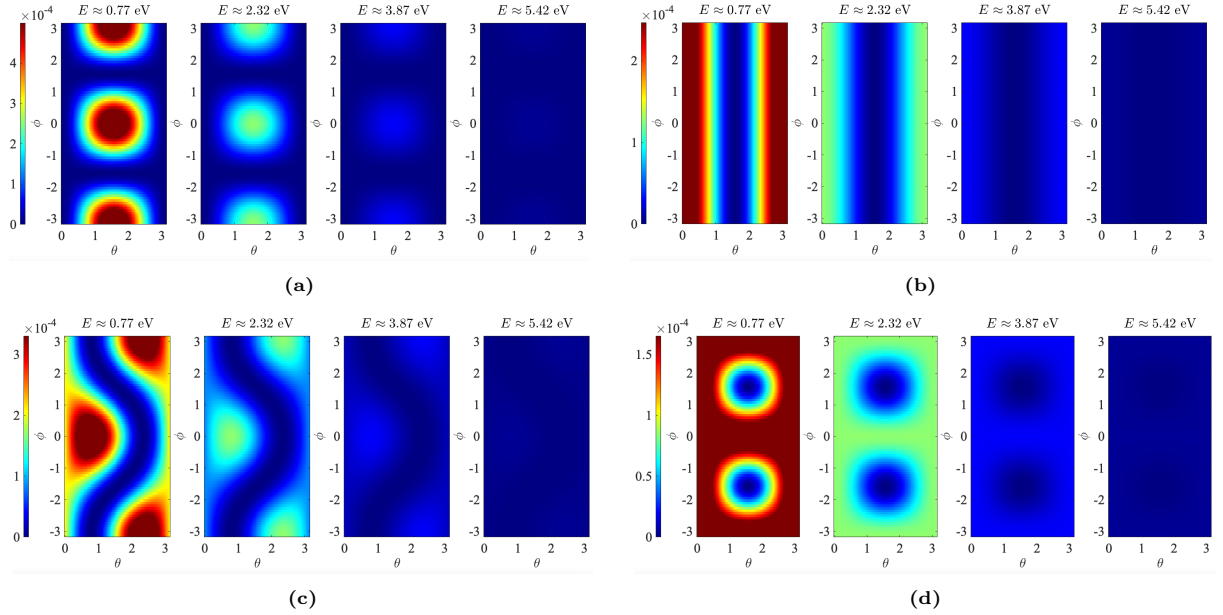


Figure 3.2: Illustration of four different angular distribution generated using the provided SFA simulation. The energy values above distributions correspond to the lower bound of each energy range. (a): x -polarization. (b): z -polarization. (c): 45° linear polarization (d): Circular polarization.

3.2.3 Additive Gaussian Noise

Additive Gaussian noise simulates the impact of random, uncorrelated noise on signals, also called white noise. This type of noise can correspond to electronic detector fluctuations, background noise from measurement electronics or false particle detection [21]. The noise is additive since it is added linearly to the signal, and as a consequence it is independent of the signal intensity. It is referred to as Gaussian because the amplitude follows a normal distribution centered around zero. If the noise is uniformly distributed across all data, it is also referred to as white [22].

The noisy pixel can be represented as the sum of the original pixel value p and a gaussian noise term η drawn from a Gaussian distribution with zero mean value and variance s^2 [21] [22]:

$$x = p + \eta, \quad G(\eta) = \frac{1}{s\sqrt{2\pi}} e^{-\frac{(\eta-\mu)^2}{2s^2}}, \quad \mu = 0 \quad (3.7)$$

The signal-to-noise ratio (SNR), often given in decibels, can be written as

$$\text{SNR} = 10 \log_{10}\left(\frac{x}{p}\right) \quad (3.8)$$

This parameter can be varied systematically to study the noise-tolerance of the reconstruction models, where increasing η corresponds to progressively stronger deviations from the ideal signal [22] [23]. In order to get a broad variety of noise levels, SNR was varied from 40 to -20 decibels using 20 equally spaced values. An example of a distribution with different levels of noise is illustrated in figure 3.3

The built-in MATLAB function `awgn` was used to generate additive Gaussian noise. The original angular distribution and the signal-to-noise ratio in decibels was used as input parameters [23].

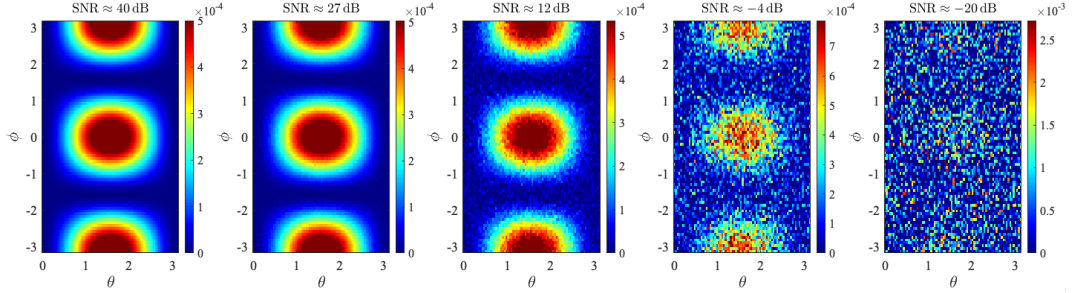


Figure 3.3: Illustration of how the angular distribution is affected by additive Gaussian noise at different noise levels. The angular distribution corresponds to x -polarized light at highest signal intensity.

3.2.4 Poisson Noise

Poisson noise is introduced by resampling the original signal using a Poisson distribution. The noise is therefore correlated to the signal intensity of the original signal, in contrast to additive Gaussian noise. It is also referred to as photon noise since this type of noise naturally arises when the photon counts in detectors are low, for example due to short measurement time, as the discrete nature of photon detection introduces statistical fluctuations in the measured signal [21]. The statistics can be modeled using a Poisson distribution since individual photon detections can be treated as independent random events [24] [21].

The noisy signal is generated by sampling from a Poisson distribution [24]:

$$P(K = k) = \frac{e^{-\lambda t} (\lambda t)^k}{k!} \quad (3.9)$$

t represents the measurement time interval, and the expectation value λ is the expected number of photons per unit time interval, which is proportional to the underlying true signal intensity [24]. Since the simulated noise was generated using photon counts rather than a time-dependent signal, the expected photon count per pixel is used instead, given by $N = \lambda t$ [25] [24].

For a Poisson distribution, the variance is equal to the expectation value, meaning the standard deviation scales with the square root of the signal $\sigma = \sqrt{\lambda}$. Consequently, the signal-to-noise ratio, defined as average over standard deviation, increases as $\sqrt{\lambda t} = \sqrt{N}$ [24] [25]. Hence, by varying the number of photons per pixel, different noise strengths can be generated, which allows for the robustness of the reconstruction methods to be evaluated. In order to get a broad range of noise levels, simulations were

performed using photon counts spanning from 10^8 to 10^2 . Examples of these simulations are presented in figure 3.4.

The MATLAB function `poissrnd` was used to introduce Poisson noise. The input parameter N was calculated using the probability density of the angular distribution multiplied by the photon count per pixel [26]. The result was then rescaled to again represent an angular distribution.

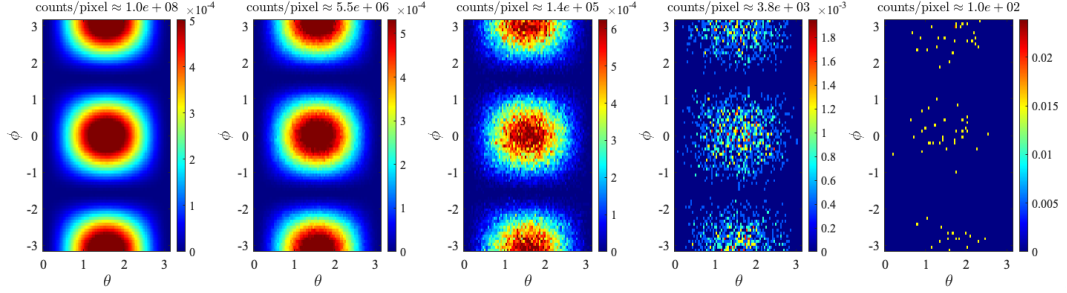


Figure 3.4: Illustration of different levels of Poisson noise added to an angular distribution, corresponding to x -polarized light at highest signal intensity.

3.3 Reconstructional methods

Three different approaches were used in the reconstructional computational work, all implemented using MATLAB. The first method was a direct fitting method based on an established state of the art approach to the problem. The second method was a projective method that had recently been developed. Both of these methods were based on numerical optimization. In contrast, the third method utilized machine learning and was a new approach developed in this work. This third method followed the same principle as the direct fitting method.

For the projective method, an existing MATLAB implementation was provided as a starting point, whereas the direct fitting method and machine learning approach were implemented from scratch.

3.3.1 Spherical Harmonic Decomposition

The angular distributions were represented using coherent sum of spherical harmonics. For single-ionization from helium $1s$ ground state, the resulting final wavefunction can be expressed as a coherent sum of p -shell spherical harmonics according to equation 2.23, where the complex coefficients $c_{l,m}$ contains information about the polarization and relative phase of the ionizing light field, characterizing the angular distribution of the photoelectrons.

3.3.2 Physical Constraints

The amplitudes of the $m = \pm 1$ components were assumed equal $|c_{1,1}| = |c_{1,-1}|$, as given by 2.17, and the relative phase difference between these components was set to a fixed value of π . The implementation of these constraints can therefore reduced the number of parameters to only three:

$$c_{1,0} = a, \quad c_{1,1} = be^{i\varphi}, \quad c_{1,-1} = be^{i(\varphi+\pi)} \quad (3.10)$$

where a and b are amplitudes and φ is the relative phase.

Since only the shape of the angular distribution was of interest, and not the absolute magnitudes of the coefficients, coefficients could additionally be normalized, reducing the number of parameters to two. However, this further reduce of parameters was not implemented in the current reconstruction methods.

The reconstructed probability distribution for these parameters could therefore be written as:

$$P_{\text{model}}(\theta, \phi) = |aY_{1,0} + be^{i\varphi}Y_{1,1} + be^{i(\varphi+\pi)}Y_{1,-1}|^2 \quad (3.11)$$

3.3.3 Numerical Optimization Procedure

Both the direct fitting and projective method use the same optimization procedure to determine the parameters that best reconstruct the angular distribution. The parameters were optimized by minimizing the difference between the input data and the reconstructed values. The minimization was performed numerically using a standard nonlinear constrained optimizer `fmincon` in MATLAB, with the problem formulation [27]:

$$\min_{\mathbf{X}} f(\mathbf{X}), \quad \mathbf{X} = [a, b, \varphi]^T \quad (3.12)$$

During each iteration, the error was calculated and parameters updated accordingly. The iterative procedure continued until one of the convergence criteria was satisfied, so to say the change between iterations became sufficiently small, the parameter updates fell below a the tolerance or the maximum number of iterations was reached [27].

The optimizer `fmincon` requires an initial guess as a starting point for the minimization procedure, as well as upper and lower bounds. Since `fmincon` is a deterministic, gradient-based solver, which means it follows the local gradient to find the closest minimum, it is very sensitive to these parameter choices. A change in any of them can result in completely different solutions due to convergence to different local minima [27].

Furthermore, the optimizer seems to struggle to find solutions that lie on the boundary, probably due to difficulty in finding gradients on both sides of the point. Therefore, in order to allow the optimizer to reach the value zero, the lower bound for the amplitude parameters was chosen slightly below zero. Similarly, the boundaries for the phase parameter were set slightly beyond $-\pi$ and π to avoid restricting the optimizer.

The choices of initial guess and boundaries were determined through trial and error, and give a sufficiently stable and accurate result. All optimization settings used are presented in table 3.2.

Parameter	Value
Maximum number of iterations	1000
Step tolerance	10^{-9}
Upper bound	$[a, b, \varphi] = [0.08, 0.08, 1.4\pi]$
Lower bound	$[a, b, \varphi] = [-0.01, -0.01, -1.4\pi]$
Initial guess	$[a, b, \varphi] = [0.01, 0.01, \pi/2]$

Table 3.2: Optimization parameters used in the calculation.

3.3.4 Direct Fitting Method

An approach investigated was to compare the angular distribution of the input data directly to a reconstructed angular distribution. That is, to fit a coherent sum of spherical harmonics obtained from the reconstruction to the simulated distribution directly, using equation 2.23. In the optimization procedure, the difference was therefore calculated as

$$\text{Diff} = \sum_{\theta, \phi} |(P - P^{\text{model}})| \quad (3.13)$$

$$P = P(\theta, \phi), \quad P^{\text{model}} = |c_{1,0}Y_{1,0}(\theta, \phi) + c_{1,1}Y_{1,1}(\theta, \phi) + c_{1,-1}Y_{1,-1}(\theta, \phi)|^2 \quad (3.14)$$

P_n corresponds to the input angular distribution, whereas P^{model} denotes the angular distribution calculated using the reconstructed coefficients $c_{1,0}$, $c_{1,1}$ and $c_{1,-1}$.

The direct fitting method can be seen as a full image-based comparison where all grid points contribute to the optimization, which for the chosen angular grid was 80×80 points. Therefore, the direct method retains the whole two-dimensional angular structure of the distribution during the optimization.

3.3.5 Projective Method

The direct fitting method compares the reconstructed angular distribution directly with the input distribution, pixel by pixel. With the angular distributions consisting of 80×80 pixels, a very large number of data points need to be evaluated in the optimization. An alternative to the direct fitting method is to utilize the fact that the angular distribution consists of a coherent sum of spherical harmonics, and can therefore be described by a limited number of parameters.

However, evaluating the expression for the angular probability, will result in products of spherical harmonics of the form $Y_{l_1, m_1}^*(\theta, \phi) Y_{l_2, m_2}(\theta, \phi)$. These products do not form an orthogonal basis, which makes a direct projection of the probability distribution onto this set of functions complicated. This problem can be solved by transforming the products into a new orthogonal basis using Gaunt coefficients [28]. The Gaunt coefficients are numerical factors that describe how products of spherical harmonics decompose into a linear spherical harmonic expansion, and are hence of great use for evaluating angular distributions. They provide a structured way to express these products as linear combinations of orthogonal spherical harmonic functions, thereby simplifying the numerical treatment and reducing computational work [28].

The resulting quadratic terms from evaluating the angular distribution can be linearized using Gaunt decomposition, where products of spherical harmonics are rewritten as sums over a new orthogonal basis weighted with Gaunt basis coefficients $k_{l_1, m_1, l_2, l, (m_2 - m_1)}$ [29]:

$$\begin{aligned} (c_{l_1, m_1} Y_{l_1, m_1})^*(\theta, \phi) c_{l_2, m_2} Y_{l_2, m_2}(\theta, \phi) = \\ \sum_{l=l_{\min}}^{l_{\max}} {}^{(2)} \langle l_2 m_2 | l_1 m_1 | l (m_2 - m_1) \rangle (c_{l_1, m_1})^* c_{l_2, m_2} Y_{l, (m_2 - m_1)} = \\ \sum_{l=l_{\min}}^{l_{\max}} {}^{(2)} k_{l_1, m_1, l_2, l, (m_2 - m_1)} Y_{l, (m_2 - m_1)} \end{aligned} \quad (3.15)$$

The superscript (2) indicates that the summation is performed in steps of two [29].

The Gaunt basis coefficients will equal to zero if certain selection rules are not fulfilled. One such selection rule is that the spherical harmonic quantum numbers must be valid, meaning [29] [30]:

$$l_1 \leq |m_1|, \quad l_2 \leq |m_2|, \quad l \leq |m_2 - m_1| \quad (3.16)$$

Furthermore, the summation limits in equation 3.15 is given by the triangle relation [30]:

$$|l_1 - l_2| \leq l \leq l_1 + l_2 \quad (3.17)$$

For ionisation from an s -subshell the angular quantum numbers are $l_1 = l_2 = 1$, and there is hence only contribution from basis functions with $l = 0$ and $l = 2$. The resulting basis therefore consists of the spherical harmonics

$$[Y_{0,0}, Y_{2,-2}, Y_{2,-1}, Y_{2,0}, Y_{2,1}, Y_{2,2}] \quad (3.18)$$

This projective method uses Gaunt decomposition before optimization to reduce the number of parameters. First, the angular distribution is projected onto the Gaunt basis given by equation 3.18, resulting

in a set of Gaunt basis coefficients:

$$\begin{aligned}
P(\theta, \phi) &= |c_{1,0}Y_{1,0} + c_{1,1}Y_{1,1} + c_{1,-1}Y_{1,-1}|^2 = \\
&Y_{0,0}(k_{10,10,00} + 2k_{11,11,00} + k_{1-1,1-1,00}) + \\
&Y_{2,0}(k_{10,10,20} + k_{11,11,20} + k_{1-1,1-1,20}) + \\
&Y_{2,1}(k_{11,10,21} + k_{10,1-1,21}) + \\
&Y_{2,-1}(k_{1-1,10,2-1} + k_{10,11,2-1}) + \\
&Y_{2,2}(k_{1-1,11,22}) + \\
&Y_{2,-2}(k_{11,1-1,2-2}) + \\
&= k_{00}Y_{0,0} + k_{20}Y_{2,0} + k_{21}Y_{2,1} + k_{2,-1}Y_{2,-1} + k_{22}Y_{2,2} + k_{2,-2}Y_{2,-2}
\end{aligned} \tag{3.19}$$

These resulting Gaunt basis coefficients are stored as:

$$k_n = [k_{00}, k_{20}, k_{21}, k_{2,-1}, k_{22}, k_{2,-2}] \tag{3.20}$$

During optimization, the spherical harmonic coefficients $c_{l,m}$ are varied to generate the modeled coefficients k_n^{model} . These coefficients are compared against the Gaunt basis coefficients obtained from the projection of the input angular distribution. The aim is therefore to find the set of $c_{l,m}$ that best reproduces the projected Gaunt basis coefficients by minimizing the difference:

$$\text{Diff} = \sum_{n=1}^6 |\text{Re}(k_n - k_n^{\text{model}})| + \sum_{n=1}^6 |\text{Im}(k_n - k_n^{\text{model}})| \tag{3.21}$$

An illustration of the projective method, showing the projection, comparison, and optimization steps used to reconstruct the angular distribution, can be found in figure 3.5.

The provided MATLAB implementation of the projective method was optimized with respect to both computational efficiency and accuracy. By precalculating the Gaunt coefficients in the optimization step, instead of evaluating them at every iteration, the computational time was reduced from minutes to seconds, allowing a large number of different angular distributions to be analyzed efficiently. Also, by tuning the boundary conditions, the initial guess, and minimization parameters, the performance was somewhat improved compared to the original implementation.

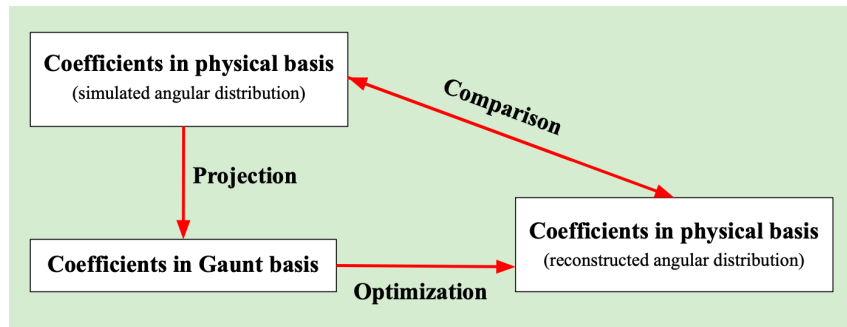


Figure 3.5: A schematic illustration of the projective method.

3.3.6 Machine Learning Approach

In addition to the direct fitting and projection-based reconstruction methods, a machine learning approach was tried. The core idea was for the network to learn mapping between angular distributions and the underlying physical coefficients describing the polarization state of the incident field. A convolutional neural network (CNN) based encoder with Adam optimizer was used to extract features from the angular distribution images. Since these networks are designed to capture spatial relationships between inputs and learn features characterizing underlying structures in the data, they are well suited for this task [31].

A fixed physics-based decoder was then used to reconstruct the underlying physical coefficients describing the polarization state of the incident light field.

Approximately $N \approx 5000$ samples were used, with each sample corresponding to an angular distribution:

$$P(\theta, \phi) \in R^{N_\theta \times N_\phi}, \quad N_\theta \times N_\phi = 80 \times 80 \quad (3.22)$$

where the energy range of each angular distribution, given by equation 3.6, was treated as an independent sample.

The dataset included distributions with varying polarization, intensities and noise levels in order to expose the network to a broad range of possible scenarios during training. The dataset was normalized in the same manner as for the other methods, but also additionally rescaled by a factor of 1000, as the magnitude was otherwise found to be too small for the network to learn during training. Only the first four energy ranges were considered, since initial tests showed poor performance for angular distributions with lower intensity, particularly for noisy distributions. Training a network on data that cannot be reliably reconstructed would not be meaningful. Also, noise was included during training to improve the robustness and generalization performance for realistic experimental data. 20 different SNR values between 40 and -10 decibels was used to generate data with Gaussian noise, and 20 different photon count values in the range of $[10^8, 10^3]$ were used to generate data with Poisson noise. The choice of these noise ranges was again motivated by the results of the initial testing.

The dataset was divided into both a training and validation set, using the so called hold out method for dataset splitting. 30% of the data were used for validation, leaving 70 % for training [31]. The training set was used for the network to learn relevant features from the angular distributions during training, whereas the validation set was used to estimate the generalization performance, that is, the performance on unseen data. A validation evaluation was performed after each training epoch, which allowed the tuning of hyperparameters and network architecture to improve performance [32]. In addition, a test set was also used for the final evaluation of the trained model, containing all relevant cases needed for comparison with the other reconstruction methods.

A typical CNN layer consists of three layers, namely a convolution, activation, and pooling. In the convolution step, learnable filters, known as kernels, are applied to the input to extract features in spatial patterns. These filters are trained using backpropagation and shared across the input. Padding can also be applied, which conserves the image size. In the activation step, a nonlinear activation function is applied to introduce nonlinearity. In order to avoid sampling every pixel for every layer in the network, a stride can be used to reduce for data reduction. This is typically introduced along with pooling [31]. In the pooling step, the spatial resolution is reduced in order to decrease computational complexity, while still preserving the all important features in the input [33]. After all layers are applied, the network is typically flattened into a one-dimensional vector using a fully connected layer, which produces a final output representation of the input that combines all features used to classify the image [31] [34]. The architecture used in the CNN is presented in table 3.3 [34]:

Input layer	$80 \times 80 \times 1$ angular distribution image
Convolutional layer 1	Convolution: 8 filters, 3×3 kernel, padding Activation: ReLU Max pooling: 2×2 (stride 2)
Convolutional layer 2	Convolution: 16 filters, 3×3 kernel, padding Activation: ReLU Max pooling: 2×2 (stride 2)
Convolutional layer 3	Convolution: 32 filters, 3×3 kernel, padding Activation: ReLU Max pooling: 2×2 (stride 2)
Convolutional layer 4	Convolution: 64 filters, 3×3 kernel, padding Activation: ReLU
Fully connected layers	64 neurons, ReLU 32 neurons, ReLU
Output layer	4 parameters: $[a, b, \cos \varphi, \sin \varphi]$

Table 3.3: The architecture used in the CNN.

To encoded the input in RGB format, three channels were used for each kernel, and padding was implemented so that spatial dimensions were reserved [34]. The rectified linear unit (ReLU) was used as activation function since it is the standard choice for CNNs [31]. Max pooling was applied to reduce the dimensions, meaning the highest value in each region was selected [34]. The last fully connected layer produces the physical coefficients representing the angular distribution. Two parameters were used to represent the phase φ , $\cos(\phi)$ and $\sin(\phi)$, due to the networks inability to handle complex outputs. This representation also avoids possible discontinuities arising from the periodic nature of the phase, and limit the values to $[0, 2\pi]$. The optimization parameters for the network were therefore:

$$[a, b, \cos(\phi), \sin(\phi)] \quad (3.23)$$

The network was trained using mini-batch gradient descent along with an Adam optimizer [35]. This optimizer is a popular and robust expansion to the stochastic gradient descent optimization algorithm [31] [33]. Additionally, the hyperparameters presented in table 3.4 were also used.

Parameter	Value
Number of epochs	50
Mini-batch size	64
Learning rate	$5e - 4$

Table 3.4: Hyperparameters used for model optimization.

In order to enforce physical constraints on the reconstructed parameters, a fixed physics-based decoder was used to reconstruct the angular distribution from the predicted coefficients. The network was trained using a reconstruction loss defined as the mean squared error between the reconstructed and target angular distributions:

$$\text{Loss} = \frac{1}{N_{\text{batch}} N_{\theta, \phi}} \sum_{n=1}^{N_{\text{batch}}} \sum_{\theta, \phi}^{N_{\theta, \phi}} (P_{\text{model}}^n(\theta, \phi) - P_{\text{target}}^n(\theta, \phi))^2 \quad (3.24)$$

P_{model} was calculated using the relation in equation 3.11. Since the network could not handle complex-valued parameters, P_{model} was evaluated using only the real value of the optimized amplitude parameters

$$c_{1,0} = \text{real}(a), \quad c_{1,1} = \text{real}(b)e^{i\phi}, \quad c_{1,-1} = \text{real}(b)e^{i(\phi-\pi)}, \quad \phi = \text{atan2}(\sin(\phi), \cos(\phi)) \quad (3.25)$$

where atan2 is the four-quadrant inverse tangent function that retrieves the phase angle from its sine and cosine components, while preserving the correct sign [36]. For the test and validation performance, the relation given by equations 3.10 could be used for P_{model} .

3.4 Evaluation Metrics

Several quantitative metrics were used to evaluate and compare the performance of the different reconstruction approaches. The main focus were on comparing how well each method fitted the data and comparison of retrieved spherical harmonic coefficients.

3.4.1 Reconstruction Error

The primary metric used to compare the performance in terms of robustness to noise and variations in intensity was the coefficient of determination, denoted R^2 . The coefficient of determination is a statistical measure that quantifies how well the reconstructed distributions reproduce the original data. It is defined as

$$R^2 = 1 - \frac{SSE}{SST} \quad (3.26)$$

where SSE is the error/regression sum of squares and SST is the total sum of squares [37] [38].

The error sum of squares measures the deviation between the original and reconstructed values,

$$SSE = \sum_i (y_i - \hat{y}_i)^2 \quad (3.27)$$

where $\hat{y}_i$ denotes the reconstructed value. The error sum of squares can be interpreted as measure of how much variation in the reconstructed value cannot be explained by a linear relationship [37].

The total sum of squares measures the total variation present in the observed data and is given by

$$SST = \sum_i (y_i - \bar{y})^2 \quad (3.28)$$

where $\bar{y}$ is the mean value of the original input data. In contrast to the error sum of squares, the same number is subtracted from each data point instead of using corresponding the fitted value [37].

The coefficient of determination has a upper-bound of one, but is not lower-bounded. The interpretation of the coefficient of determination can be summarized according to table 3.5 [38] [39].

R^2 value	Interpretation
$R^2 = 1$	Perfect agreement between reconstructed and original distribution, hence the reconstruction error is zero $SSE = 0$
$R^2 = 0$	The model performs no better than predicting the mean value of the original data for all points, corresponding to $SSE/SST = 1$
$R^2 < 0$	The reconstruction is worse than a constant mean-value model, which occurs when the reconstruction error exceeds the variance of the original data $SSE/SST > 1$

Table 3.5: Interpretation of the coefficient of determination R^2 .

In the case where $SST = 0$, meaning all values are equal to their mean value, the coefficient of determination R^2 becomes ill-defined due to the lack of variance in the data [40]. Therefore, it is also useful to consider the sum of squared errors SSE . This metric is lower-bounded by zero but has no finite upper bound, and its absolute value is only meaningful in comparison between models evaluated on the same dataset and under the same conditions.

Since R^2 is scale independent, it provides a convenient metric for comparing different fitting approaches and datasets with varying intensity scales [39]. It is mostly commended for linear models, but can be used for models that are intrinsically nonlinear and fitted with nonlinear methods, as well as for linear models fitted using methods other than regression [40].

3.4.2 Coefficient Comparison

To quantitatively evaluate the reconstruction of spherical harmonic coefficients, the extracted coefficients were compared against the known expected values for specific polarizations presented in figure 2.4.

A comparison of the reconstructed parameters can reveal possible degeneracies in the methods used, that is, if the same angular distribution can be reconstructed for different sets of coefficient values. Since $c_{1,\pm 1}$ and $c_{1,1}$ were assumed to be equal, it is enough to investigate coefficients $c_{1,0}$ and $c_{1,1}$.

4 Results

In this chapter, the results obtained from the evaluation of the reconstruction methods are presented. The main aspects considered are dependence on signal intensity, robustness with respect to noise, and comparison of the reconstructed coefficients. The results are interpreted and discussed, and the limitations of each reconstruction method are addressed.

4.1 Dependence on Signal Intensity

Figure 4.1 presents the results for the three reconstructional methods, showing the average R^2 value over 30 simulated angular distribution as a function of energy, along with the variation in signal intensity for increasing energy. No noise have been added to the simulated distributions. The y-axis was set to show $R^2 \in [-1.5, 1]$, and the variance is included as a shaded region around the curves. The same result is also presented in table 4.1. It is apparent that all three methods produce similarly good results at high signal intensity, although the direct fitting method appears much more robust at lower signal levels in comparison to the projective method and the machine learning method. The projective method shows the largest decrease in performance with decreasing intensity, and exhibits the highest variance. The performance of the machine learning method also decreases at the same energy as the projective method, although it maintains a higher R^2 and lower variance, indicating a slightly more stable reconstruction performance.

In figure 4.2, the average result over all samples of the projective method is plotted along with the gaunt basis projection, again for $R^2 \in [-1.5, 1]$. The full reconstruction refers to the final result obtained after optimization, whereas the Gaunt basis projection represents only the projection of the original input distribution onto the Gaunt basis before the optimization step. The gaunt basis projection seems to accurately describe the angular distribution for all signal intensities, with very small variance, indicating that the projection is not the reason for the degradation in performance with increasing energies for the projective method.

To investigate whether the decline in performance at decreasing signal intensity for the projective method is related to the magnitude of the coefficients, a quick test was done where the coefficients used in the optimization were multiplied by different factors before the optimization step. As shown in figure 4.6, there is a considerable improvement in performance at lower signal intensities, where the coefficients have smaller magnitudes, when a rescaling factor is used. The best performance at the lowest intensities is obtained with a factor of 1000. This behaviour is not unexpected, especially since the machine learning method also required the data to be rescaled by a factor of 1000 for the network to learn effectively. However, the performance decreases at higher signal intensities when a large rescaling factor is used, suggesting that excessive rescaling does not improve the overall performance. In other words, rescaling when the coefficient magnitudes are already sufficiently large will have a negative effect on the performance.

Figure 4.3 shows the corresponding results for different polarization states with $R^2 \in [0, 1]$, and figure 4.4 illustrates respective reconstructed angular distribution for the lowest energy ranges. For the direct fitting method, completely x - and z -polarized light, as well as 45° -polarized light appears to give highly accurate results across all energies, whereas circularly polarized light performs noticeably worse as the energy is increased. The projective method appears to perform the worse for fully x - and 45° -polarized

light, with the fastest decline in performance for the latter case. The machine learning method shows relatively similar performance across all polarization states, with its best performance observed for 45°-polarized light. However, its performance decreases more rapidly than the other two methods as the intensity is reduced for the other cases. It appears that the average performance of the machine learning method is higher than for the projective method, while its performance is generally lower for individual cases. This indicates that the machine learning method gives a more consistent reconstruction, whereas the projective method appears to be more dependent on the polarization of the light, performing very well for some polarizations but significantly worse for others.

It is also noteworthy that the apparent default solution of the projective method resembles the distribution obtained for circular polarization. It is therefore difficult to determine whether the good performance in the circularly polarized case reflects accurate reconstruction, or if it is simply due to inherent bias of the reconstruction method. This behaviour is expected, since the initial guess for the phase is the same as for circularly polarized light. It is therefore reasonable that the method converges towards the initial value if no other solution is found. As figure 4.5 shows, changing the initial guess consequently results in the other default solutions. This also explains why the performance of the projective method is dependent on the polarization state. In contrast to the two other reconstruction methods, no initial guess is provided for the machine learning method. However, the model still appears to converge towards the same default solution for all cases, although this solution differs from those obtained with the other two methods. The reason for this behavior is unclear, but could possibly be related to the composition of the training data, where the model may have learned to reproduce the most common patterns present in the dataset.

Energy (eV)	Direct fitting		Projective		Machine learning	
	μ	σ	μ	σ	μ	σ
0.7750	1.0000	0.0000	0.9997	0.0006	0.9994	0.0017
2.3249	1.0000	0.0000	0.9957	0.0078	0.9987	0.0030
3.8748	0.9999	0.0003	0.9643	0.1312	0.9936	0.0124
5.4547	1.0000	0.0000	0.9725	0.0924	0.9654	0.0364
6.9747	0.9998	0.0008	-0.1436	1.0403	0.4874	0.4262
8.5246	0.9047	0.5030	-0.1088	1.0422	0.3384	0.5328
10.0745	0.8587	0.4201	-0.1057	1.0405	0.3194	0.5458
11.6244	0.8511	0.4315	-0.3385	0.8989	0.3129	0.5466

Table 4.1: Comparison of average values μ and standard deviation values σ for the three different reconstructional methods using eight different energy ranges. The energy values given in the table correspond to the lower bound of each energy range.

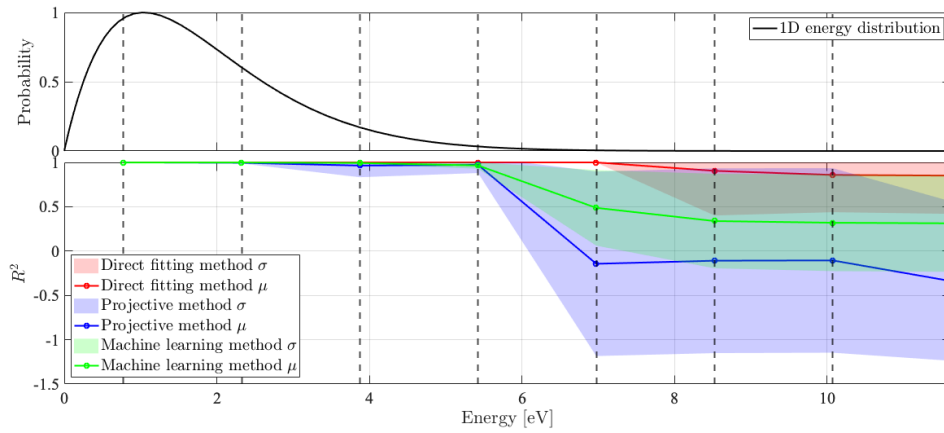


Figure 4.1: Comparison of the average R^2 values for the three reconstructional methods with regards to signal intensity.

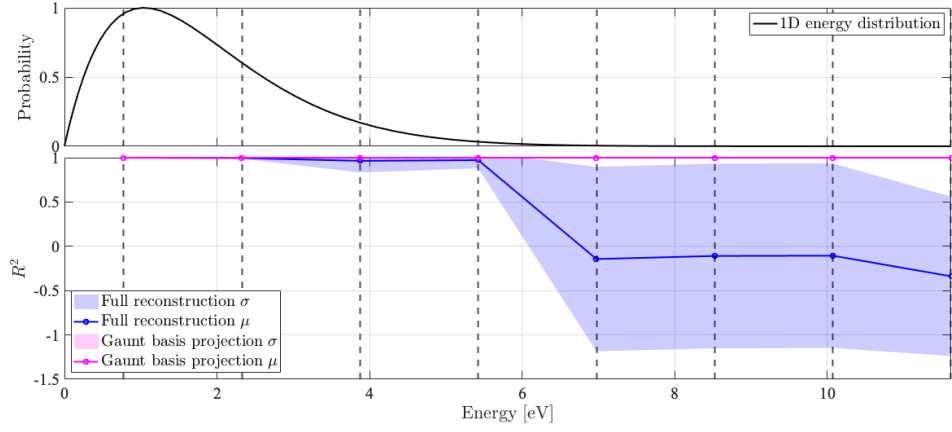


Figure 4.2: Comparison of the average R^2 values of the full reconstruction and the Gaunt basis projection for the projective method, as a function of signal intensity.

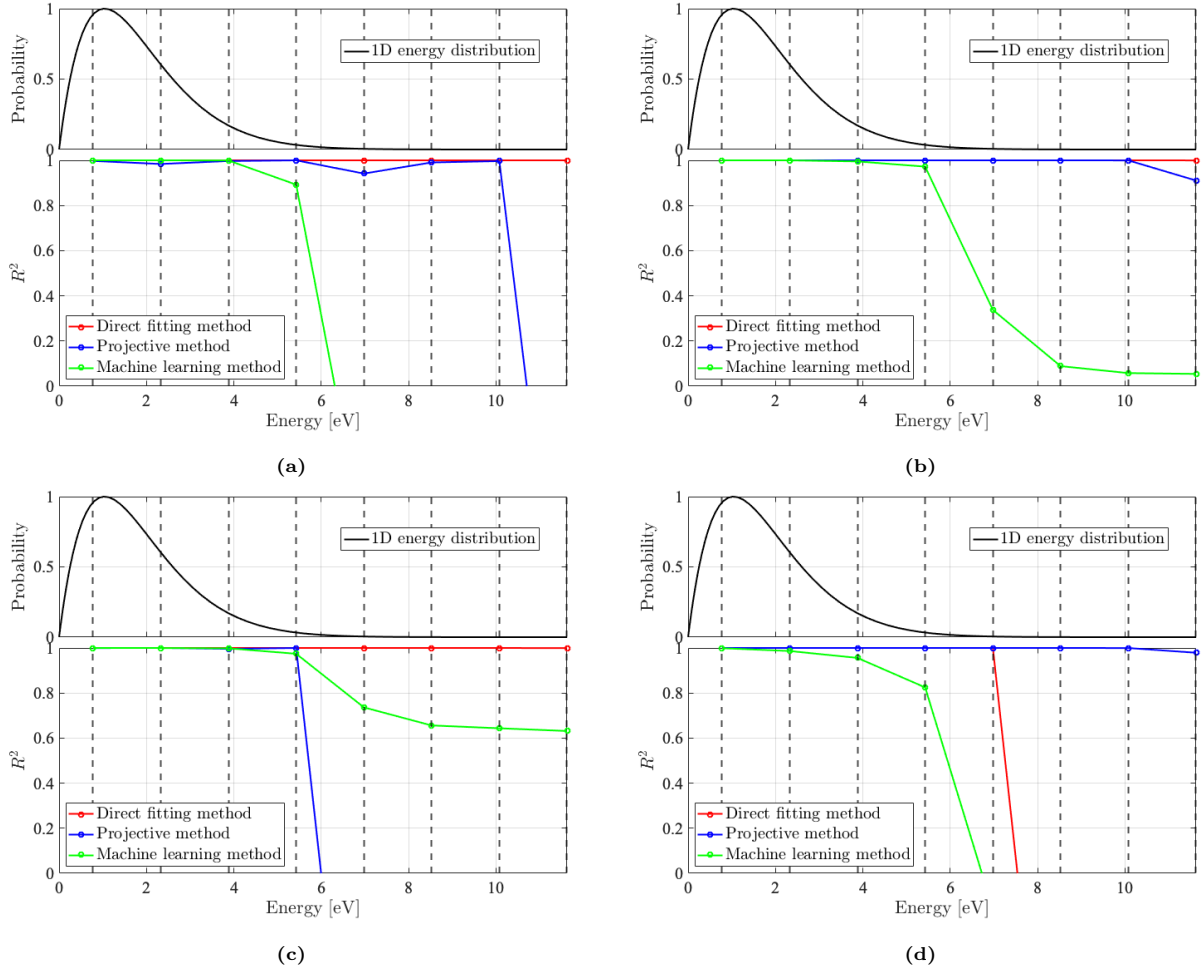


Figure 4.3: The R^2 values for four different angular distributions under varying signal intensity (a): x -polarization. (b): z -polarization. (c): 45° linear polarization (d): Circular polarization.

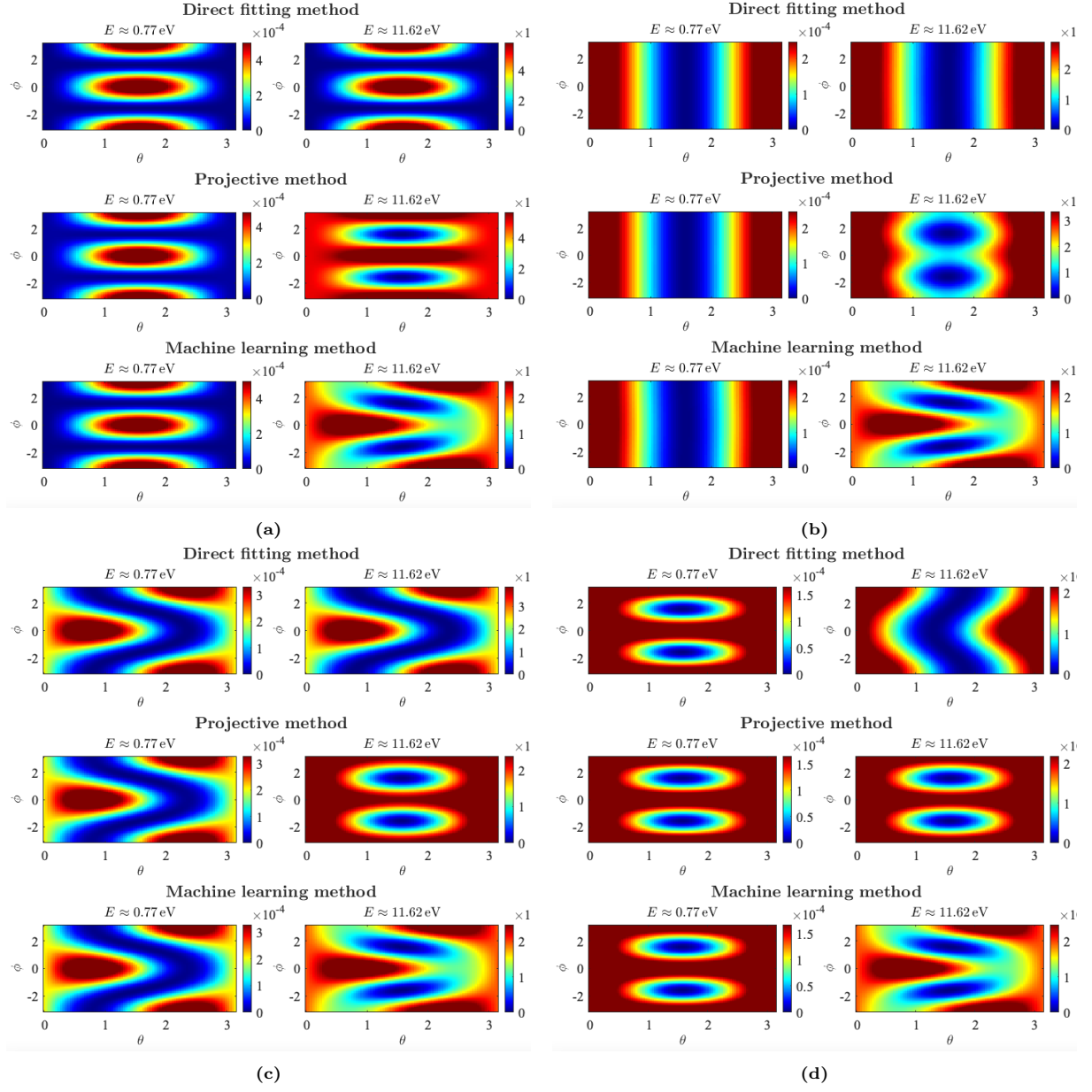


Figure 4.4: Reconstructed angular distributions for four different angular distributions. Only the first and last energy ranges are shown. (a): x -polarization. (b): z -polarization. (c): 45° linear polarization (d): Circular polarization.

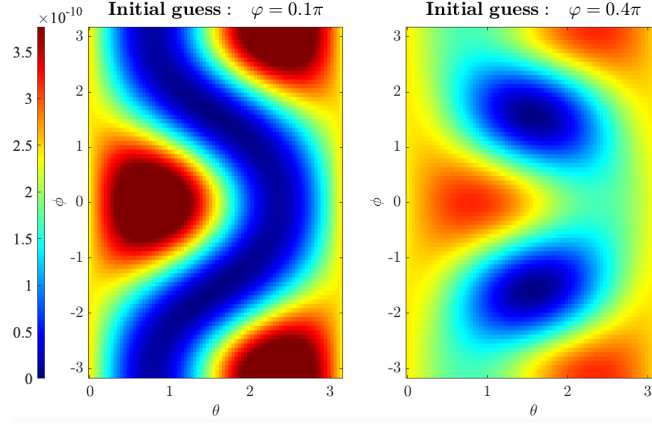


Figure 4.5: Comparison of reconstructed angular distributions for the highest energy range using the projective method with different initial guesses for the phase parameter.

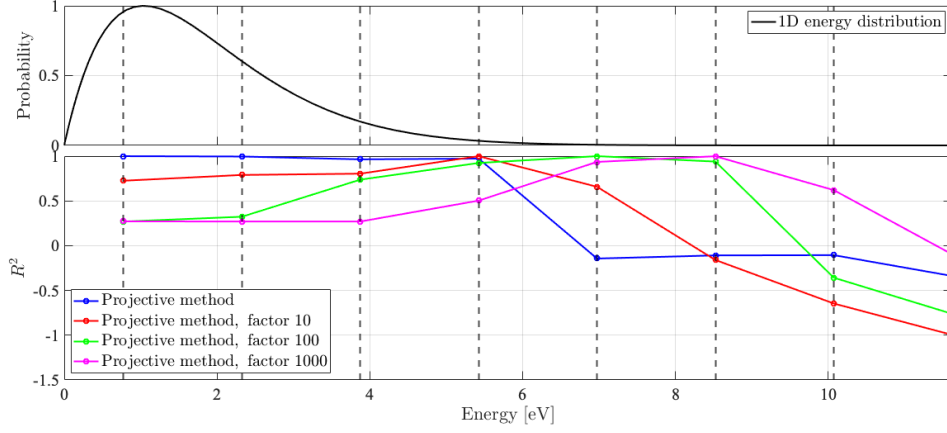


Figure 4.6: Results for the projective method using different rescaling factors.

4.2 Noise Robustness Analysis

In figure 4.7, the average R^2 performance over 30 samples for both additive Gaussian noise and Poisson noise are shown for all three reconstruction methods, with the y-axis limited to $R^2 \in [0.5, 1]$. Only the first energy range, corresponding to highest signal intensity, is considered. The variance is shown by the shaded areas. Respective average and variance values for each method are also presented in table 4.2 and table 4.3.

All three methods perform similarly well for Gaussian noise levels up to an SNR of approximately 0 dB. The performance thereafter decreases for all methods. The projective method maintains the highest performance, followed by the direct fitting method and then the machine learning method. The variance remains very small for the projective and direct fitting methods when the reconstruction performance is high, whereas the machine learning method shows a larger variance at lower noise levels as well as generally in comparison to the other two methods. For Poisson noise, all methods again perform well at low noise levels, with a decrease in performance occurring at approximately $1e4$ counts/pixel. Compared to Gaussian noise, the direct fitting method shows a significantly larger and faster decrease in performance. Again, the projective method achieves the highest performance, while the machine learning method follows a similar trend but with a slightly steeper decrease. Similar to the case of Gaussian noise, the machine learning method shows a higher variance compared to the other two methods at lower noise levels.

Figure 4.8 compares the gaunt basis projection in the projective method against the full reconstruction for both additive Gaussian noise and Poisson noise. Again, the plot only represents the result for the

first energy range and only $R^2 \in [0.5, 1]$ are shown. The Gaunt basis projection shows significantly lower performance compared to the full reconstruction for decreasing signal intensity. This behavior is expected, since the Gaunt basis projection aims to describe the input distribution as accurate as possible, and will therefore give a low performance when compared to the ground truth for noisy input. In contrast, during the optimization step, variations not describing the underlying structure of the distribution are disregarded since noise should not be reconstructed. Nevertheless, this still raises the question of how such a high reconstruction accuracy can be achieved with a noisy basis.

Figure 4.9 and 4.10 shows the results for different polarization states at the highest signal intensity, for additive Gaussian noise and Poisson noise respectively. For both figures, the y-axis shows $R^2 \in [0.5, 1]$. For Gaussian noise in figure 4.9, circular polarization consistently appears to give the poorest reconstruction performance for all methods, and the machine learning method also performs particularly poor for z -polarization. Apart from this observation, there is no strong indication of dependence on polarization state for any of the methods. For Poisson noise in figure 4.10, the direct fitting method performs similarly across all polarization states. As observed for Gaussian noise, circular polarization gives the poorest performance for the projective and machine learning methods. These two methods also show unexpectedly large variations in performance, particularly for fully x -polarized and circularly polarized light, where the performance is, in some cases, better at higher noise levels than at lower ones.

Gaussian noise SNR (dB)	Projective method		Direct fitting method		Machine learning method	
	μ	σ	μ	σ	μ	σ
~ 40	0.9995	0.0010	1.0000	0.0000	0.9994	0.0017
~ 37	0.9967	0.0159	1.0000	0.0000	0.9994	0.0016
~ 34	0.9996	0.0008	1.0000	0.0000	0.9994	0.0017
~ 31	0.9995	0.0010	1.0000	0.0000	0.9994	0.0018
~ 27	0.9995	0.0010	1.0000	0.0000	0.9994	0.0018
~ 24	0.9973	0.0114	1.0000	0.0000	0.9994	0.0017
~ 21	0.9995	0.0011	1.0000	0.0000	0.9994	0.0020
~ 18	0.9995	0.0010	1.0000	0.0000	0.9993	0.0018
~ 16	0.9994	0.0010	1.0000	0.0000	0.9991	0.0026
~ 12	0.9995	0.0010	0.9999	0.0001	0.9992	0.0025
~ 8	0.9994	0.0010	0.9998	0.0001	0.9990	0.0027
~ 5	0.9994	0.0009	0.9996	0.0002	0.9990	0.0029
~ 2	0.9992	0.0009	0.9992	0.0005	0.9987	0.0033
~ -1	0.9975	0.0070	0.9985	0.0010	0.9970	0.0057
~ -4	0.9980	0.0022	0.9969	0.0020	0.9948	0.0136
~ -7	0.9962	0.0044	0.9938	0.0042	0.9951	0.0067
~ -11	0.9920	0.0093	0.9872	0.0083	0.9891	0.0212
~ -14	0.9838	0.9742	0.0161	0.9569	0.0934	0.0192
~ -17	0.9659	0.0403	0.9487	0.0295	0.8847	0.2785
~ -20	0.9303	0.0732	0.8982	0.0534	0.7233	0.5253

Table 4.2: Average values and standard deviations of R^2 for the three reconstruction methods for different levels of additive Gaussian noise.

Poisson noise Counts/pixel	Projective method		Direct fitting method		Machine learning method	
	μ	σ	μ	σ	μ	σ
$\sim 1.0e8$	0.9995	0.0010	1.0000	0.0000	0.9994	0.0018
$\sim 4.8e7$	0.9995	0.0010	1.0000	0.0000	0.9994	0.0017
$\sim 2.3e7$	0.9996	0.0009	1.0000	0.0000	0.9993	0.0019
$\sim 1.1e7$	0.9973	0.0114	1.0000	0.0000	0.9993	0.0021
$\sim 5.5e6$	0.9994	0.0011	1.0000	0.0000	0.9993	0.0019
$\sim 2.6e6$	0.9981	0.0076	1.0000	0.0000	0.9992	0.0023
$\sim 1.3e6$	0.9995	0.0009	1.0000	0.0000	0.9990	0.0032
$\sim 6.1e5$	0.9995	0.0011	1.0000	0.0001	0.9991	0.0025
$\sim 3.0e5$	0.9994	0.0010	0.9999	0.0001	0.9988	0.0034
$\sim 1.4e5$	0.9976	0.0078	0.9999	0.0002	0.9977	0.0063
$\sim 6.9e4$	0.9994	0.0010	0.9998	0.0004	0.9980	0.0056
$\sim 3.4e4$	0.9992	0.0011	0.9996	0.0006	0.9968	0.0052
$\sim 1.6e4$	0.9990	0.0012	0.9945	0.0166	0.9974	0.0058
$\sim 7.9e3$	0.9972	0.0043	0.9588	0.1014	0.9948	0.0099
$\sim 3.8e3$	0.9952	0.0138	-0.6880	1.3863	0.9913	0.0160
$\sim 1.8e3$	0.9957	0.0040	-1.1531	1.0819	0.9893	0.0149
$\sim 8.9e2$	0.9922	0.0099	-1.0552	0.9811	0.9429	0.1243
$\sim 4.3e2$	0.9697	0.0436	-1.2061	0.9306	0.8930	0.2347
$\sim 2.0e2$	0.8992	0.1328	-1.3056	0.8790	0.8939	0.1773
$\sim 1.0e2$	0.8860	0.2220	-1.2792	0.9779	0.7688	0.5389

Table 4.3: Average values and standard deviations of R^2 for the three reconstruction methods for different levels of Poisson noise.

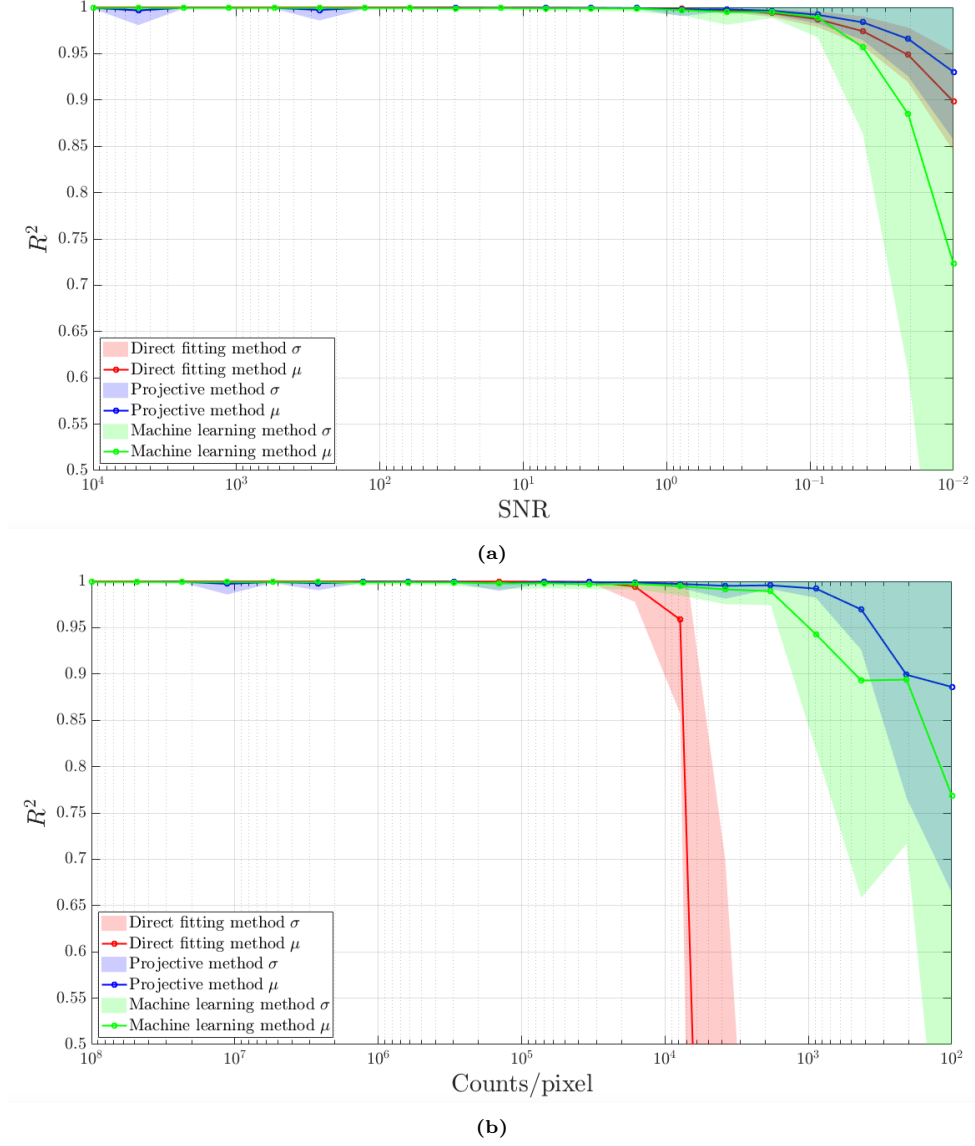


Figure 4.7: Comparison of the average R^2 values for the three reconstructional methods at different levels of noise (a): additive Gaussian noise. (b): Poisson noise.

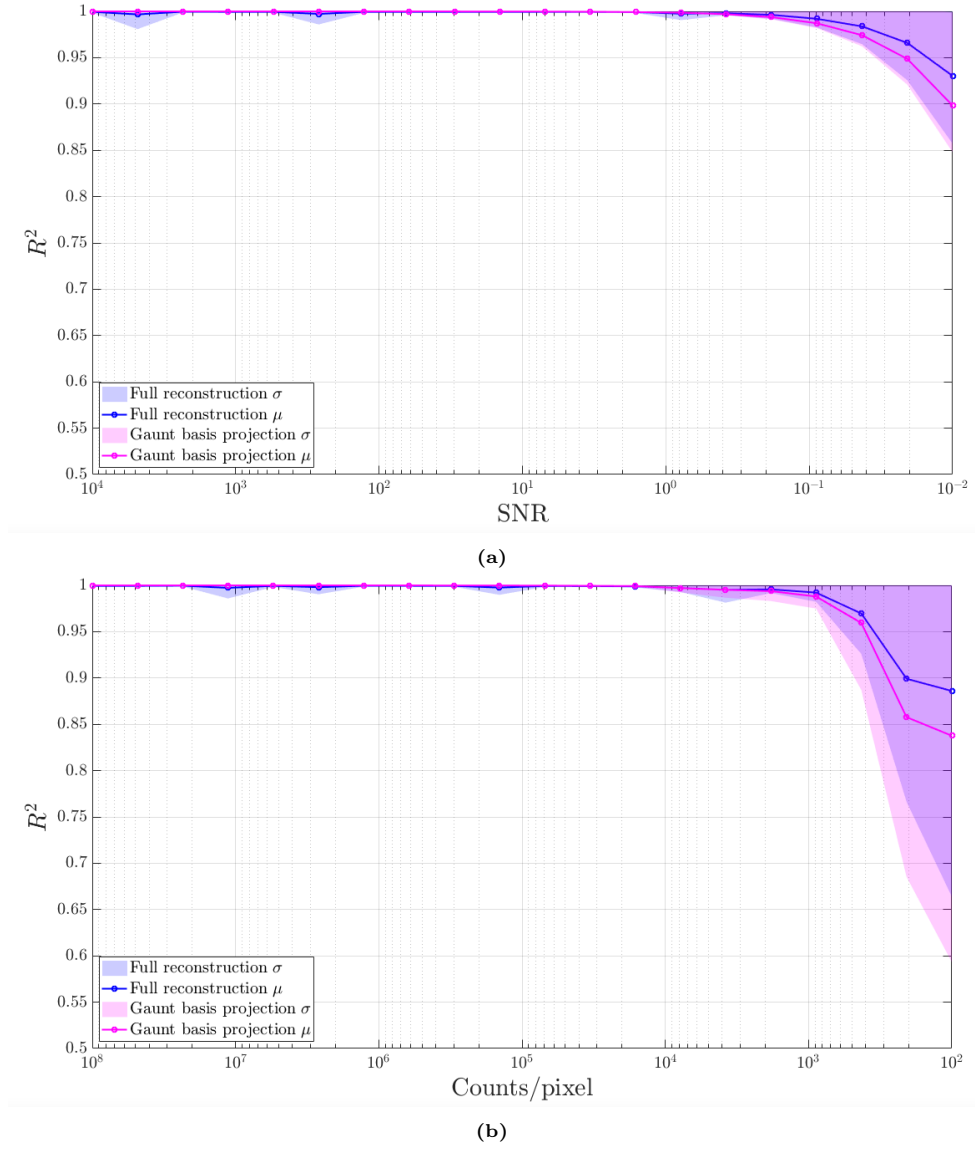


Figure 4.8: Comparison of the average R^2 values of the full reconstruction and the Gaunt basis projection for the projective method. (a): additive Gaussian noise. (b): Poisson noise.

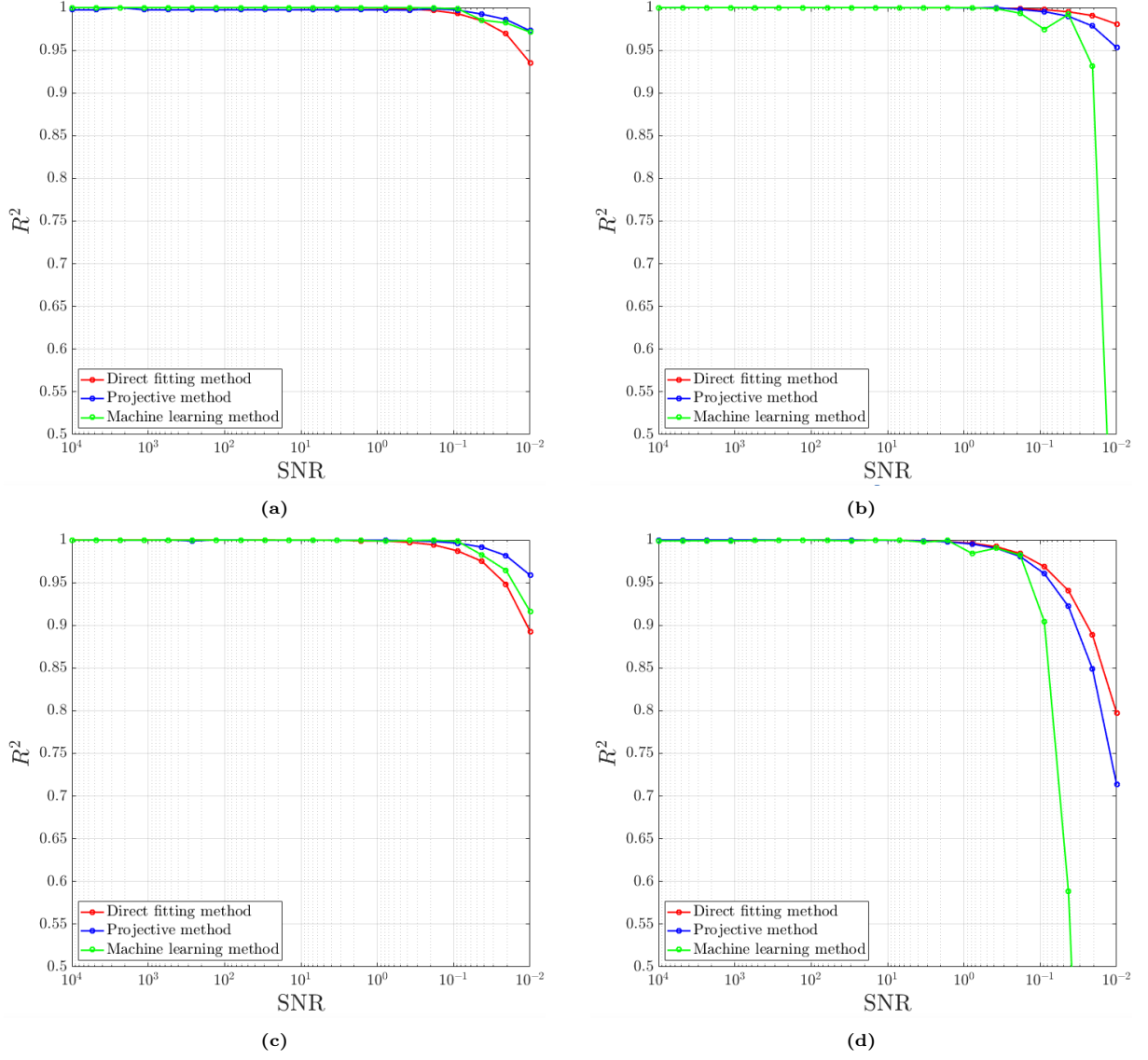


Figure 4.9: The result for four different angular distributions with varying level of additive Gaussian noise (a): x -polarization. (b): z -polarization. (c): 45° linear polarization (d): Circular polarization.

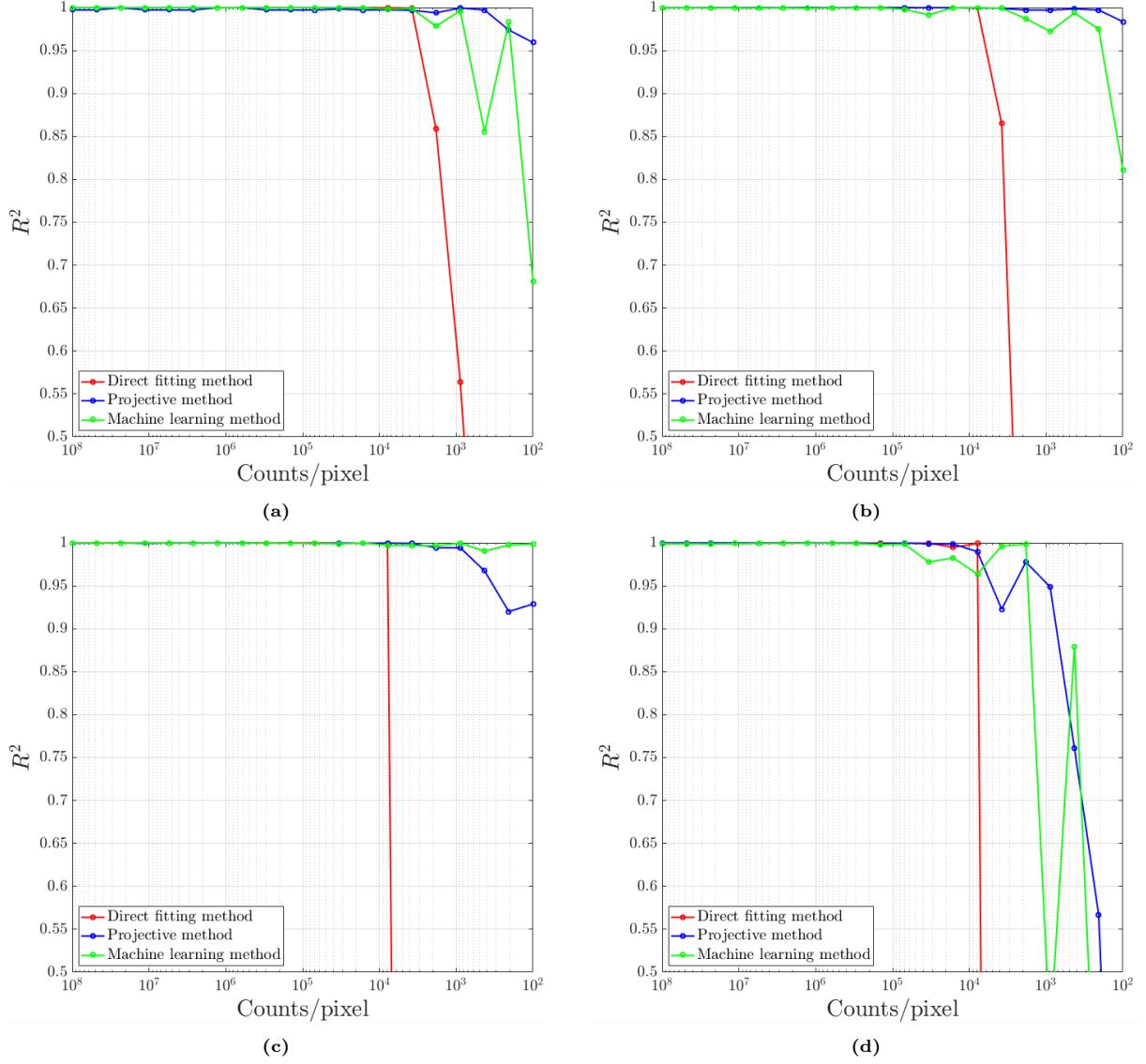


Figure 4.10: The result for four different angular distributions with varying level of Poisson noise (a): x -polarization. (b): z -polarization. (c): 45° linear polarization (d): Circular polarization.

4.3 Comparison Reconstructed Coefficients

Figure 4.11 shows the polarization scan plots along with normalized reconstructed coefficients for each of the reconstructual method, as a function of the inclination angle α , whereas figure 4.12 illustrates the same results for varying relative phase β . A zoomed-in version of figure 4.11(a), using 16 samples centered around x -polarization and only including the amplitude coefficients, is shown in figure 4.13. Figure 4.14 illustrates the normalized reconstructed coefficients for each of the reconstructual method under Gaussian noise with $SNR \approx 1\text{dB}$, and figure 4.15 shows the corresponding reconstructed coefficients under Poisson noise with approximately $6.9 \cdot 10^4$ counts per pixel.

For varying inclination angle, shown in figure 4.11, the largest deviation from the expected amplitudes is found for the $c_{1,0}$ coefficient in the case of fully x -polarized light, where the projective method appears to have the greatest discrepancy. One possible reason why all three methods give values greater than zero in this case is that they all have difficulty converging exactly to zero, and instead return a small non-zero value for coefficients that are expected to vanish. All three methods also give phase values lower than expected, which could be related to the increased amplitude values. This behaviour is also

seen for the corresponding case of fully z -polarized light, where the $c_{1,1}$ coefficient should vanish, though not as apparent. The phases $\varphi_{1,\pm 1}$ show similar behaviour and remain mainly constant, with the largest deviations occurring for fully x - and z -polarized light. This is expected, since the corresponding amplitude coefficients become very small as the polarization approaches a pure state, making the phase harder to determine and causing the result to remain close to the initial guess.

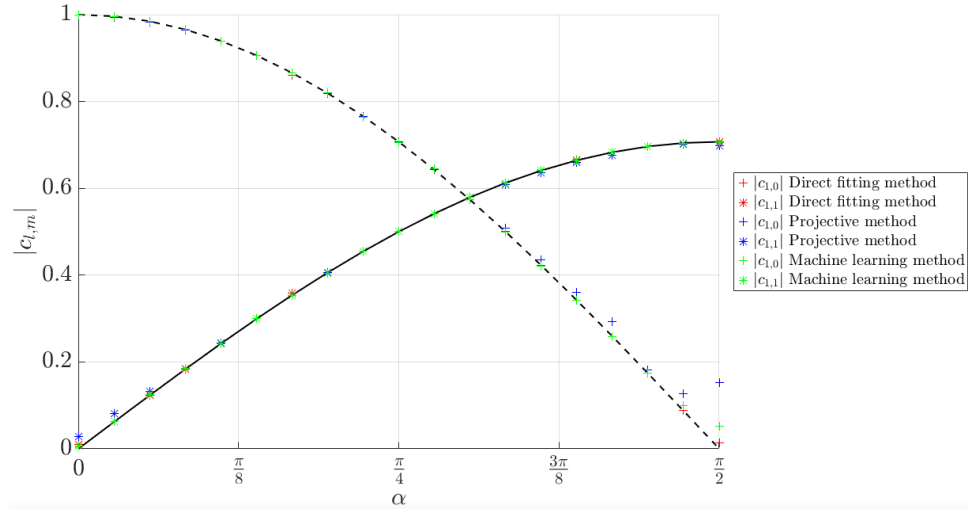
The projective method shows the largest variations, followed by the machine learning method. It should be mentioned that the machine learning phase values shown have been plotted using their absolute values, instead of the actual negative values reconstructed by the method. The reason for these negative phase values could possibly be related to the use of the `atan2` function, which returns the phase in the correct quadrant. However, these differences do not necessarily indicate degeneracy, where different physical solutions are found, since phases differing by 2π are equivalent.

Examining the results for varying relative phase, shown in figure 4.12, no large deviations can be observed. All methods show some difficulty in accurately resolving the phase for 45° -polarization, with both the projective method and machine learning method showing somewhat larger phase fluctuations compared to the direct fitting method. Again, this could possibly be explained by the difficulty of the reconstruction methods to assign exactly zero as a value. Similar to figure 4.11, the plotted phase values $\varphi_{1,\pm 1}$ given by the machine learning method are actually the absolute values of the reconstructed phase values.

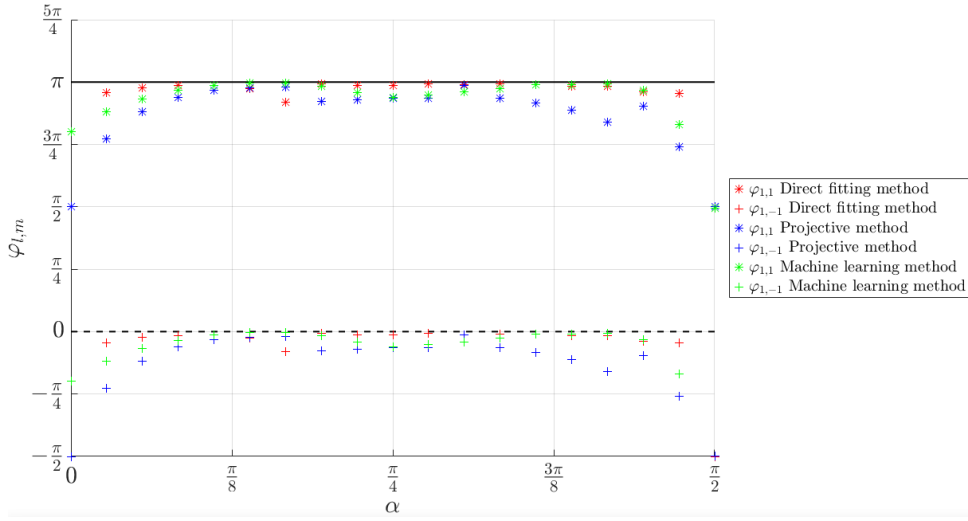
The zoomed-in version of figure 4.11(a), presented in figure 4.13, shows that all methods perform similarly well for the $c_{1,1}$ coefficient. The projective method shows the largest deviation from the expected value of $c_{1,0}$, as previously seen in figure 4.11. Interestingly, for four of the 16 samples, the projective method reconstructs the coefficient much more accurately. It appears that even very small changes in α can result in large differences in reconstruction accuracy.

In figure 4.14 showing the coefficient values with added Gaussian noise, the largest deviations from the expected values are observed for the machine learning method, with is not surprising as it is consistent with the previously presented noise robustness analysis. For most polarization states, the machine learning method tends to underestimate the value of $c_{1,0}$, while overestimating $c_{1,1}$ as well as the phases φ_{11} and φ_{1-1} . An exception is the fully x -polarized state, where $c_{1,0}$ is overestimated and both φ_{11} and φ_{1-1} are underestimated.

For the Poisson noise case, presented in figure 4.15, the machine learning method shows a similar behavior and larger deviations can also be observed for the direct method, which is again consistent with the noise robustness analysis. The direct method generally follows trends similar to the machine learning method. Some changes can also be observed for the projective method, but in contrast to the other two methods, it overestimates $c_{1,0}$ for polarization states between fully x -polarized and 45° polarization, and underestimating the phase values in some cases. In general, the largest deviations are observed in the phase values for β values corresponding to elliptically polarized light. The phase values reconstructed using the direct method show particularly large deviations, and become essentially meaningless in some cases. This may help explain the poor performance of the direct method for high Poisson noise levels.

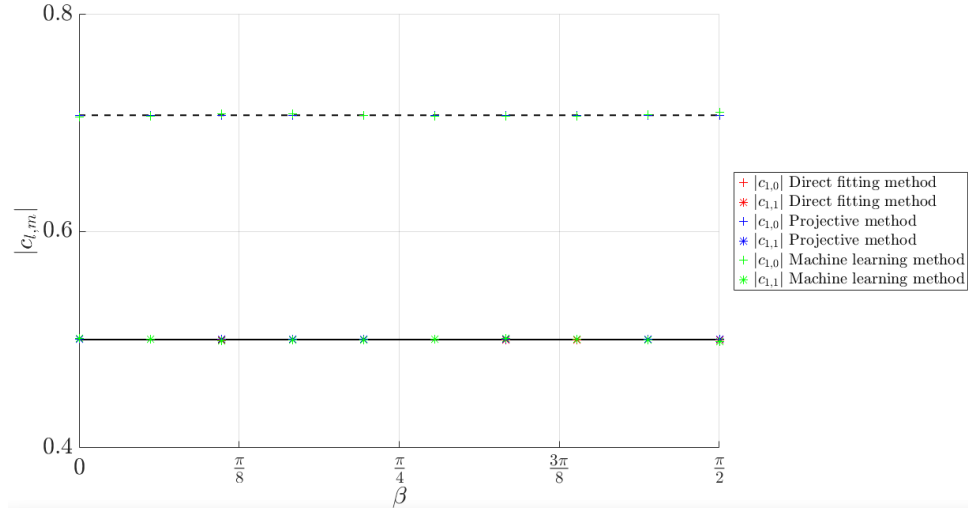


(a)

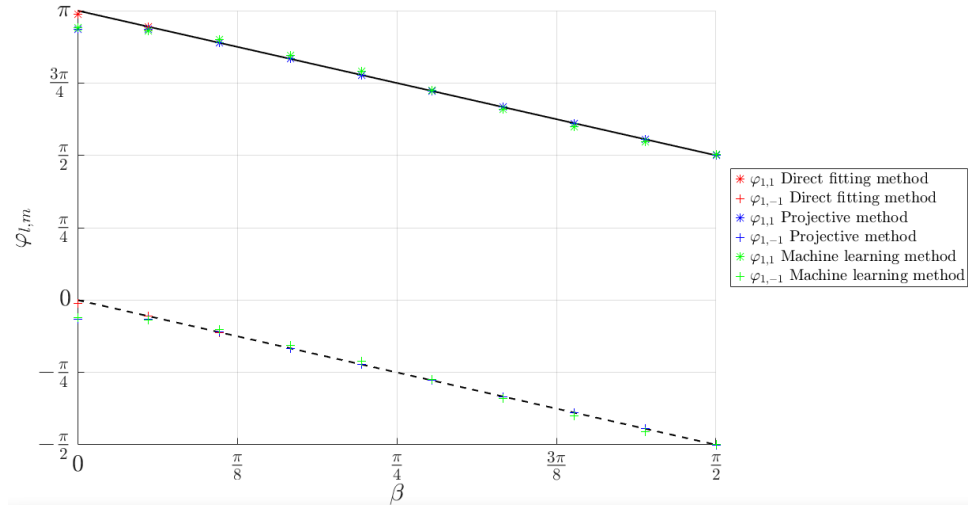


(b)

Figure 4.11: Comparison of reconstructed coefficients against expected values, varying the inclination angle $\alpha \in [0, \pi/2]$: (a) the amplitude coefficients $c_{1,0}$ and $c_{1,1}$ (b) the phase coefficients $\varphi_{1,1}$ and $\varphi_{1,-1}$



(a)



(b)

Figure 4.12: Comparison of reconstructed coefficients against expected values, varying the relative phase $\beta \in [0, \pi/2]$: (a) the amplitude coefficients $c_{1,0}$ and $c_{1,1}$ (b) the phase coefficients $\varphi_{1,1}$ and $\varphi_{1,-1}$

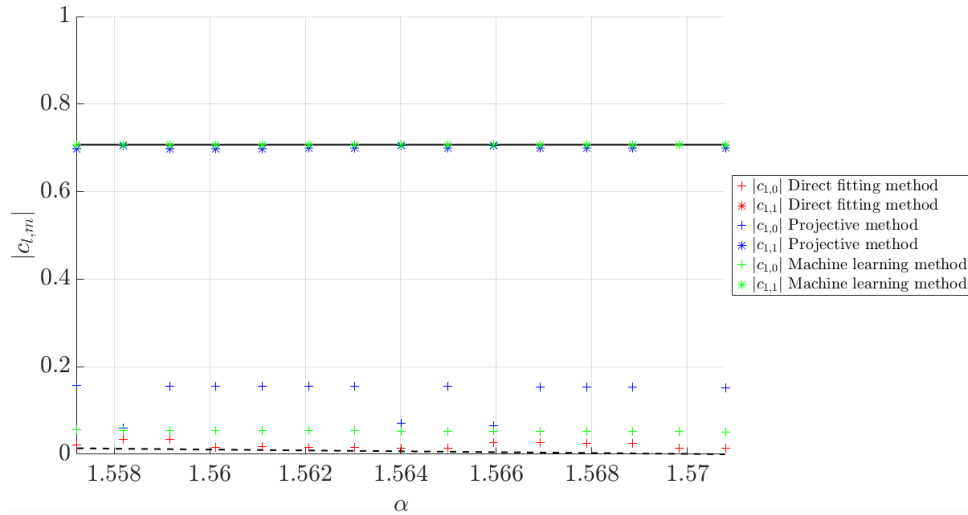


Figure 4.13: Comparison of reconstructed amplitude coefficients against expected values, varying the inclination angle $\alpha \in [\sim \frac{89\pi}{180}, \frac{\pi}{2}]$.

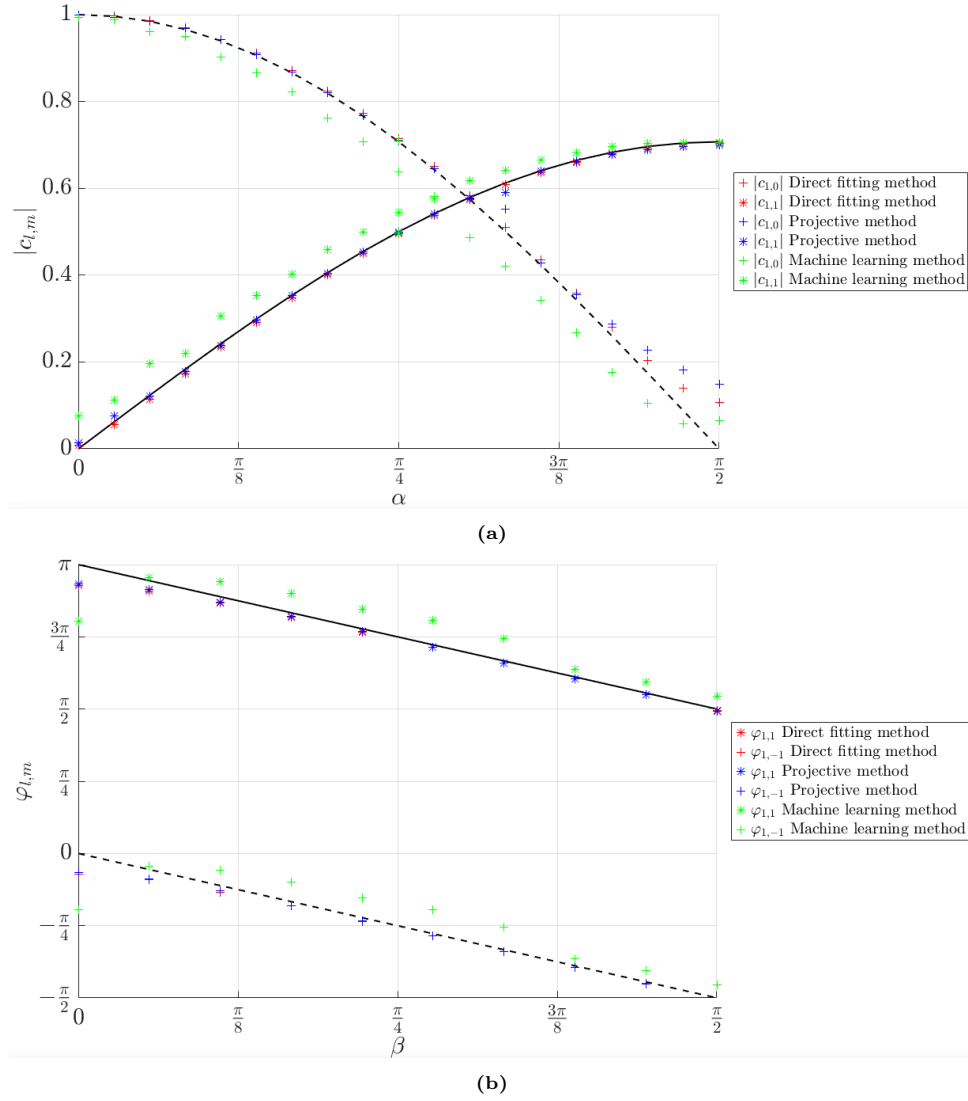


Figure 4.14: Comparison of reconstructed coefficients against expected values using additive Gaussian noise with SNR $\approx -1\text{dB}$. (a) the amplitude coefficients $c_{1,0}$ and $c_{1,1}$ for varying inclination angle α (b) the phase coefficients $\varphi_{1,1}$ and $\varphi_{1,-1}$ for varying relative phase β

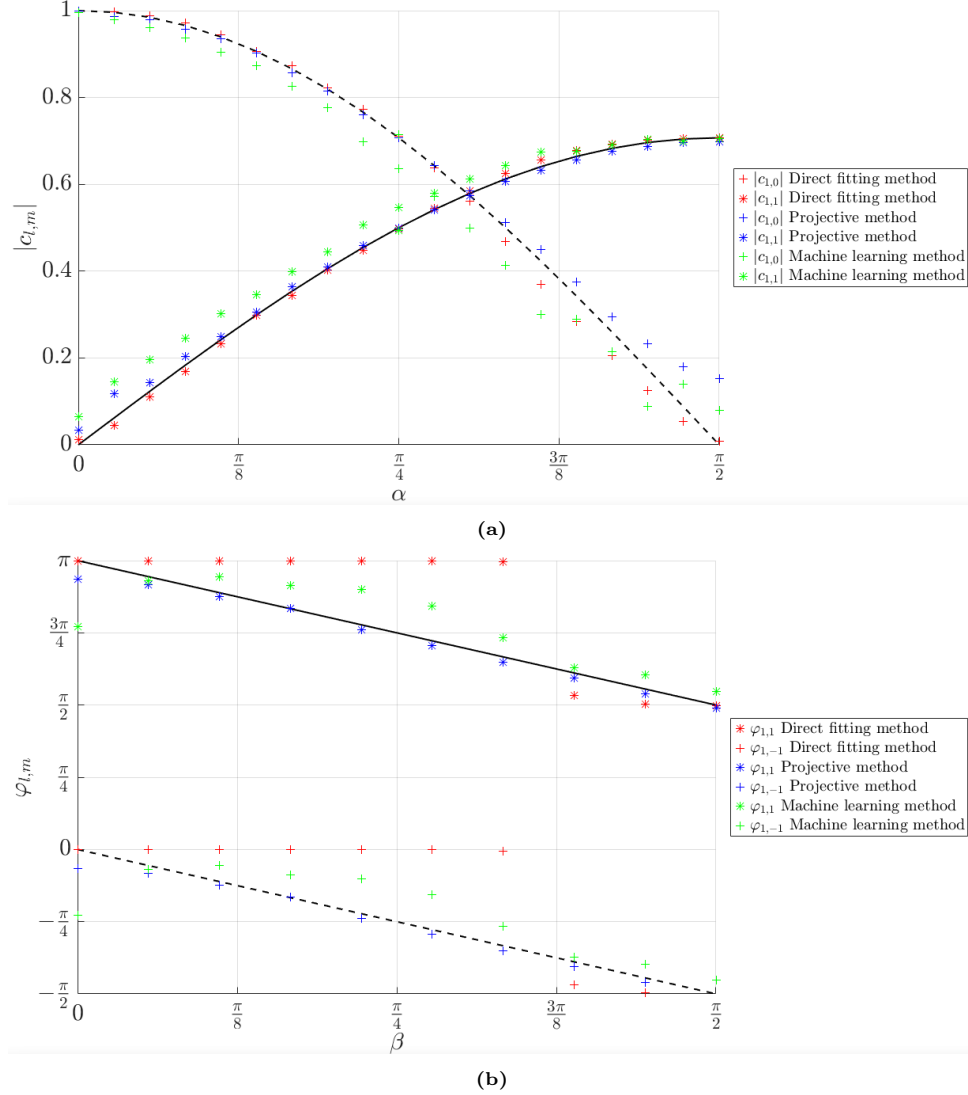


Figure 4.15: Comparison of reconstructed coefficients against expected values using Poisson noise with counts/pixel $\approx 6.9 \cdot 10^4$ (a) the amplitude coefficients $c_{1,0}$ and $c_{1,1}$ for varying inclination angle α (b) the phase coefficients $\varphi_{1,1}$ and $\varphi_{1,-1}$ for varying relative phase β

4.4 Limitations in Reconstructual Methods

There are several limitations for each of the three reconstruction methods. For the direct fitting and projective method, the reconstructed parameters are obtained by using a minimization function very much dependent on both initial guess and boundary conditions. Although the dependence on boundary conditions is not directly apparent in the results presented, variations in these settings during the early stages of this work were found to affect the reconstruction performance to a great extent. These dependencies contribute to highly varying performance between different polarization states, with significantly better reconstruction accuracy for some states than for others. Ideally, a reconstruction method should provide consistent performance for different polarization states, rather than strongly depending on the input data.

For the projective method, the reconstruction accuracy does not appear to be limited by the accuracy of the Gaunt basis projection itself, but rather by the magnitude of the coefficients used for optimization. However, any errors introduced in projection stage will affect the reconstructed coefficients and therefore limit overall performance.

In contrast to the two other methods, the machine learning method does not require an initial guess or boundary conditions for optimizing the reconstructed parameters. However, since the model is trained on a predefined dataset, the choice and diversity of the training data strongly influence the performance. The neural network may perform better for distributions that are similar to those included during training, while its performance might decrease for cases outside the training distribution, such as lower intensities or higher noise levels. Furthermore, since no physical constraints are imposed on the reconstructed parameters, such as parameter boundaries, the model might converge towards solutions that do not correspond to physically meaningful parameters.

A challenge observed for all three methods is the difficulty of accurately reconstruct the physical coefficients describing polarization states where one coefficient is supposed to vanish. None of the methods seems able to converge close to zero, and instead to return small non-zero values. Also, the finite resolution and sampling of the angular distributions might cause fundamental limitations in the reconstruction accuracy for all three reconstruction methods, as well as not having used extra decimals in MATLAB calculations. Another limitation is the inherent redundancy for special cases, such as left and right circular polarization states, which produce identical angular distributions. Therefore, these polarization states cannot be distinguished when using only their angular distributions.

5 Conclusions

This chapter summarizes the previously presented results and discusses possible improvements to the current reconstruction methods. Future work is also investigated, and final conclusions regarding the project are presented.

5.1 Summary of Results

All reconstruction methods achieve accurate results for high-intensity distributions. As the signal intensity is reduced, the direct fitting method maintains the highest performance, indicating a lower sensitivity to small coefficient magnitudes. In comparison, both the projective and machine learning method show a stronger decrease in performance at lower intensities, indicating a higher sensitivity. It is worth noting that higher performance may be achieved with the projective method for lower intensities by rescaling before optimization. Both the numerical optimization-based methods show strong dependence on selected initial guess and parameter boundaries, while the machine learning method appears less dependent on both initial guess and specific polarization state.

For noisy data, the projective method generally gives the most robust performance, especially for Poisson noise. However, all methods maintain high performance over a relatively large noise range, especially for Gaussian noise. The direct fitting method shows a more significant decrease in performance than the other methods, while the machine learning method follows a similar trend as the projective method, but with larger variance. The difficulty of the direct fitting method in adapting to distributions that deviate from the expected structure is most likely due to optimization being performed pixel by pixel, which makes it more sensitive to local deviations. The projective method on the other hand considers the overall structure of the distribution through the basis projection, making it more tolerant towards small variations. Furthermore, the machine learning method gives a more consistent performance for different polarizations, whereas the numerical optimization-based methods show stronger variations depending on the polarization state.

Regarding the reconstructed coefficient values, all methods provide results close to the expected values when no noise is added, with the largest deviations observed in the phase parameters for varying inclination angle. All methods show difficulties in accurately reconstructing coefficients that are expected to be zero, instead returning small non-zero values.

Overall, the results demonstrate that the coefficients obtained from decomposing a photoelectron momentum distributions into a coherent sum of spherical harmonics provide sufficient information to characterize the ionizing light field. Furthermore, several different reconstruction approaches can be successfully

used to retrieve these spherical harmonic coefficients, although their performance differs with regard to factors such as signal intensity and noise level. The machine learning method also shows potential to match the performance of the other methods if further improved, for example through larger training dataset or further optimized network architecture.

5.2 Improvements and Future Work

There are several alterations that could improve the overall performance of the reconstructional methods, and hence should be considered in future related works. Also, the results of this work could be used as guidance for experimental work, and physical scope of the methods could be extended.

5.2.1 Improvement of Current Methods

Applying the reconstruction methods to real experimental data would be a natural next step in order to obtain a better understanding of their performance under realistic conditions. While simulated data and artificial noise provide a controlled environment for evaluating the methods, they cannot fully represent all additional effects and uncertainties that may be present in experimental data. Alternatively, simulated with both Gaussian and Poisson noise at the same time could be considered. In a real experiment, both background noise and a limited number of detected electrons are present, so including both types of noise would provide more realistic testing of the methods. The study could also be extended to more complex angular distributions, such as those given by multiple ionizing pulses with different time delays, examining the resulting interference patterns.

Another possible improvement is to reduce the number of parameters to optimize by parameterizing the coefficients using only the inclination angle α and the relative phase φ . By normalizing the spherical-harmonic coefficients, the two amplitude parameters a and b can be expressed in terms of α

$$|c_{1,0}| = \cos(\alpha), \quad |c_{1,1}| = |c_{1,-1}| = \sin(\alpha) \quad (5.1)$$

where equation 2.24 and 2.26 have been used. Using these two parameters instead of the current three would reduce the dimensionality of the optimization problem, and could possibly improve the performance.

For the numerical optimization-based methods, the choice of initial guesses and parameter boundaries strongly affects the reconstruction results, since changes can lead the optimization algorithm to converge to a different local minimum. One possible improvement would therefore be to explore other minimization functions that do not require initial guesses, and is less sensitive to parameter boundaries.

For the projective method, rescaling the data used in the optimization procedure reduces its dependence on the magnitudes of the spherical harmonic coefficients, and increases the performance for low-intensity distributions. However, in order to maintain good performance across the full range of signal intensities, the rescaling factor needs to be adjusted according to the magnitude of the coefficients. The simplest approach would be to use the first energy range to determine the rescaling factor and apply the same factor to all other energy ranges. This would ensure that the coefficients values are sufficiently large for the optimization, while avoiding the need to adjust the rescaling factor for each signal intensity.

A further improvement of the machine learning approach would be to add physical constraints into the model, ensuring that the predicted solutions remain physically meaningful. The use of k-fold cross-validation for dataset splitting could also be investigated, as this would allow a more efficient use of the available data, and pruning could be explored in order to reduce the computational complexity while maintaining the same accuracy. The method could also be further optimized through additional tuning of hyperparameters, and the investigation of alternative network architectures. Furthermore, the test set used for the final evaluation was not fully unbiased, as it contained data that were very similar or identical to those used during training and validation, and hence the presented performance may be overestimated. Ideally, a completely independent dataset should be used for the final evaluation to ensure an unbiased assessment. Also, the machine learning approach investigated in this work is based only on the same

principles as the direct method. It would therefore be interesting to investigate whether instead using the projective method as a basis for a machine learning approach could improve its performance.

5.2.2 Experimental Application of Results

The results of this work can hopefully provide meaningful guidance for experimental measurements. The results show that all three methods can tolerate relatively high levels of Gaussian noise, while the number of detected electrons appears to be a much more limiting factor for the performance. If the experimental count rate is known, it should be possible to estimate the minimum measurement time needed to collect a sufficient number of counts for a reliable reconstruction. Since the three methods have different Poisson noise thresholds, the required measurement time will depend on the reconstruction method used.

Furthermore, it would be very interesting to investigate whether the Gaussian and Poisson noise limits predicted from the results are consistent with those observed experimentally. This would provide a way to evaluate how well the artificial noise used in this work describes the conditions encountered in actual experiments.

5.2.3 Extension of the Physical Scope

In addition to improving the current reconstruction methods, future work should also focus on extending the physical scope of the study. The work presented in the thesis is limited to ionization from an s -subshell, resulting in relatively simple descriptions of the angular distributions. A natural extension would be to consider ionization from a p -subshell with either specific or arbitrary polarization, which would result in more complex angular distributions and therefore a more challenging reconstruction problem. This would allow the reconstruction methods to be evaluated for more general atomic systems, making them relevant to a wider range of applications. Additionally, investigating more complex processes, such as two-step photoionization, could provide further insight into the ability of the reconstruction methods to describe more complicated electron dynamics.

References

- [1] Y. Cheng et al. “Controlling Photoionization Using Attosecond Time-Slit Interferences”. In: *Proceedings of the National Academy of Sciences* 117.20 (2020), pp. 10727–10732. DOI: 10.1073/pnas.1921138117.
- [2] J.M. Dahlström et al. “Theory of attosecond delays in laser-assisted photoionization”. In: *Chemical Physics* 414 (2013). Attosecond spectroscopy, pp. 53–64. ISSN: 0301-0104. DOI: <https://doi.org/10.1016/j.chemphys.2012.01.017>. URL: <https://www.sciencedirect.com/science/article/pii/S0301010412000298>.
- [3] Sara Mikaelsson et al. “A high-repetition rate attosecond light source for time-resolved coincidence spectroscopy”. In: *Nanophotonics* 10.1 (2021), pp. 117–128. DOI: <https://doi.org/10.1515/nanoph-2020-0424>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1515/nanoph-2020-0424>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1515/nanoph-2020-0424>.
- [4] MIT OpenCourseWare. *5.111 Lecture Summary 6: Principles of Chemical Science*. https://ocw.mit.edu/courses/5-111sc-principles-of-chemical-science-fall-2014/5c828fce769b146e5820a7d6ed0abfc2/MIT5_111F14_Lec6.pdf. 2014.
- [5] Federico Vismarra. “Strong Field-Electron-XUV Interaction Simulator: Study of Ensemble Effects in Attosecond Streaking Spectroscopy”. Master’s Thesis in Physics Engineering. Politecnico di Milano, 2019-2020.
- [6] Christoph Dittel. *Quantum Theory of Attosecond Light-Matter Interaction*. Lecture notes, Spring semester 2025, February version. Chapter 3.3.1, pages 34-38. 2025.
- [7] Kasra Amini et al. “Symphony on strong field approximation”. In: *Reports on Progress in Physics* 82.11 (Oct. 2019), p. 116001. DOI: 10.1088/1361-6633/ab2bb1. URL: <https://doi.org/10.1088/1361-6633/ab2bb1>.
- [8] Christoph Dittel. *Quantum Theory of Attosecond Light-Matter Interaction*. Lecture notes, Spring semester 2025, February version. Chapter 3.3.3, pages 42-47. 2025.
- [9] Marcus J. Dahlström. *Strong Field Approximation for High Order Harmonic Generation with $\omega/2\omega$ Laser Fields*. Tech. rep. LRAP-381. Lund, Sweden: Lund University, Atomic Physics, 2007.
- [10] M. Gisselbrecht et al. “Optimization of momentum imaging systems using electric and magnetic fields”. In: *Review of Scientific Instruments* 76.1 (Dec. 2004), p. 013105. ISSN: 0034-6748. DOI: 10.1063/1.1832411. eprint: https://pubs.aip.org/aip/rsi/article-pdf/doi/10.1063/1.1832411/15830403/013105_1_online.pdf. URL: <https://doi.org/10.1063/1.1832411>.
- [11] Division of Mathematical Physics, Department of Physics, Lund University. *Atomic Structure and Dynamics*. <https://www.matfys.lu.se/research/atomic-structure-and-dynamics/>. 2021.
- [12] Carsten Winter. *High-order Harmonic Generation*. <https://www.uni-muenster.de/Physik.PI/Zacharias/en/research/hh/hhg.html>. AG Prof. H. Zacharias, University of Münster, 2026.
- [13] Yu-Chen Cheng et al. “Controlling photoionization using attosecond time-slit interferences”. In: *Proceedings of the National Academy of Sciences* 117.20 (2020), pp. 10727–10732. DOI: 10.1073/pnas.1921138117. URL: <https://www.pnas.org/doi/abs/10.1073/pnas.1921138117>.
- [14] Meng Han et al. “Separation of photoionization and measurement-induced delays”. In: *Science Advances* 10.4 (2024), eadj2629. DOI: 10.1126/sciadv.adj2629. eprint: <https://www.science.org/doi/pdf/10.1126/sciadv.adj2629>. URL: <https://www.science.org/doi/abs/10.1126/sciadv.adj2629>.
- [15] Mark Fox. *A Student’s Guide to Atomic Physics*. 1st ed. Cambridge: Cambridge University Press, 2018, pp. 1–17. ISBN: 978-1-108-44631-0.
- [16] David Ayuso Molinero. “Coherence in Molecular Photoionization”. PhD thesis. Universidad Autónoma de Madrid, 2015.
- [17] JN Das. “Partial Wave Expansion of Three-particle Continuum. States in Hyperspherical Coordinates: Application to Ionisation Problems”. In: *Australian Journal of Physics* 47.6 (Dec. 1994), pp. 743–750. ISSN: 0004-9506. DOI: 10.1071/PH940743. eprint: <https://connectsci.au/ph/article-pdf/47/6/743/1141932/ph940743.pdf>. URL: <https://doi.org/10.1071/PH940743>.

- [18] Eric W. Weisstein. *Spherical Harmonic*. From MathWorld-A Wolfram Resource. URL: <https://mathworld.wolfram.com/SphericalHarmonic.html>.
- [19] Giuseppe Lingetti and Paolo Pani. “General spherical harmonic bra-ket overlap integrals of trigonometric functions”. In: *Classical and Quantum Gravity* 40.5 (Feb. 2023), p. 057001. DOI: 10.1088/1361-6382/acb880. URL: <https://doi.org/10.1088/1361-6382/acb880>.
- [20] Christopher J. Foot. *Atomic Physics*. 1st ed. Oxford: Oxford University Press, 2005, pp. 22–41. ISBN: 978-0-198-50696-6.
- [21] Shahzaib Iqbal et al. “Recent trends and advances in fundus image analysis: A review”. In: *Computers in Biology and Medicine* 151 (Nov. 2022), p. 106277. DOI: 10.1016/j.combiomed.2022.106277.
- [22] IB-Lenhardt AG. *AWGN (Additive White Gaussian Noise)*. <https://ib-lenhardt.com/kb/glossary/awgn>. 2025.
- [23] MathWorks. *awgn: Add white Gaussian noise to signal*. <https://www.mathworks.com/help/comm/ref/awgn.html>. 2026.
- [24] Google Inc. Samuel W. Hasinoff. *Photon, Poisson noise*. <https://people.csail.mit.edu/hasinoff/pubs/hasinoff-photon-2012-preprint.pdf>.
- [25] University of Glasgow. *Quantum Mechanics Lecture Notes (QF3)*. https://radio.astro.gla.ac.uk/old_0A_course/pw/qf3.pdf. n.d.
- [26] MathWorks. *poissrnd: Random numbers from Poisson distribution*. <https://www.mathworks.com/help/stats/poissrnd.html>. 2026.
- [27] MathWorks. *fminon: solve constrained nonlinear multivariable minimization problem*. <https://se.mathworks.com/help/optim/ug/fmincon.html>, year = 2026.
- [28] Silay Aytaç Yükcü et al. “Calculation of the Gaunt coefficients over real spherical harmonics”. In: *Canadian Journal of Physics* 103.4 (2025), pp. 321–327. DOI: 10.1139/cjp-2024-0161. URL: <https://doi.org/10.1139/cjp-2024-0161>.
- [29] Silay Aytaç Yükcü, Niyazi Yükcü, and Emin Öztekin. “New representations for Gaunt coefficients”. In: *Chemical Physics Letters* 735 (2019), p. 136769. ISSN: 0009-2614. DOI: <https://doi.org/10.1016/j.cplett.2019.136769>. URL: <https://www.sciencedirect.com/science/article/pii/S000926141930750X>.
- [30] The Sage Development Team. *Wigner, Clebsch-Gordan, Racah, and Gaunt coefficients*. <http://sporadic.stanford.edu/reference/functions/sage/functions/wigner.html>. Sage 9.3.beta9 Reference Manual. 2021.
- [31] Mattias Ohlsson, Patrik Edén, and Vincent Brebion. *Introduction to Artificial Neural Networks and Deep Learning*. Lecture notes, Fall 2025. Course codes: BERN04/EXTQ41/NMV012F. Lund University, 2025.
- [32] Pragati Baheti. *Train Test Validation Split: How To Best Practices [2024]*. <https://www.v7darwin.com/blog/train-validation-test-set>. 2021.
- [33] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. <http://www.deeplearningbook.org>. MIT Press, 2016.
- [34] MathWorks. *Create Simple Deep Learning Neural Network for Classification*. <https://se.mathworks.com/help/deeplearning/ug/create-simple-deep-learning-network-for-classification.html>. 2026.
- [35] MathWorks. *adamupdate: Update parameters using adaptive moment estimation (Adam)*. <https://se.mathworks.com/help/deeplearning/ref/adamupdate.html>. 2026.
- [36] MathWorks. *atan2: Four-quadrant inverse tangent*. <https://www.mathworks.com/help/matlab/ref/double.atan2.html>. 2026.
- [37] Jay L. Devore. *Probability and Statistics for Engineering and the Sciences*. 8th ed. Boston, MA: Brooks/Cole, Cengage Learning, 2012, pp. 483–485. ISBN: 978-0-538-73352-6.
- [38] George Casella and Roger L. Berger. *Statistical Inference*. 2nd ed. Pacific Grove, CA: Duxbury, Thomson Learning, 2001, pp. 556–557. ISBN: 0-534-24312-6.

- [39] Davide Chicco, Matthijs J. Warrens, and Giuseppe Jurman. “The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation”. In: *PeerJ Computer Science* 7 (2021), e623. DOI: 10.7717/peerj-cs.623.
- [40] Tarald O. Kvålseth. “Cautionary Note about R²”. In: *The American Statistician* 39.4 (1985), pp. 279–285. ISSN: 00031305, 15372731. URL: <http://www.jstor.org/stable/2683704>.
- [41] Mark Fox. *A Student’s Guide to Atomic Physics*. 1st ed. Cambridge: Cambridge University Press, 2018, pp. 84–102. ISBN: 978-1-108-44631-0.
- [42] Mark Fox. *A Student’s Guide to Atomic Physics*. 1st ed. Cambridge: Cambridge University Press, 2018, pp. 40–56. ISBN: 978-1-108-44631-0.
- [43] Matthew Reed. *Solving the Helium Atom. Or: Why does Chemistry Exist?* Final project report MATH164 HM: Scientific Computing. Claremont, CA: Harvey Mudd College, 2007.

A Basic Quantum Mechanics

A.1 Commutation Relations

The commutation of two quantum mechanical operators is defined as [41]:

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \quad (\text{A.1})$$

If two operators do not commute, their values cannot be known simultaneously, and the uncertainty principle must be obeyed [41]:

$$\Delta A^2 \Delta B^2 \geq \frac{1}{4} |[\hat{A}, \hat{B}]|^2 \quad (\text{A.2})$$

A.2 Parity Operator

In quantum mechanics, parity refers to the sign change under inversion about the origin of the position vector. The parity operator is defined as [20]

$$\hat{P}\psi = -P\psi \quad (\text{A.3})$$

It can easily be proven that $\hat{P}^2\psi = P^2\psi$, and since two parity operations corresponds to there being no net charge, the parity operator must have eigenvalues $P = \pm 1$ [20]. This corresponds to even and odd parity wavefunctions defined as [20] [42]:

$$\hat{P}\psi = \psi \quad \hat{P}\psi = -\psi \quad (\text{A.4})$$

The use of the operator on spherical harmonics gives [20]

$$\hat{P}Y_{l,m} = (-1)^l Y_{l,m} \quad (\text{A.5})$$

A.3 Orbital Angular Momentum

Electrons orbiting the nucleus in an atom possess an angular momentum defined by the radial position and linear momentum operator as [20] [41]:

$$\hat{\mathbf{L}} = \mathbf{r} \times \hat{\mathbf{p}}, \quad \hat{\mathbf{p}} = -i\hbar\nabla \quad (\text{A.6})$$

The Cartesian components of the operator can be written as [41]:

$$\hat{l}_x = -i\hbar(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}), \hat{l}_y = -i\hbar(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}), \hat{l}_z = -i\hbar(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}) \quad (\text{A.7})$$

The following commutation relations apply [41]:

$$[\hat{l}_x, \hat{l}_y] = i\hbar\hat{l}_z, \quad [\hat{l}_y, \hat{l}_z] = i\hbar\hat{l}_x, \quad [\hat{l}_z, \hat{l}_x] = i\hbar\hat{l}_y, \quad [\hat{\mathbf{L}}^2, \hat{L}_z] = 0 \quad (\text{A.8})$$

Since $\hat{\mathbf{L}}^2$ and $\hat{L}_z$ commute, they have a common set of eigenstates and can be simultaneously measured [41]. The orbital angular momentum for an electron is characterized by quantum numbers l and m_l , with [41]:

$$\mathbf{L}^2 = l(l+1)\hbar^2, \quad L_z = m_l\hbar \quad (\text{A.9})$$

For each l , the magnetic quantum number can take $2l+1$ values, giving rise to degeneracy in isolated atoms. This degeneracy can be broken by external perturbations such as magnetic or electric fields. If

look at system multiple electrons, total orbital angular momentum for system is denoted L along with operator $\mathbf{L}$ [41].

Since the magnetic quantum number m_l appears frequently in the notation of spherical harmonics, the notation m will be used throughout this work instead of m_l , unless otherwise specified.

A.4 Spin Angular Momentum

Apart from orbital angular momentum, electrons do also possess an additional type of angular momentum called spin. Similar to orbital angular momentum, the spin of an electron is described by quantum numbers s and m_s for which [41]:

$$\mathbf{s}^2 = s(s+1)\hbar^2, \quad s_z = m_s\hbar \quad (\text{A.10})$$

The quantum number m_s can take on $(2s+1)$ values, causing degeneracy in isolated atoms. The capital letter S is used to denote whole spin for a system including more than one single electron [20].

Since atoms with single s-shell valence electron only two directions, so to say $(2s+1) = 2$, it is apparent that $s = 1/2$ and $m_s = \pm 1/2$ [41]. Though unlike for orbital angular momentum, spin is a more abstract concept that cannot be written using eigenstates expressed in angular coordinates. Instead, eigenstates are usually expressed using the Dirac's ket notation $|sm_s\rangle$. The following shorthand notation is commonly used [20]:

$$|s = \frac{1}{2}, m_s = \frac{1}{2}\rangle = |\uparrow\rangle, \quad |s = \frac{1}{2}, m_s = -\frac{1}{2}\rangle = |\downarrow\rangle \quad (\text{A.11})$$

A.5 Total Angular Momentum

The total angular momentum is defined as the sum of its orbital and spin angular momenta [20] [41]:

$$\mathbf{j} = \mathbf{l} + \mathbf{s}, \quad |\mathbf{j}|^2 = j(j+1)\hbar^2, \quad j_z = m_j\hbar \quad (\text{A.12})$$

Allowed values are given from applying $j = l \oplus s$, and using $s = 1/2$ gives $j = l \oplus s = (l \pm 1/2)$ with expectation when $l = 0$ which gives $j = 1/2$ [20] [41].

Total angular momentum for whole system is denoted $\mathbf{J}$, and specified by the quantum number J . The total angular momentum must be constant in the absence of external perturbation, which requires an isolated atom to have well-defined J -states. Changes in total angular momentum due to external perturbations are governed by selection rules [41].

B Quantum Mechanical Representation of Atoms

B.1 Hydrogenic Atoms

In the development of quantum theory, the hydrogen atom has played a central role because of its simple structure and the spherical symmetry of its potential [20]. Hydrogenic atoms are atoms consisting of a nucleus of arbitrary charge and a single bound electron, whose structure can be described using the solution for the hydrogen atom. One example of a hydrogenic atom is singly ionized helium [43].

B.2 Solution of the Angular Part of the Schrödinger Equation

The Schrödinger equation for an electron in a spherically symmetric potential $V(r)$ is given as

$$\left(\frac{-\hbar^2}{2m_e}\nabla^2 + V(r)\right)\psi = E\psi \quad (\text{B.1})$$

where $\hbar$ is the reduced Planck constant, m_e is the electron mass, ∇^2 is the Laplacian operator describing spatial variation of the wavefunction, ψ is the wavefunction of the electron, and E is the corresponding eigenvalue energy. [20] [43].

Using spherical coordinates (r, θ, ϕ) , the Laplacian operator can be written as [20]:

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 \frac{\partial}{\partial r}) - \frac{1}{r^2 \hbar^2} \mathbf{l}^2, \quad \mathbf{l}^2 = -\hbar^2 \left\{ \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} (\sin \theta \frac{\partial}{\partial \theta}) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right\} \quad (\text{B.2})$$

This partial differential equation is solved by using a separable ansatz $\psi = R(r)Y(\theta, \phi)$ which separates the wavefunction into radial and angular components. Substitution into equation B.1 and separation of variables show that both the radial and angular parts must equal a constant, resulting in the following angular eigenvalue equation [20]:

$$\mathbf{l}^2 Y = bY \quad (\text{B.3})$$

where b is the eigenvalue associated with the angular momentum operator.

Assuming that the function can be separated into components depending on different variables, $Y(\theta, \phi) = \Theta(\theta)\Phi(\phi)$, the separation of variables again shows that both parts must equal a constant. The function $\Phi(\phi)$ gives the harmonic motion of the electron, and has solution [20]:

$$\Phi(\phi) = Ae^{im\phi} + Be^{-im\phi}, \quad \Phi(\phi + 2\pi) = \Phi(\phi), \quad m \in Z \quad (\text{B.4})$$

Function $\Phi(\phi)$ is a sum of eigenfunctions of the z-component of the orbital angular momentum operator $\hbar l_z = -i\hbar \frac{\partial}{\partial \phi}$, with eigenvalue m^2 [20].

To find function $Y(\theta, \phi)$ and its eigenvalues, the so called ladder operators can be utilized [20]:

$$l_+ = l_x + il_y = e^{i\phi} \left(\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right), \quad l_- = l_x - il_y = e^{-i\phi} \left(-\frac{\partial}{\partial \theta} + i \cot \theta \frac{\partial}{\partial \phi} \right) \quad (\text{B.5})$$

The operator l_- is called the lowering operator and transforms function with eigenvalue n into another angular momentum eigenfunction with eigenvalue $n - 1$, whereas operator l_+ is a raising operator that transforms function with eigenvalue n into another angular momentum eigenfunction with eigenvalue $n + 1$ [20].

These operators commute with $\mathbf{l}^2$, hence Y , $l_+ Y$ and $l_- Y$ are eigenfunctions of $\mathbf{l}^2$ with same eigenvalue $b = l(l + 1)$. For each b there is a maximum value of magnetic quantum number $m_{max} = l$. Using the raising operator on an eigenfunction with m_{max} will therefore yields zero, and solving $l_+ Y = 0$ gives eigenfunctions with $m_{max} = l$ of the form [20]:

$$Y_{l, m_{max}}(\theta, \phi) \propto \sin^l \theta e^{il\phi} \quad (\text{B.6})$$

A general solution is obtained by using the lowering operator [20]:

$$Y_{l, m} \propto (l_-)^{l-m} \sin^l \theta e^{il\phi}, \quad l = k, \quad m_l = m = -k, \dots, 0, \dots, k \quad (\text{B.7})$$

Noteworthy is that function $Y_{0,0}$ have no angular dependence, and is therefore a constant. These angular components of the wavefunction, $Y_{l, m_{max}}$, are called spherical harmonic functions, and they are normalized using [20]:

$$\int_0^{2\pi} \int_0^\pi |Y_{l, m}|^2 \sin \theta d\theta d\phi = 1 \quad (\text{B.8})$$

The total probability distribution of an electron given by the angular and radial distribution squared [20]

$$|\psi(r, \theta, \phi)|^2 = R_{n, l}^2(r) |Y_{l, m}(\theta, \phi)|^2 \quad (\text{B.9})$$

C MATLAB Scripts

C.1 Direct Fitting Method

```
1 function [res_three] = func_opt_three_sshell_direct(P_opt,theta_axis,phi_axis,
2     init)
3     lb = [-0.01 -0.01 -1.4*pi];
4     ub = [0.08 0.08 1.4*pi];
5     [theta_slice, phi_slice] = meshgrid(theta_axis, flip(phi_axis));
6
7     Y10 = func_Ylm(theta_slice,phi_slice,1,0);
8     Y11 = func_Ylm(theta_slice,phi_slice,1,1);
9     Y1m1 = func_Ylm(theta_slice,phi_slice,1,-1);
10
11     j = 2; k = 3; p = 4;
12
13     fun_min = @(X) optimization_error(X,P_opt,Y10,Y11,Y1m1);
14
15     options = optimoptions("fmincon", 'Algorithm', 'interior-point', '
16         MaxIterations', 1000,'StepTolerance', 10^(-9), 'MaxFunctionEvaluations',
17         1000);
18
19     coeff = fmincon(fun_min, init, [], [], [], [], lb, ub, [], options);
20
21     res_three = [j, k, p, coeff(1), coeff(2), coeff(3), coeff(2), coeff(3)-pi,
22         fun_min(coeff)];
23 end
```

Listing 1: Main optimization function

```
1 function err = optimization_error(X,P_opt,Y10,Y11,Y1m1)
2
3     c10 = X(1);
4     c11 = X(2) * exp(1i*X(3));
5     c1m1 = X(2) * exp(1i*(X(3)-pi));
6
7     P_model = abs(c10*Y10 + c11*Y11 + c1m1*Y1m1).^2;
8
9     err = sum(sum(abs(P_opt - P_model)));
10
11 end
```

Listing 2: Optimization error function

C.2 Projective Method

```
1 function [res_three] = func_opt_three_sshell_projective(coeff_opt, init)
2
3     lb = [-0.01 -0.01 -1.4*pi];
4     ub = [0.08 0.08 1.4*pi];
5
6     gaunt_idx = [
7     1, 0, 1, 0, 0, 0; % conj(Y10)*Y10
8     1, 1, 1, 1, 0, 0; % conj(Y11)*Y11
9     1, -1, 1, -1, 0, 0; % conj(Y1m1)*Y1m1
10    1, 0, 1, 0, 2, 0; % conj(Y10)*Y10
11    1, 1, 1, 1, 2, 0; % conj(Y11)*Y11
```

```

12 1, -1, 1, -1, 2, 0; % conj(Y1m1)*Y1m1
13 1, 0, 1, -1, 2, 1; % conj(Y1m1)*Y10
14 1, 1, 1, 0, 2, 1; % conj(Y10)*Y11
15 1, 0, 1, 1, 2, -1; % conj(Y11)*Y10
16 1, -1, 1, 0, 2, -1; % conj(Y10)*Y1m1
17 1, 1, 1, -1, 2, 2; % conj(Y1m1)*Y11
18 1, -1, 1, 1, 2, -2; % conj(Y11)*Y1m1
19 ];
20
21 gaunt_func = arrayfun(@(idx) func_gaunt(gaunt_idx(idx,1), gaunt_idx(idx,2),
    gaunt_idx(idx,3), gaunt_idx(idx,4), gaunt_idx(idx,5), gaunt_idx(idx,6)),
    1:size(gaunt_idx,1));
22
23 j = 2; k = 3; p = 4;
24
25 fun_min = @(X) optimization_error(X, coeff_opt, gaunt_func);
26
27 options = optimoptions("fmincon", 'Algorithm', 'interior-point', '
    MaxIterations', 1000, 'StepTolerance', 10^(-9), 'MaxFunctionEvaluations',
    1000);
28
29 coeff = fmincon(fun_min, init, [], [], [], [], lb, ub, [], options);
30
31 res_three = [j, k, p, coeff(1), coeff(2), coeff(3), coeff(2), coeff(3)-pi,
    fun_min(coeff)];
32
33 end

```

Listing 3: Main optimization function

```

1 function err = optimization_error(X, coeff_opt, gaunt_func)
2
3 c10 = X(1);
4 c11 = X(2) * exp(1i*X(3));
5 c1m1 = X(2) * exp(1i*(X(3)-pi));
6
7 c_00 = gaunt_func(1) *abs(c10)^2 + ... % l=0 m=0
8       gaunt_func(2)*abs(c11)^2 + ...
9       gaunt_func(3)*abs(c1m1)^2;
10
11 c_20 = gaunt_func(4)*abs(c10)^2 + ... % l=2 m=0
12       gaunt_func(5)*abs(c11)^2 + ...
13       gaunt_func(6)*abs(c1m1)^2;
14
15 c_21 = gaunt_func(7)*c10*conj(c1m1) + ... % l=2 m=1
16       gaunt_func(8)*c11*conj(c10);
17
18 c_2m1= gaunt_func(9)*c10*conj(c11) + ... % l=2 m=-1
19       gaunt_func(10)*c1m1*conj(c10);
20
21 c_22 = gaunt_func(11)*c11*conj(c1m1); % l=2 m=2
22
23 c_2m2 = gaunt_func(12)*c1m1*conj(c11); % l=2 m=-2
24
25 coeff_model = [c_00; c_20; c_21; c_2m1; c_22; c_2m2];
26
27 err = sum(abs(real(coeff_opt) - real(coeff_model))) + sum(abs(imag(coeff_opt) -
    imag(coeff_model)));
28
29 end

```

Listing 4: Optimization error function

C.3 Machine Learning Method

```
1 layers = [  
2     imageInputLayer([n_theta n_phi 1], 'Normalization', 'none', 'Name', 'input')  
3  
4     convolution2dLayer(3,8,'Padding','same')  
5     reluLayer  
6     maxPooling2dLayer(2,'Stride',2)  
7  
8     convolution2dLayer(3,16,'Padding','same')  
9     reluLayer  
10    maxPooling2dLayer(2,'Stride',2)  
11  
12    convolution2dLayer(3,32,'Padding','same')  
13    reluLayer  
14  
15    fullyConnectedLayer(64)  
16    reluLayer  
17  
18    fullyConnectedLayer(32)  
19    reluLayer  
20  
21    fullyConnectedLayer(4)  
22 ];  
23  
24 lgraph = layerGraph(layers);  
25 net = dlnetwork(lgraph);
```

Listing 5: CNN Architecture

```
1  
2 for epoch = 1:num_epochs  
3  
4     X_dl = dldarray(X_train, 'SSCB');  
5     X_dl_noise = dldarray(X_train_noise, 'SSCB');  
6  
7     shuffle = randperm(n_slices_train);  
8     X_dl_shuffle = X_dl(:,:, :, shuffle);  
9     X_dl_shuffle_noise = X_dl_noise(:,:, :, shuffle);  
10  
11     num_batches = 0;  
12     train_loss = 0;  
13  
14     for i = 1:mini_batch_size:n_slices_train  
15  
16         idx = i:min(i+mini_batch_size-1, n_slices_train);  
17         X_batch = X_dl_shuffle(:,:, :, idx);  
18         X_batch_noise = X_dl_shuffle_noise(:,:, :, idx);  
19  
20         iter = iter + 1;  
21  
22         [loss, gradients] = dlfeval(@loss_function, net, X_batch, X_batch_noise,  
23                                     Y10, Y11, Y1m1);  
24  
25         [net, average_grad, average_sq_grad] = adamupdate(net, gradients,  
26                                                         average_grad, average_sq_grad, iter, learning_rate);  
27  
28         train_loss = train_loss + loss;  
29  
30         num_batches = num_batches + 1;  
31  
32     end
```

```

31
32     train_loss_tot(epoch) = train_loss / num_batches;
33
34 end

```

Listing 6: Training loop

```

1  function [loss, gradients] = loss_function(net, X, X_noise, Y10, Y11, Y1m1)
2
3     coeff = forward(net, X_noise);
4     batch_size = size(X,4);
5     loss = 0;
6
7     a = coeff(1,:);
8     b = coeff(2,:);
9     cphi = coeff(3,:);
10    sphi = coeff(4,:);
11    norm = sqrt(cphi.^2 + sphi.^2);
12    phi = atan2(sphi ./ norm, cphi ./ norm);
13
14    for k = 1:batch_size
15        c10 = real(a(k));
16        c11 = real(b(k)).* exp(1i*phi(k));
17        c1m1 = real(b(k)).* exp(1i*(phi(k) - pi));
18
19        P_pred = abs(c10 .* Y10 + c11 .* Y11 + c1m1 .* Y1m1).^2;
20
21        target = X(:, :, 1, k);
22        loss = loss + mean((P_pred - target).^2, 'all');
23
24    end
25
26    loss = loss / batch_size;
27    gradients = dlgradient(loss, net.Learnables);
28
29 end

```

Listing 7: Loss function