Ptychography for imaging and spatial beam characterization at soft X-ray free-electron lasers

Dissertation
zur Erlangung des Doktorgrades
an der Fakultät für Mathematik, Informatik und
Naturwissenschaften
Fachbereich Physik
der Universität Hamburg

vorgelegt von
Konstantin Kharitonov

University of Hamburg
Hamburg
February-June
2022

Gutachter/innen der Dissertation:	Dr. Elke Plönjes-Palm Prof. Dr. Christian Schroer
Zusammensetzung der Prüfungskommission:	Dr. Elke Plönjes-Palm Prof. Dr. Tais Gorkhover Prof. Dr. Christian Schroer Prof. Dr. Henry Chapman Prof. Dr. Sven-Olaf Moch
Vorsitzende/r der Prüfungskommission:	Prof. Dr. Sven-Olaf Moch
Datum der Disputation:	16.09.2022
Vorsitzender Fach-Promotionsausschuss PHYSIK:	Prof. Dr. Wolfgang J. Parak
Leiter des Fachbereichs PHYSIK:	Prof. Dr. Günter H. W. Sigl
Dekan der Fakultät MIN:	Prof. Dr. Heinrich Graener

Eidesstattliche Versicherung / Declaration on oath

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben.

Hamburg, October 7, 2022

Unterschrift der Doktorandin / des Doktoranden

Ich versichere, dass dieses gebundene Exemplar der Dissertation und das in elektronischer Form eingereichte Dissertationsexemplar (über den Docata-Upload) und das bei der Fakultät (zuständiges Studienbüro bzw. Promotionsbüro Physik) zur Archivierung eingereichte gedruckte gebundene Exemplar der Dissertationsschrift identisch sind.

Hamburg, October 7, 2022

Unterschrift der Doktorandin / des Doktoranden

Abstract

Ptychography is a scanning coherent diffraction imaging (CDI) technique capable of performing high-resolution imaging of extended samples. Moreover, ptychography can reconstruct the illuminating wavefield, thus making it applicable for beam characterization. Currently, ptychography is broadly used at synchrotron light sources, while implementation at X-ray free-electron-lasers (FEL) is still novel.

Self-amplified spontaneous emission (SASE) FELs produce extremely bright, ultra-short, and highly-coherent light pulses, making them ideal for CDI experiments. However, SASE FEL light pulses are only partially spatially-coherent and undergo significant spatial fluctuations due to the stochastic nature of the SASE process. This complicates ptychography experiments at SASE FELs and requires a modification of the existing ptychography algorithms. However, most of them are difficult to modify because they use a hand-derived closed-form gradient expression. This can be solved by a computational technique called automatic differentiation (AD), allowing on-the-fly computation of the gradients of differentiable functions. This way, AD removes the need for a manual derivation of the gradients enabling the creation of flexible and efficient ptychography algorithms.

Within this work, an AD-based ptychography framework was developed, computationally realized, and applied for classical ptychography at SASE FELs. Both forward model and experimental setup are validated during a first ptychography experiment at FLASH2, resulting in a high-resolution reconstruction of the sample and the shot-to-shot unique probe wavefield. Potential improvement of achievable throughput and resolution of the ptychography at FLASH2 by using the more advanced PERCIVAL detector is demonstrated. Finally, the results of a practical application of ptychography at FELs for imaging of plasma-treated thin plastic samples are shown.

The scanning nature of ptychography hinders its use for imaging of dynamical samples or any form of time-resolved imaging. For the visible range, this problem was solved in the single-shot ptychography setup proposed by Sidorenko et al. [1]. However, it is based on refractive optics making it hard to implement in the X-ray range. A novel X-ray compatible single-shot ptychography setup and a suitable forward model for the AD-based ptychography framework were developed to demonstrate for the first time a single-shot ptychography at soft X-ray wavelengths. The proposed setup was evaluated during the first single-shot experiment performed at FLASH2, resulting in a high-resolution reconstruction of extended samples.

Overall, this work shows how computational and technical developments enable the use of various ptychography techniques at FELs for high-throughput high-resolution imaging and shot-to-shot beam characterization.

Zusammenfassung

Ptychographie ist eine Technik der scannenden kohärenten Beugungsbildgebung (CDI), die in der Lage ist, eine hochauflösende Bildgebung der gesamten Probe durchzuführen. Darüber hinaus kann auch das Beleuchtungswellenfeld rekonstruiert werden, womit sie auch für die Strahlcharakterisierung einsetzbar ist. Ptychographie an Synchrotronstrahlungsquellen ist weit verbreitet, die Anwendung an Freie-Elektronen-Lasern (FEL) im Röntgenbereich ist jedoch noch ganz neu.

FELs, die auf dem Prinzip der selbst verstärkenden, spontanen Emission (SASE) beruhren, erzeugen extrem helle, ultra kurze und hoch kohärente Lichtpulse, was sie für CDI Experimente ideal macht. Die Lichtpulse sind jedoch nur teilweise räumlich kohärent und unterliegen aufgrund der stochastischen Natur des SASE Prozesses, erheblichen räumlichen Schwankungen. Dies erschwert Experimente an SASE FELs und erfordert eine Modifikation der bestehenden Ptychographie Algorithmen. Die meisten von ihnen sind jedoch schwer zu ändern, da sie einen von Hand abgeleiteten Gradientenausdruck in geschlossener Form verwenden. Dies kann durch eine Berechnungstechnik gelöst werden, die als automatische Differenzierung (AD) bezeichnet wird und eine Berechnung der Gradienten differenzierbarer Funktionen während der Laufzeit ermöglicht. Auf diese Weise macht AD die manuelle Ableitung der Gradienten überflüssig und ermöglicht die Erstellung flexibler und effizienter Ptychographie Algorithmen.

Der erste Teil dieser Arbeit beschreibt die Entwicklung eines AD-basierten Ptychographie Frameworks, präsentiert Details der AD-basierte Ptychographie und diskutiert die Bestandteile ihrer computergestützten Realisierung.

Der zweite Teil zeigt die Verwendung dieses Frameworks für klassische Ptychographie an SASE FELs. Das Vorwärtsmodell für die AD-basierte Ptychographie und der experimentelle Aufbau, der speziell für klassische Ptychographie an SASE FELs entwickelt wurde, werden gezeigt. Beides wurde in einer ersten Messzeit bei FLASH2 validiert. Es gelang sowohl eine hochauflösenden Rekonstruktion der Probe als auch des von Puls zu Puls vairiierenden einfallenden Wellenfeldes. Die Verbesserung der Durchsatzrate und der Aufloösung wird durch den Einsatz des fortschrittlichen PERCIVAL Detektors demonstriert. Die Ergebnisse einer praktischen Anwendung von Pythographie an FELS wird an Hand von dünnen, plasmabehandelten Kunststoffproben gezeigt.

Da in der Pythographie die Proben normalerweise gescannt werden, können weder zeitaufgelösten Messungen erfolgen noch dynamische Proben untersucht werden. Im sichtbaren Spektralbereich wurde dieses Problem in dem von Sidorenko et al. [1] vorgeschlagenen Aufbau für Single-Shot Ptychographie gelöst, der refraktive Optiken verwendet und somit nicht direkt auf den Röntgenbereich übertragbar ist. Ein neuartiger experimenteller Aufbau für pulsaufgelöste Ptychographie im Röntgenspektralbereich sowie das Vorwärtsmodell für das AD-basierte Ptychographie-Framework werden im dritten Teil vorgestellt. Hiermit konnten erstmals Single-

Shot-Experimente an FLASH2 durchgeführt und eine hochaufgelöste Rekonstruktion von ausgedehnten Proben berechnet werden.

Diese Arbeit zeigt, wie rechnerische und technische Entwicklungen den Einsatz der verschiedenen Techniken der Ptychographie an FELs für eine hochaufgelöste Bildgebung von Proben mit hoher Durchsatzrate sowie pulsaufgelöster Strahlcharakterisierung ermöglichen.

List of Publications

M. Ravandeh, M. Mehrjoo, **K. Kharitonov**, et al., "X-ray ptychographic imaging and spectroscopic studies of plasma-treated plastic films," Polymers, vol. 14, no. 13, 2022.

K. Kharitonov, M. Mehrjoo, M. Ruiz-Lopez, et al., "Single-shot ptychography at a soft x-ray free-electron laser, PREPRINT (version 1)," available at Research Square [https://doi.org/10.21203/rs.3.rs-1629463/v1], 2022.

K. Kharitonov, M. Mehrjoo, M. Ruiz-Lopez, et al., "Flexible ptychography platform to expand the potential of imaging at free electron lasers," Optics express, vol. 29, no. 14, pp. 22 345–22 365, 2021.

List of Abbreviations

AD ..... Automatic Differentiation
ADAM ..... ADAptive Moment estimation
AFM ..... Atomic Force Microscopy
CCD ..... Charge-Coupled Device
CDI ..... Coherent Diffraction Imaging
CSD ..... Cross-Spectral Density
DFT ..... Discrete Fourier Transform
FEL ..... Free-Electron Laser
FOV ..... Field Of View
FRC ..... Fourier Ring Correlation
FTH ..... Fourier Transform Holography
FWHM ..... Full Width at Half Maximum
GD ..... Gradient Descent
GPU ..... Graphics Processing Unit
HHG ..... High Harmonic Generation
IR ..... Impulse Response
KB mirror ..... Kirkpatrick-Baez mirror
LSE ..... Least Squares Error
MCF ..... Mutual Coherence Function
MGD ..... Momentum Gradient Descend
OPR ..... Orthogonal Probe Relaxation
PET ..... PolyEthylene Terephthalate
ePIE ..... extended Ptychographical Iterative Engine
mPIE ..... momentum Ptychographical Iterative Engine
PNLL ..... Poisson Negative Log Likelihood
RMS ..... Root Mean Square
ROI ..... Region Of Interest
SASE ..... Self-Amplification of Spontaneous Emission
SDC ..... Spectral Degree of Coherence
SEM ..... Scanning Electron Microscopy
TF ..... Transfer Function

TSVD ..... Truncated Singular Value Decomposition
TV ..... Total Variation
TVD ..... Total Variation Denoising
XUV ..... eXtreme UltraViolet

List of Mathematical Symbols

(q_x, q_y) ..... reciprocal coordinate system
(x, y),(v, u),(ζ, η) ..... coordinate systems
A_θ ..... affine transformation matrix
$\mathcal{A}_\theta$ ..... affine transformation operator
α_i ..... modal weights in multimodal ptychography
α_s, β_s ..... step size parameters in gradient optimization
α_{samp},β_{samp} ..... ptychography sampling criteria
$\mathbb{B}$ ..... magnetic field
B_j ..... j-th beamlet wavefield in single-shot ptychography
β_n ..... modal weights
β_{gr_j} ..... diffraction grating efficiency for j-th diffraction ofrder
c ..... speed of light in vacuum
D ..... source size, beam size
d ..... distance perpendicular to the optical axis
δx ..... resolution in the sample plane
e ..... elementary charge
$\mathcal{E}$ ..... error function
E_e ..... electron energy
$\mathbb{E}$ ..... electric field
F ..... Fresnel number
$\mathcal{F}, \mathcal{F}^{-1}$ ..... forward and inverse Fourier transform operators
Γ ..... mutual coherence function
γ ..... Lorentz contraction factor
H_{z_1} ..... transfer function of free space
h_{z_1} ..... impulse response of free space
$I, \tilde{I}$ ..... intensity, approximated intensity
$\mathbb{I}$ ..... current
K ..... undulator parameter
L ..... distance along the optical axis
$\mathcal{L}$ ..... loss function
λ ..... wavelength

M_i ..... modes in multimodal ptychography
m_e ..... electron rest mass
μ ..... spectral degree of coherence
$\mathbb{N}$ ..... natural numbers
$n = 1 - \delta + i\beta$ ..... refraction index
O ..... object function
Ω_f ..... defocus term
ω ..... frequency
$\mathcal{ORT}$ ..... orthonormality operator
P ..... probe function
$\mathbb{P}$ ..... power
$\mathcal{P}_z$ ..... operator of propagation (propagator) in z direction
p_{det} ..... detector pixel size
Ψ ..... complex wavefield
$\mathcal{R}$ ..... regularization term of the loss function
ρ_{FEL} ..... characteristic FEL parameter
S ..... support in phase retrieval
$\mathbb{S}$ ..... spectral density
τ ..... time interval
θ ..... angle
$\mathcal{TV}$ ..... total variation operator
W ..... cross-spectral density
ξ ..... global degree of coherence
ξ_t, ξ_l ..... transverse and longitudinal coherence lengths
z ..... distance along the optical axis

Contents

Chapter 1

Introduction

> Vision is the art of seeing what is invisible to others.
>
> Jonathan Swift

The resolution of classical imaging is limited by the numerical aperture of the imaging optics and the wavelength of the employed light [2]. The appearance of X-ray sources with shorter wavelengths dramatically improved the achievable resolution. However, producing high-precision aberration-free optics with simultaneously high numerical aperture becomes a challenging task for decreasing wavelength. Thus, the creation of an X-ray microscope with atomic-scale resolution is still impossible [3].

Coherent diffraction imaging (CDI) [4] provides a solution to this problem by replacing the low numerical-aperture imaging optics with a 'computational lens.' This can be done by measuring the light scattered by the sample and propagated to the detector plane and, later, computational back-propagation of it to the sample plane. The numerical resolution of this imaging technique can be directly linked to the highest scattering angle measured and thus to the size of the detector, allowing to achieve nanometer resolution. The back-propagation of the scattered wave to the sample plane is possible only when its amplitude and phase are known at each detector pixel. However, due to the high frequency of the X-ray radiation, detectors are able only to measure intensity which is proportional to the modulus of the amplitude of the scattered wave. The missing phase can be retrieved through the complex computational process of phase retrieval [5] enabling the back-propagation

and reconstruction of the sample.

An emergence of a highly-coherent X-ray sources such as synchrotrons [6], high harmonic generation (HHG) [7], and FELs [8], has led to development of a wide range of experimental techniques such as the variations of the full-field CDI [9], including Bragg [10] and the specular [11] reflection configuration, Fourier transform holography (FTH) [12] and Ptychography [13]. All these methods have benefits and drawbacks, determining their area of application. Full-field CDI is capable of high-resolution imaging of objects utilizing *diffraction before destruction* approach [14], [15], but relies on complex iterative phase retrieval algorithms, is sensitive to noise, has limited FOV, and not always results in a unique solution[16]. FTH can produce diffraction-limited [17] single-shot imaging of the sample utilizing the simpler reconstructional algorithms but has low FOV and requires the use of complex sample holders with specially designed reference sources [18]. However, recently an in-flight FTH was demonstrated where the use of nanoclusters as the reference scattering sources solved this problem for imaging of a Mimivirus [19]. Finally, ptychography is capable of routinely producing the high-resolution imaging of the extended samples [13] and, due to the redundancy of data, even capable of producing the sliced 3D reconstructions. However, it is susceptible to the precision of scanning[20] and requires high stability of the beam[21].

Most of the CDI techniques retrieve only the exit-wave at the sample plane. The exit-wave in the projection approximation can be expressed as the product of the complex wavefield and the complex transfer function of the sample. In turn, the transfer function of the sample can be found by the division of the exit-wave by the wavefield function. Thus, the direct reconstruction of the sample requires either to use of a perfectly uniform wavefield, which is not easily experimentally achievable or to perform a beam characterization before the experiment.

Ptychography [13] stands out from the other CDI techniques as it can reconstruct the probe and the sample individually without preliminary knowledge of the probe wavefield and thus can be used for both imaging and beam characterization. Ptychography achieves this by scanning the sample such that multiple diffraction patterns are measured from mutually overlapping regions. The overlap provides redundancy in the measured data, which allows to perform a deconvolution and thus reconstruct the probe and the sample individually. Moreover, it provides addi-

tional constraints allowing ptychographical reconstructions to converge reliably and making them robust to noise.

The dense scanning of the sample is a core of the ptychography technique. It allows reliable high-resolution imaging of extended samples but makes the measurements very time-consuming. This prevents ptychographical imaging of dynamical processes in the sample or time-resolved ptychography. Recently, a solution to this problem was proposed in the form of so-called *single-shot* ptychography. The setup proposed by Sidorenko et al. [1] for the visible range utilizes a lens system to cross several beams on the sample plane simultaneously and thus allows to perform the ptychographic measurement in one shot. However, the use of refractive optics prevents the straightforward transfer of this setup to X-ray wavelengths.

All CDI techniques, including ptychography, rely on measuring the intensity distribution produced by the diffraction of the radiation scattered by the sample. The diffraction effects can only be observed in the case of highly coherent radiation. Additionally, the best achievable resolution of CDI depends on the total amount of scattered photons measured [22]. Therefore, CDI experiments require a very large photon flux and place high demands on the coherence of the photon sources. Free-electron lasers (FELs) [23] are capable of producing such high amounts of coherent photons due to their unprecedented brilliance, which makes them ideal facilities for performing CDI experiments. However, their radiation is still only partially spatially coherent [24] and simultaneously spatially-fluctuating shot-to-shot. This, in turn, increases the need for beam-characterization methods capable of reconstructing the wavefield of the individual FELs pulses and their coherent properties.

Ptychography, being well established at synchrotrons due to the high stability of their beams, still finds a problem with routine implementation at FELs due to their partially-spatially-coherent and spatially-fluctuating radiation. Prior to this work, several succesfull experiments were performed at FELs [25], [26]. However, they had several limitations in comparison to this work. In the first of them [25] the fully coherent forward model was used, thus giving no information about the coherent properties of the beam. Additionally, compound refractive lenses (CRLs) are used in [25] for focusing the beam simultaneously, thus suppressing all the spatial fluctuations and selecting the most coherent part of the beam at a price of limited photon flux. The second one [26] was performed at the seeded FERMI FEL [27]

with highly-spatially-coherent radiation and thus proposed no solution for the shot-to-shot fluctuations of the coherent properties of the illumination.

Development of a ptychographical framework capable of working with shot-to-shot spatially-fluctuating and partially-spatially-coherent beams would be beneficial both for imaging and FEL beam characterization. First, the high brightness of the FEL beams would dramatically decrease the required scanning time and increase the achievable resolution. Second, such a framework would allow high-resolution shot-to-shot beam characterization together with the retrieval of its coherent properties, providing viable insight into machine and beamline optics performance. Additionally, the transfer of single-shot ptychography to X-ray FELs with their ultra-bright and ultra-short beams would allow performing imaging of extended samples in a pump-probe manner with unprecedented time resolution.

This work describes the development of a ptychography model capable of working with simultaneously partially-spatially-coherent and spatially-fluctuating sources. It presents a flexible computational framework for ptychography and demonstrates the results of the classical ptychography experiments performed at a free-electron laser in Hamburg (FLASH). Additionally, it reports on the implementation of a novel ptychographical technique, single-shot ptychography at X-ray FELs. Single-shot ptychography is capable of retrieving the probe and the extended sample from the data obtained from a single measurement without the need to scan the sample. Here, an X-ray-compatible setup for this technique and its first use at FLASH2 are described.

1.1 Outline

This work is structured as follows: Chapter 2 introduces the theoretical fundamentals of FELs. Then it briefly describes methods of the numerical propagation of radiation wavefields and gives a short overview of the coherence properties of radiation sources. Finally, it introduces the phase problem and gives the motivation for the CDI.

Chapter 3 introduces the ptychography technique and describes the various ptychography experimental setups and the main parameters of the ptychography exper-

iments. Additionally, it demonstrates the derivation of the classical ptychography reconstruction algorithm based on the hand-derived closed-form gradients and highlights its main disadvantages. Finally, it introduces an automatic differentiation as an alternative tool for creating efficient ptychographical algorithms.

Chapter 4 introduces automatic differentiation (AD) as a tool for creating a flexible ptychographical framework. It shows how an AD avoids the need for manually derived closed-form derivatives in gradient-based ptychography algorithms. Then, it discusses different components of the AD-based ptychography engine and the various error metrics and regularization strategies. Additionally, Chapter 4 presents details of the computational realization of the AD-based ptychography algorithms. At the end of the Chapter 4, an outlook on possible future developments of AD-based ptychography is presented.

Chapter 5 demonstrates how an AD-based flexible framework can be utilized to perform classical ptychography at SASE FELs. It discusses the main obstacles preventing the use of traditional ptychography engines at SASE FELs and shows how they were mitigated during the first ptychography measurement at FLASH. Both the design of the experimental setup as well as the corresponding forward model used for the reconstructions are discussed in detail. This chapter presents the use of ptychography for the simultaneous imaging and beam characterization at SASE FEL and how its results are utilized to determine the single-shot coherence degree of the FEL pulses. Results of a practical application of ptychography at FELs for imaging of plasma-treated thin plastic samples are also shown. The text of this Chapter closely follows the work of Kharitonov et al. [28] and Ravandeh et al. [29].

Chapter 6 describes the first implementation of single-shot ptychography at an FEL. It starts with discussing the setup for visible light proposed by Sidorenko et al. [1] and its adaptation for X-ray wavelengths. Then, it presents the single-shot AD-based ptychography algorithm. Finally, it reports the first single-shot ptychography experiment at an FEL performed at FLASH2 and the details of the data treatment and the reconstruction procedure. Finally, there is an outlook on the perspectives of the single-shot ptychography at FELs and possible future improvements. The text of this Chapter closely follows the work, currently available as a preprint at [30].

Finally, Chapter 7 summarizes this thesis and provides the outlook on the future

development of the classical and single-shot ptychography at FELs.

Chapter 2

Theoretical fundamentals

2.1 SASE FELs

X-ray FELs are capable of producing light pulses of extremely high brilliance significantly outperforming all of the synchrotron-based X-ray sources as shown in Figure 2.1. Moreover, these pulses are highly spatially-coherent and extremely short (down to 10 fs) and may be generated with an up to MHz repetition rate. These characteristics make FELs the tool of choice for imaging experiments. As at synchrotrons, the light pulses at the FELs are produced by the electrons accelerated to relativistic speeds and forced by alternating magnetic fields of the undulators onto the sinusoidal trajectory. However, at the FELs, the undulators are much longer then at synchrotrons (27 m and 30 m for FLASH1 and FLASH2 respectively [32], [33]), which allows the interaction between the electron beam and the photons emitted by it. These interactions cause modulations in the electron bunch density forming the so-called 'microbunches' moving in phase with the radiation. This way, the highly-correlated microbunches emit coherently and amplify the initial radiation through the positive feedback. This process is called the self-amplified spontaneous emission (SASE) and is the basis for SASE FELs. The SASE process allows obtaining ultra-bright and highly coherent pulses. However, due to its stochastic nature, each shot produces unique pulses resulting in significant shot-to-shot fluctuations of both the intensity and the wavefront. This section, gives some basic characteristics of the SASE FELs following [23], [34]and [35].

The schematic of the SASE generation is illustrated in Figure 2.2. The initially

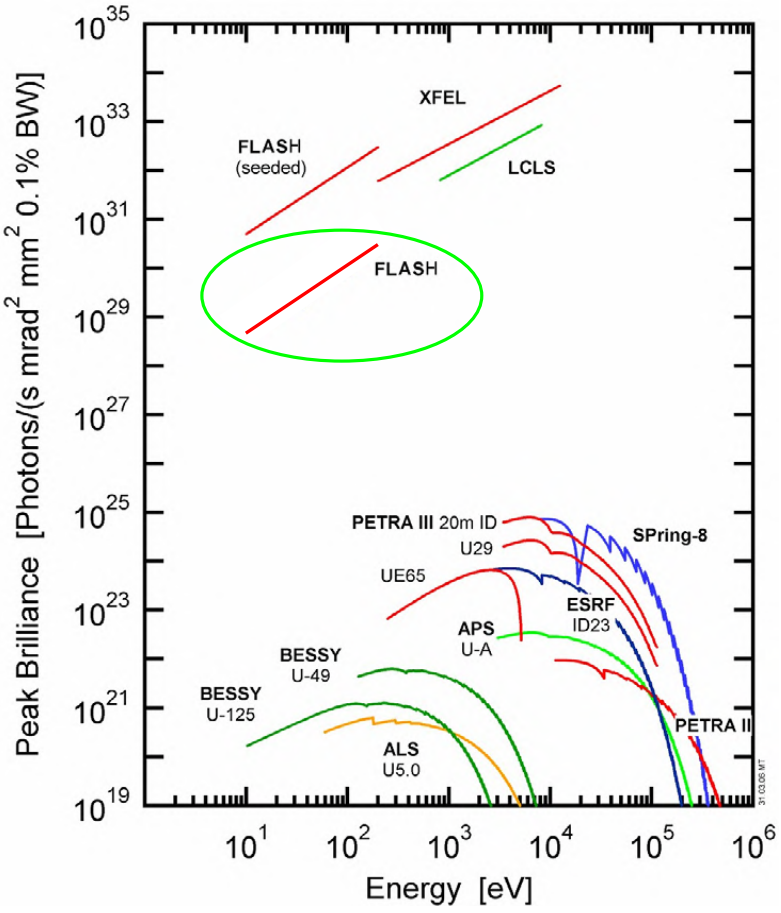

Figure 2.1: **Peak brilliance of storage rings and FEL sources.** FLASH highlighted with green ellipse. Adapted from [31].

homogeneous electron bunch moves through the planar undulators in the z direction. The undulator is composed of magnets with alternating directions of the magnetic field such that its strength in the y direction varies as follows:

$$\mathbb{B}_y(z) = \mathbb{B}_0 \cdot sin(\frac{2\pi}{\lambda_w} z), \tag{2.1}$$

where λ_w is the undulator period and the $\mathbb{B}_0$ is the maximum magnetic field determined by the undulator gap and the strengths of the magnets. While passing through the undulator along the z axis, the electron bunch will be forced on the sinusoidal trajectory by the Lorentz force. Due to the acceleration, electrons spontaneously emit radiation which in the laboratory frame of reference has a wavelength:

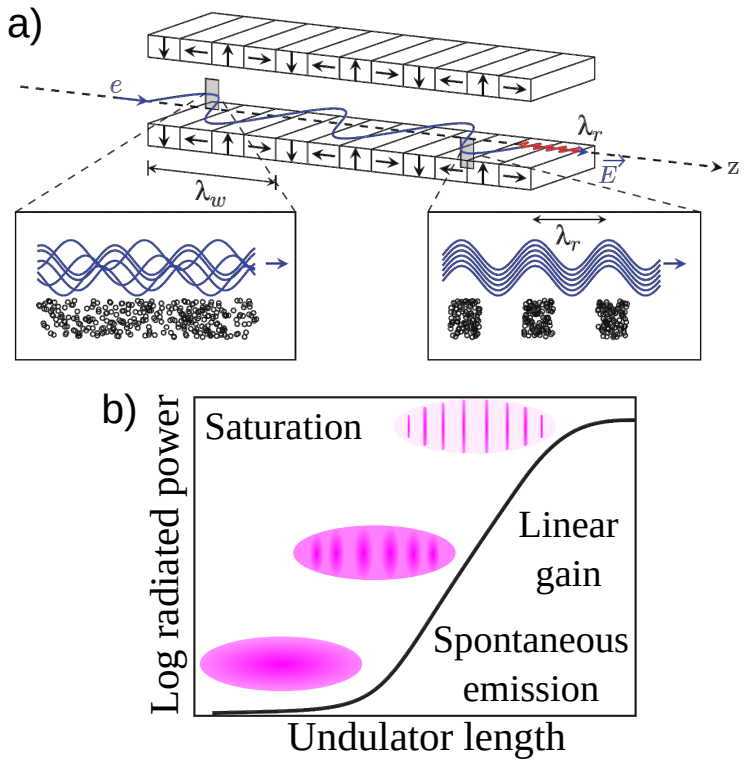

Figure 2.2: **(a) Undulator and microbunching.** The electron bunch passes through the undulator with an undulator period λ_w in the positive z direction emitting the radiation with a wavelength of λ_r. Insets show the spatial structure of the electron bunch and the corresponding radiation. *Left*: initially homogeneous electron bunch structure and corresponding incoherent emission. *Right*: fully spatially-modulated microbuches and coherent emission. **(b) Typical power gain curve for a SASE FEL**, as a function of the propagation through a length of the undulator. Initially, spatially unmodulated electron bunch corresponds to the linear growth of power (left-most part of the graph). Then, passing through the undulator length, 'microbunches' start to emerge, and the radiation power of the main mode exponentially increases. Finally, the full spatial modulation of the electron density is achieved, corresponding to the maximum radiated power. Adapted from [36] and [35].

$$\lambda_r = \frac{\lambda_w}{2\gamma^2}(1 + \frac{K^2}{2} + \gamma\theta^2), \tag{2.2}$$

observed in the stationary frame of reference, where γ is the Lorentz contraction factor $\gamma = 1/\sqrt{(1 - v^2/c^2)}$, $K = \frac{e\mathbb{B}_0\lambda_w}{2\pi m_e c}$ is the dimensionless undulator parameter, e is the elementary charge, m_e is the electron rest mass, θ is the radiation angle

measured from the z axis in the forward direction. The interaction between the radiation and the electrons results in the modulation of the electron density, causing the appearance of the microbunch structure with the inter-microbunch distance equal to the wavelength of the radiation λ_r. Due to the emergence of the microbunch structure, electrons start to emit more coherently, which leads to the amplification of the particular radiation mode (linear amplification in Figure 2.2(b). The amplification continues until the peak saturation power is limited by the efficiency of the energy transfer between the electrons and the electro-magnetic field achieved (the end of the undulator in Figure 2.2 (a) and the saturation in Figure 2.2 (b)). The peak power can be approximated as:

$$\mathbb{P}_{sat} = \rho_{FEL} \cdot \gamma m_e c^2 \cdot \frac{\mathbb{I}_e}{e} = \rho_{FEL} \cdot \mathbb{P}_e, \tag{2.3}$$

where $\mathbb{P}_e$ and $\mathbb{I}_e$ refer to the highest current and the power of the electron flow respectively, and ρ_{FEL} is the dimensionless characteristic FEL parameter describing the energy transfer efficiency between electrons and photons [34]. The spectral bandwidth at the saturation can be evaluated as $\Delta\omega/\omega \approx 2.35\rho_{FEL}$.

It should be further emphasized that the SASE process is essentially stochastic since the different modes are amplified at the different FEL shots. This causes significant shot-to-shot fluctuations both in the intensity and the wavefront of the SASE radiation. Additionally, around the saturation point, microbunching tends to have a noticeable harmonic content. Thus, the density modulations happen not only on the inter-microbunch distance of λ_r but also on harmonic distances $\lambda_n = \lambda_r/n,\ n \in \mathbb{N}$. Even harmonics are not observed directly on the undulator axis due to symmetry but can be observed at points located at an angle to it. At the same time, odd harmonics can be observed even directly on the axis. The intensity of harmonics is orders of magnitude lower than the intensity of the fundamental wavelength. However, they can also be used in the FEL experiments, provided that the fundamental wavelength is blocked by transmission filters[37].

All of the experiments described in this thesis were performed at the Free Electron Laser in Hamburg (FLASH) at the beamline FL24 with the typical parameters shown in the table 2.1 [33], [38].

In the future, it is planned to upgrade the FLASH1 and FLASH2 [39], [40] and implement the seeded mode of operation [41]. The main difference between a SASE

Parameter	**Value**
Electron energy E_e	$1.25\ GeV$
Wavelength λ_r	$4 - 90\ nm$
Spectral bandwidth at FWHM $\Delta\omega/\omega$	$10^{-2} - 10^{-3}$
Pulse Duration (FWHM) $\Delta\tau$	$10 - 200\ fs$
Photons/pulse	$3 \cdot 10^{12}$
Average Pulse Energy (single bunch)	$1 - 500\ \mu J$
Repetition rate	$1\ MHz$ @ $10\ Hz$

Table 2.1: Key parameters of the FLASH2.

and a seeded FEL is that at the seeded FELs, the lasing happens by amplification of the initial *seed* light pulse, which can be produced by an external laser or high harmonic generation source [41], or through the process of self-seeding [42]. The seeded FEL inherits the coherent and spatial properties of the *seed* [41] resulting in a significant decrease in beam fluctuations. Also, seeding dramatically increases both coherence and spectral brilliance of the light pulses and even allows to achieve statistically *laser-like* coherence properties [43]. The FLASH1 line will be seeded by the external laser [40], while the FlASH2 line could potentially obtain the harmonic lasing self-seeding capabilities [40], [42]. All imaging techniques described in this work would greatly benefit from the seeded mode of operation due to the increase in the total available photon flux, shot-to-shot beam stability, and coherence.

2.2 Light propagation

Maxwell's equations can fully describe electromagnetic radiation, its evolution, and its interaction with the medium, tracing the evolution of the electric field vector $\mathbb{E}(\boldsymbol{r}, t)$ and magnetic induction vector $\mathbb{B}(\boldsymbol{r}, t)$ through four coupled equations. However, the process of finding the exact solution for them is tedious and impractical for the purposes of CDI. Nevertheless, assuming the case of the propagation of light in the homogeneous medium and neglecting the polarization effects, it is possible to obtain a pair of uncoupled symmetrical equations describing the evolution of $\mathbb{E}$ and $\mathbb{B}$. This assumption is called *scalar approximation* and uses a single scalar equation to describe the evolution of the complex electromagnetic wavefield $\Psi(\boldsymbol{r}, t)$ [44], [45].

In all CDI experiments described in this work, the radiation is propagated in the high-vacuum, and no polarization effects were involved, thus fulfilling the scalar approximation. A detailed description of the optical wave theory can be found in [45]. The following section is based on [44].

Equations describing light propagation can be derived from the Huygens-Fresnel principle. According to it, every point of the wavefront itself is a source of the secondary spherical wave, and the interference of all the secondary waves forms the new wavefront [2]. Here, the equations are derived for the case of several parallel planes of interest in rectangular coordinates, as shown in Figure 2.3. The source

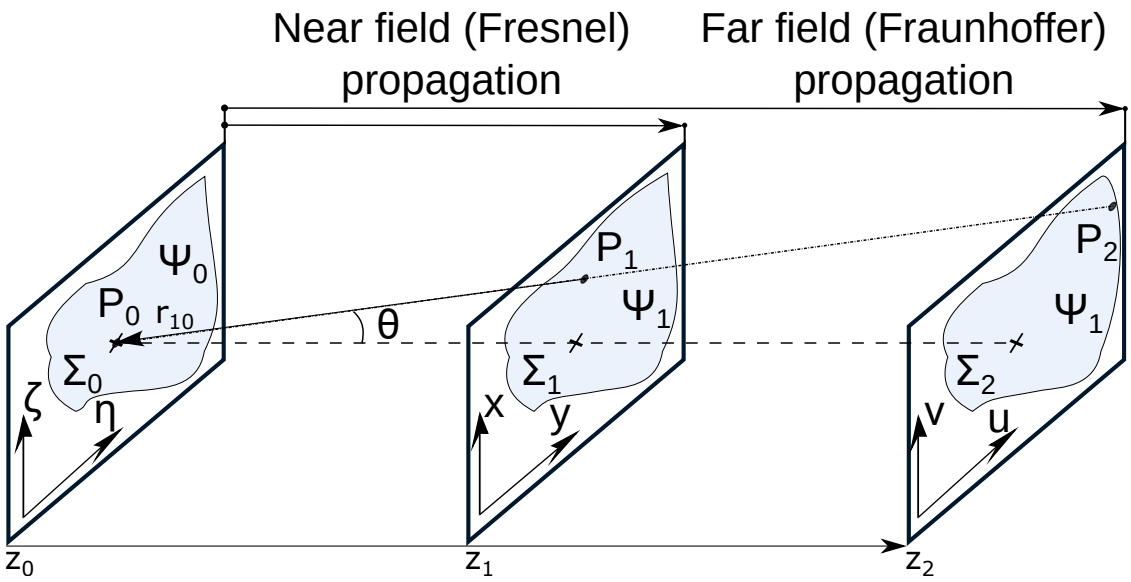

Figure 2.3: **Diffraction geometry for free-space propagation** The source plane with the wavefield Ψ_0 constrained in the area Σ_0 is located at z_0 and has the transverse coordinates (ζ, η). Each point P_0 of the Σ_0 generates a secondary wave. The propagation planes are located at distances z_1, z_2 in the near and far field respectively and have the associated transverse coordinates (x, y) and (v, u) respectively. (v, u) coordinate system is scaled as $(\frac{kx}{z_2}, \frac{ky}{z_2})$ due to the far field approximation. The resulting vawefields Ψ_1, Ψ_2 are constrained in Σ_1, Σ_2 in planes z_1, z_2 are produced by the sum of all the secondary waves emitted by Σ_0. Adapted from [44].

wavefield Ψ_0 is restricted in the area of Σ_0 in plane (ζ, η) and propagates in the positive z direction. The wavefields of interest are placed in the planes (x, y) and (v, u) parallel to the plane (ζ, η) and separated from it by the propagation distances

z_1 and z_2 respectively. The distances z_1, z_2 fulfill $\frac{D^2}{z_1\lambda} >> 1 >> \frac{D^2}{z_2\lambda}$, where D is the characteristic size of the source wavefield in the (ζ, η) plane, and λ is its the wavelength. The $F = \frac{D^2}{z_1\lambda}$ expression is called the Fresnel number [44] and can be used as a criterion to distinguish between the near- ($F >> 1$) and far- ($F << 1$) field cases.

According to the Huygens-Fresnel principle, the wavefield Ψ in the point P_1 in plane (x, y) can be found as

$$\Psi(P_1) = \frac{1}{i\lambda} \iint_{\Sigma} \Psi(P_0) \frac{\exp(ik\boldsymbol{r}_{10})}{\vec{r}_{10}} \cos(\theta)\ d\Sigma, \tag{2.4}$$

where $\Psi(P_0)$ is the wavefield in the point P_0, $k = \frac{2\pi}{\lambda}$ is the wavenumber, θ is the angle between the z axis and vector $\boldsymbol{r}_{10}$ pointing from the point P_1 to the P_0, and Σ is the area in which the wavefield Ψ is constrained with elemental area part $d\Sigma = d\zeta d\eta$ in the (ζ, η) coordinate system [44]. The $\cos\theta$ can be expressed as

$$\cos(\theta) = \frac{z_1}{\boldsymbol{r}_{10}}. \tag{2.5}$$

So, the Equation 2.4 can be rewritten as

$$\Psi(x, y) = \frac{z}{i\lambda} \iint_{\Sigma} \Psi(\zeta, \eta) \frac{\exp(ikr_{10})}{r_{10}^2}\ d\zeta d\eta, \tag{2.6}$$

with the distance r_{10} given as

$$r_{10} = \sqrt{z_1^2 + (x - \zeta)^2 + (y - \eta)^2)}. \tag{2.7}$$

The equation 2.6 allows calculating the wavefield in an arbitrary plane provided the field in the source plane. Nevertheless, it still requires a lot of computations (its computational complexity grows as N^4 for the propagation of the $N \times N$ points of the wavefield). Thus, it is undesirable for use in iterative calculations. However, additional assumptions on the propagation distance z and the point-to-point distance r_{10} will lead to further simplification and help with the calculations: the Fresnel approximation is suitable for the near- and intermediate-distance propagation, and the Fraunhofer approximation is suitable for the far-field propagation.

2.2.1 Fresnel approximation

Under Fresnel approximation the propagation distance z is assumed to be significantly larger than the size of the source wavefield. In the case of all the points P_0 being close enough to the optical axis to fulfill paraxial approximation:

$$\frac{|\zeta - x|}{z} \ll 1, \quad \frac{|\eta - y|}{z} \ll 1, \tag{2.8}$$

the Equation 2.7 can be simplified by using only the linear term in the binomial expression of the square root resulting in

$$r_{10} = z_1 \sqrt{1 + \left(\frac{x-\zeta}{z}\right)^2 + \left(\frac{y-\eta}{z}\right)^2} \approx z \cdot \left[1 + \frac{1}{2}\left(\frac{x-\zeta}{z}\right)^2 + \frac{1}{2}\left(\frac{y-\eta}{z}\right)^2\right]. \tag{2.9}$$

This expression can be used to substitute the r_{10} in the exponent term of the Equation 2.6 since the precision of the approximation there is more critical due to the high values of k ($\sim 10^7 - 10^{11} m^{-1}$) for the extreme ultraviolet (XUV) and X-ray ranges. For the denominator, all the terms except z_1 can be omitted, thus resulting in

$$\Psi(x,y) = \frac{\exp(ikz_1)}{i\lambda z_1} \iint_\Sigma \Psi(\zeta,\eta) \exp\left(\frac{ik\left[(x-\zeta)^2 + (y-\eta)^2\right]}{2z_1}\right) d\zeta d\eta. \tag{2.10}$$

Equation 2.10 can be seen as the convolution of the source wavefield $\Psi(\zeta,\eta)$ with the Fresnel kernel

$$h_{z_1}(x,y) = \frac{\exp(ikz)}{i\lambda z} \exp\left(\frac{ik\left[x^2 + y^2\right]}{2z_1}\right) \tag{2.11}$$

also called the impulse response (IR) of the empty space. The convolution can be efficiently calculated with the help of the Fourier transform using the convolution theorem. According to the theorem, a convolution in the real space is equivalent to the multiplication in the Fourier space, and thus the Equation 2.10 can be rewritten as

$$\Psi(x,y) = \mathcal{F}^{-1}\left\{\mathcal{F}\left\{\Psi(\zeta,\eta)\right\}\mathcal{F}\left\{h_{z_1}\right\}\right\} = \mathcal{F}^{-1}\left\{\mathcal{F}\left\{\Psi(\zeta,\eta)\right\}H_{z_1}\right\}, \tag{2.12}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the forward and inverse discrete Fourier transform (DFT) respectively. H_{z_1} is called the transfer function (TF) of the empty space and can be expressed as

$$H_{z_1} = \exp(ikz) \exp\left(-i\pi\lambda z_1 \left[q_x^2 + q_y^2\right]\right), \tag{2.13}$$

where q_x, q_y are reciprocal space coordinates.

Even though IR and TF propagators are analytically identical, their numerical implementations utilizing the DFT will give different results depending on the sampling in the source plane and propagation distance. Thus the sampling conditions should always be considered while planning a CDI experiment [46].

It is also possible to rewrite equation 2.10 by factoring the term $\exp\left(\frac{ik\left[x^2+y^2\right]}{2z_1}\right)$ from the integral to obtain the following propagation formulation [47]

$$\Psi(x,y) = \frac{\exp(ikz_1)}{i\lambda z_1} \exp(\frac{jk\left[x^2+y^2\right]}{2z_1}) \mathcal{F}\left\{\Psi(\zeta,\eta) \exp\left(\frac{ik\left[\zeta^2+\eta^2\right]}{2z_1}\right)\right\}. \tag{2.14}$$

This expression utilizes only one Fourier transform and thus makes the calculations more efficient. Additionally, it introduces the scaling relation between the source and target planes, allowing to obtain a higher resolution in the CDI experiment.

2.2.2 Fraunhofer approximation

In the case where the propagation distance z_2 is much greater than the extent of the wavefield in the source plane, the Fraunhofer approximation is satisfied. Thus, if the propagation distance z_2 is much greater than the source plane extent such that

$$z_2 \gg \max\left\{\frac{k\left[\zeta^2+\eta^2\right]}{2}\right\} \tag{2.15}$$

is satisfied, then the quadratic phase factor under the integral in the Equation 2.14 is approximately unity over the whole source wavefield Σ_0. Thus, the wavefield in the target plane can be found as

$$\Psi(x,y) = \frac{\exp(ikz_2) \exp(\frac{ik\left[x^2+y^2\right]}{2z_2})}{i\lambda z_2} \iint_{\Sigma_0} \Psi(\zeta,\eta) \exp\left(\frac{-ik\left[x\zeta+y\eta\right]}{\lambda z_2}\right) d\zeta d\eta. \tag{2.16}$$

The Equation 2.16 can be further simplified by replacing the variables from x, y to the reciprocal variables u, v such that

$$\begin{aligned} v &= kx/z_2, \\ u &= ky/z_2. \end{aligned} \tag{2.17}$$

After this replacement, the Equation 2.16 becomes the Fourier transform of the source wavefield multiplied with the phase term

$$\begin{aligned} \Psi(v,u) &= \frac{\exp(ikz_2)\exp(\frac{iz_2\left[v^2+u^2\right]}{2k})}{i\lambda z_2} \iint_{\Sigma_0} \Psi(\zeta,\eta)\exp -i\left[v\zeta+u\eta\right])\ d\zeta d\eta \\ &= \frac{\exp(ikz_2)\exp(\frac{iz_2\left[v^2+u^2\right]}{2k})}{i\lambda z_2} \cdot \mathcal{F}\left\{\Psi(\zeta,\eta)\right\}, \end{aligned} \tag{2.18}$$

which is the far-field Fraunhofer propagation. It is worth noticing, that in practical applications often only the intensity distribution $I(x,y) = |\Psi(x,y)|^2$ in the target plane is of interest, so the phase prefactors in Equations 2.14, 2.18 can be omitted.

All the propagators described in this section assume a fully-coherent wavefield Ψ. However, in real experiments, it often becomes necessary to use partially coherent radiation to increase the light flux or to characterize the spatial and coherent properties of the unperturbed beam. The next section discusses the coherent properties of light and shows how partially-coherent wavefields can be propagated using modal decomposition.

2.3 Coherence

Coherent diffraction imaging heavily relies on observing the intensity produced by the light scattered by the sample in general and the diffraction fringes in particular. XUV and X-rays have a high frequency due to their short wavelengths. Because of this, even the fastest detectors cannot directly measure the intensity produced by a single oscillation of a light wave. Thus, detectors effectively measure the intensity produced by the mean field averaged across the many periods of light-wave oscillations. The diffraction and interference phenomena are only detectable if dif-

ferent parts of the wavefield have a fixed phase difference in time and space since any changes in the relative phase, being averaged, blur the interference fringes. The stability of the phase relationship can be evaluated through the correlation of the different regions of the wavefield. This correlation can be calculated in two ways. First, it can be calculated for the different points of the wavefield at the fixed time, thus resulting in the spatial (transverse) characteristic of the coherence. Alternatively, the correlation can be calculated for the same point in space, considering only the time evolution of the wavefield, thus characterizing the temporal (longitudinal) coherence. This section first gives two quantitative definitions of the transverse ξ_t and longitudinal ξ_l coherence lengths closely following [2] and then derives the multi-modal representation of the partially coherent wavefields based on the general mutual coherence function definition following [48]. A detailed description of the coherence theory can also be found in [44], [45].

2.3.1 Spatial coherence

The spatial coherence evaluates the phase correlations in planes orthogonal to the direction of the light propagation at a fixed moment of time. The spatial coherence is determined by the angular size of the radiation source since, in the case of most light sources, individual electrons emit light more or less randomly. The light emitted by an incandescent lamp represents the most chaotic part of the spectrum, where all the acts of emission are totally uncorrelated between the individual electrons. In contrast, the laser represents the least chaotic and, thus, the theoretically fully coherent case with all the electrons emitting synchronously through the stimulated emission process.

The transverse coherence length is defined by considering the double-slit diffraction from the extended source as shown in Figure 2.4 a). Two slits separated by the distance d in the lateral plane are illuminated by the extended monochromatic light source D placed at a distance L upstream. Let us assume D to be a source emitting in a fully uncorrelated manner. In the paraxial approximation, the radiation emitted by each of the source points creates the cosine-squared intensity pattern on

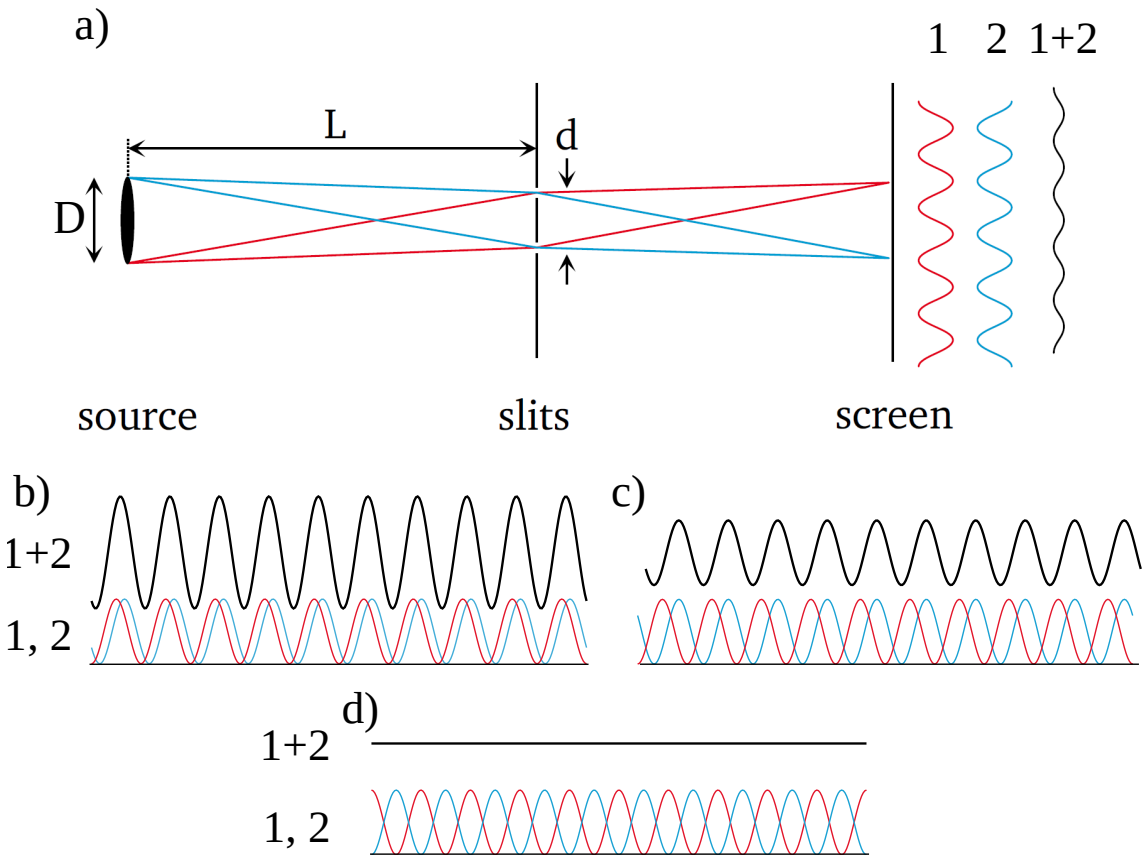

Figure 2.4: **Transverse coherence. (a)** Each point of a monochromatic, chaotic source D emits waves independent of all other points. Due to this, the waves are incoherent with each other. Because of the different geometry and direction of each wave, they each create their own diffraction pattern. Intensity distributions produced by the two limit points of the source D are shown in red and blue. With increasing source size D, the different patterns are transversely shifted as shown in **(b-d)**. Overall, this blurs the measured diffraction pattern and can possibly fully eliminate any fringes. Adapted from [49].

a screen with the maxima of the intensity visible at the angles

$$\theta_n = \frac{n\lambda}{d}. \tag{2.19}$$

Thus, the angular separation between the two neighboring fringes may be calculated

as

$$\delta\theta = \frac{\lambda}{d} \tag{2.20}$$

Let us consider two limit points of an extended source. As the source size D increases, the individual intensity patterns produced by them (shown in Figure 2.4 (b-d) with red and blue) will begin to shift, thus reducing the visibility of the fringes in total incoherent intensity distribution (shown in Figure 2.4 (b-d) black). This shift is determined by the source angular size

$$\alpha = \frac{D}{L}. \tag{2.21}$$

The transverse coherence length ξ_t in the plane of the slits is defined as the slit separation $d = \xi_t$ at which the angles defined by equations 2.20 and 2.21 are equal and thus the diffraction fringes are completely invisible as in Figure 2.4(d):

$$\frac{D}{L} = \frac{\lambda}{\xi_t} \;\Rightarrow\; \xi_t = \frac{\lambda L}{D}. \tag{2.22}$$

The coherence length ξ_t may be used as a qualitative parameter indicating the maximum size of the radiation beam, at which it is still possible to register diffraction fringes and thus utilize the radiation for CDI purposes. All the experiments described in this work were done at the beamline FL24 at FLASH2 with an approximate source size of $20-60\ \mu m$ (FWHM) at wavelength varying from 2.66 nm (3rd harmonic of 8 nm) to 13.5 nm respectively. The source was placed $\sim 80\ m$ in front of the entrance plane of the focusing optics [50], [51] resulting in $\xi_t \simeq 10-18\ mm$ at the entrance plane of the focusing optics. These approximate values are comparable with the entrance aperture of the focusing optics ($\sim 10mm$), which allows assuming the high spatial coherence of the FLASH2 radiation for experiments described in this work.

2.3.2 Temporal coherence

Another way to evaluate the coherence is by correlating the wavefields in the same point of space with different time delays. It is possible to evaluate the temporal

coherence and the so-called coherence time τ_l, a time during which the wavefields are correlated. The coherence time corresponds to the longitudinal coherent length - a distance traveled by the light during this time:

$$\xi_l = c\tau_l. \tag{2.23}$$

The expression for the longitudinal coherence distance can be derived using an example shown in the Figure 2.5. For the source of the spectral bandwidth of $\delta\lambda$

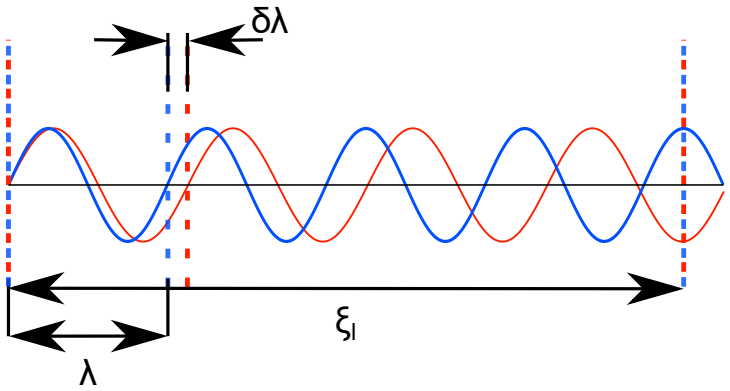

Figure 2.5: **Temporal coherence length.** Two waves of the wavelengths λ and $\lambda + \delta\lambda$ are out of phase after the longitudinal coherence distance ξ_l. Adapted from [35].

two limit waves with the almost similar wavelengths λ and $\lambda + \delta\lambda$ will be fully out of phase after the distance of ξ_l. This distance can be expressed as

$$\xi_l = N\lambda = (N - \frac{1}{2})(\lambda + \delta\lambda), \tag{2.24}$$

for the shortest and longest waves of the spectral interval $\delta\lambda$ respectively, since, the longest of the waves needs to oscillate exactly half a cycle less to travel the same distance. For the narrow-bandwidth source $\lambda \gg \delta\lambda$ and $(N - \frac{1}{2}) \approx N = \frac{\lambda}{2\delta\lambda}$. Therefore, the longitudinal coherence distance can be approximated as

$$\xi_l = \frac{\lambda^2}{2\delta\lambda}. \tag{2.25}$$

It is worth mentioning that a factor of $\frac{1}{2}$ appears here due to the initial assumption of the complete out-of-phase 180° shift and is frequently omitted for the quantitative

characterization of the temporal coherence length. The typical bandwidth of the FLASH2 pulses used for the experiments described in this work, $(\delta\lambda)/\lambda = 10^{-3} - 10^{-2}$ [33], results in a typical longitudinal coherent length $\xi_l \simeq 1.4\ \mu m$ for $\lambda = 13.5\ nm$. The longitudinal coherent length should be considered while planning the CDI experiment since it defines the maximal optical path difference between the two waves hitting the same detector pixel at which the interference is still detectable. This limits the best resolution achievable in the CDI experiment as

$$\delta x \geq \frac{D\delta\lambda}{2\lambda}, \tag{2.26}$$

where δx is the best achievable resolution in the sample plane, D is the beam size in the sample plane, and $\frac{\delta\lambda}{\lambda}$ is the spectral bandwidth of the source [52].

However, it is worth mentioning that ptychography relaxes both the spatial and temporal coherent requirements due to the inherent redundancy of the measured data allowing to perform the ptychography measurements with partially coherent radiation as described in Section 3.1.4. Thus, the ptychographical sampling condition (Equation 3.3) introduced in Section 3.1.2 becomes mode important while designing ptychographical experiments.

2.3.3 Degree of coherence and mode representation

In general, the coherence of an arbitrary wavefield $\Psi(\boldsymbol{r}, t))$ with respect to two points in space $\boldsymbol{r}_1, \boldsymbol{r}_2$ and the time interval τ can be analyzed by the mutual coherence function (MCF) given as

$$\Gamma(\boldsymbol{r_1}, \boldsymbol{r_2}, \tau) = \Gamma_{1,2}(\tau) = \langle \Psi(\boldsymbol{r}_1, t+\tau)\Psi^*(\boldsymbol{r}_2, t)\rangle\,, \tag{2.27}$$

where $\langle\cdot\rangle$ denotes the time average. The MCF is in general four-dimensional and can characterize both temporal and spatial properties of the wavefield. The Fourier-transform of the MCF $W = \mathcal{F}\{\Gamma\}$ is called a cross-spectral density (CSD):

$$W(\boldsymbol{r}_1, \boldsymbol{r}_2, \omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Gamma(\boldsymbol{r_1}, \boldsymbol{r_2}, \tau)\exp(i\omega\tau)\,d\tau. \tag{2.28}$$

The intensity I and spectral density $\mathbb{S}$ for the particular point $\boldsymbol{r}$ can be found as

$$\begin{aligned} I(\boldsymbol{r_1}) &= \Gamma_{1,1}(\tau = 0) \\ \mathbb{S}(\boldsymbol{r_1}, \omega) &= W_{1,1}(\omega). \end{aligned} \tag{2.29}$$

CSD can be normalized with the mean spectral densities at each point of space to obtain the spectral degree of coherence (SDC) μ :

$$\mu_{1,2}(\omega) = \frac{W_(\boldsymbol{r_1}, \boldsymbol{r_2}, \omega)}{\sqrt{\mathbb{S}(\boldsymbol{r_1}, \omega) S(\boldsymbol{r_2}, \omega)}}, \qquad 0 \leq |\mu_{12}(\omega)| \leq 1, \tag{2.30}$$

where the case of $|\mu_{12}(\omega)| = 0$ corresponds to the fully incoherent wavefield, while $|\mu_{12}(\omega)| = 1$ represents the full coherence. For practical applications, it is often more convenient to have a single number characterizing the overall coherence of the entire wavefield. For this purpose the SDC is frequently integrated over all the spatial coordinates to obtain the global degree of coherence ξ [53] as

$$\xi(\omega) = \frac{\int |W(\boldsymbol{r}_1, \boldsymbol{r}_2, \omega)|^2 \; d\boldsymbol{r}_1 d\boldsymbol{r}_2}{(\int \mathbb{S}(\boldsymbol{r}, \omega) \; d\boldsymbol{r})^2}, \quad 0 \leq \xi \leq 1. \tag{2.31}$$

According to Mandel et.al. [48] the CSD can always be represented as an infinite converging sum as

$$W(\boldsymbol{r}_1, \boldsymbol{r}_2, \omega) = \sum_n \beta_n(\omega) \Psi(\boldsymbol{r}_1, \omega) \Psi^*(\boldsymbol{r}_2, \omega). \tag{2.32}$$

This allows to rewrite the equation 2.29 for spectral density as

$$\mathbb{S}(\boldsymbol{r}, \omega) = \sum_n \beta_n(\omega) |\Psi_n(\boldsymbol{r})|, \tag{2.33}$$

where Ψ_n are orthonormal statistically independent coherent modes and β_n are corresponding modal weights. Using an orthonormality of the Ψ_n and substituting the equations 2.32 and 2.33 into the equation 2.31 it is possible to calculate the global degree of coherence $\xi(\omega)$ as proposed by Vartanyants et al. [53] as

$$\xi(\omega) = \frac{\sum_n \beta_n^2}{(\sum_n \beta_n)^2}. \tag{2.34}$$

equation 2.34 provides us a way to perform the characterization of the coherent properties of the wavefields buy applying the multi-modal expansion and analyzing modal weights β_n.

Another important consequence of the equation 2.32 is the possibility of using it to calculate the resulting intensity of partially coherent radiation at an arbitrary point in space. Applying the multi-modal expansion to equation 2.29, it becomes possible to express the intensity of the partially coherent wavefield at the point $\boldsymbol{r_1}$ as

$$I(\boldsymbol{r_1}) = \Gamma_{11}(\tau = 0) = \sum_n \beta_n \Psi_n(\boldsymbol{r_1}) \Psi_n^*(\boldsymbol{r_1}) = \beta_n \sum_n I_n. \tag{2.35}$$

equation 2.35 allows us to use the results obtained in the section 2.2 for the propagation of the partially coherent field. First, the partially coherent wavefield is represented as a sum of the mutually incoherent modes in the source plane. Then each of the modes is propagated individually to the target plane, where the total intensity of the partially coherent wavefield can be found as a sum of the intensities of the modes. This, in turn, opens the possibility for the use of partially-coherent radiation in CDI experiments to increase the amount of accessible flux. Additionally, it becomes possible to perform the characterization of the partially coherent wavefields both in terms of the spatial and coherent characteristics by measurement of the coherent modes Ψ_n and their respective modal weights β_n.

During the history of the FLASH, several attempts have been made to characterize the coherent properties of its radiation [24], [53]–[59]. These experiments relied on various experimental approaches such as speckle pattern analysis [56], double-slit diffraction [24], [53]–[55], Hanbury Brown–Twiss interferometry [58], [59] ,and Wigner distribution analysis [57]. Most of the measurements resulted in the values of the global degree of coherence ξ laying in the range of $0.4 - 0.8$ depending on the wavelength, beam energy, and exact experimental setup. These values agree with theoretical calculations presented in [60]. At the same time, their variation may be explained by the fact that their values strongly depend on the specific parameters of the FLASH accelerator and undulators that are changing from experiment to experiment. However, the Wigner distribution analysis resulted in the outlying results of 0.03 [57]. This discrepancy can be explained by the use of the intensity distribution averaged across several FLASH pulses without considering their spatial fluctuations and limited pointing stability. Moreover, in [57] data were analyzed

using the framework based on the ray-tracing, thus fully ignoring the diffraction effects. In this work, the novel multi-modal ptychography model with single-shot-unique modal weights was used to characterize the coherent properties of the single FLASH 2 pulses. The details about this model are given in Section 5.1.3 and its use for estimating the coherence properties of FLASH2 pulses is presented in Section 5.2.2.

2.4 Contrast forming in CDI

In general, the process of interaction between radiation and matter leading to the contrast forming in CDI is studied in the scattering theory. However, in most of the CDI experiments, the samples can be considered thin enough to fulfill

$$\Delta z < \frac{dx^2}{\lambda}, \tag{2.36}$$

where Δz is the thickness of the sample in the direction of light propagation, and dx is the numerical resolution in the sample plane. In this case, the projection approximation is valid, and volumetric scattering effects can be neglected, thus allowing for the substitution of the 3D description of the sample with a 2D transfer function. Details on the scattering theory can be found in [47].

For CDI applications, it is typically enough to use the refraction index to characterize the interactions between light and matter. For X-rays the refraction index may be represented as

$$n = 1 - \delta + i\beta. \tag{2.37}$$

β and δ are small positive numbers, where β characterizes the attenuation and $1-\delta$ characterizes the phase shift of the monochromatic beam in the sample as shown in Figure 2.6. The thin sample approximation assumes only single-scattering events are occurring when light propagates through the bulk of the sample and neglects diffraction inside the sample. Thus, the sample may be modeled by its transfer function as

$$O = \exp\left(-k\beta z_s\right) \cdot \exp\left(-ik\delta z_s\right), \tag{2.38}$$

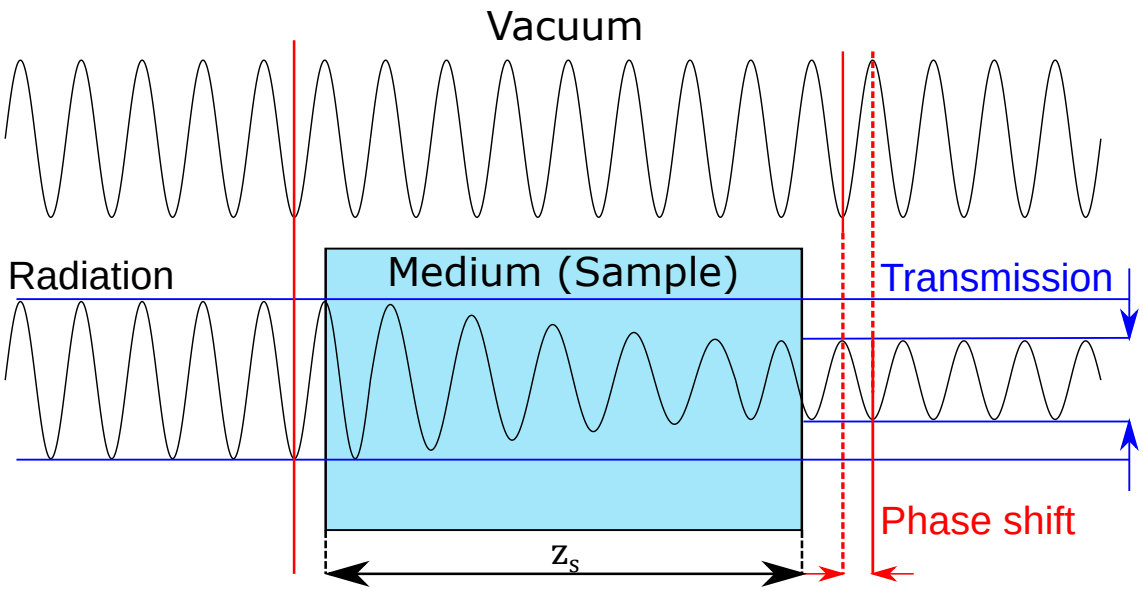

Figure 2.6: **Contrast forming mechanism behind the projection approximation.** X-ray wave passes through the vacuum without scattering (top). X-ray wave propagates through the medium characterized by its refractive index $n = 1-\delta+i\beta$ which causes scattering (bottom). The scattered wave is simultaneously attenuated and moves with higher phase velocity resulting in a phase shift and an amplitude change compared to the unscattered wave. Adapted from [61].

where $k = \frac{2\pi}{\lambda}$ is a wavenumber and z_s is the thickness of the sample. Under this assumption, the exit-wave produced by the interaction between the sample O and the wavefield of the incoming radiation P can be calculated as the product of the corresponding complex-valued functions

$$\Psi = O \cdot P. \tag{2.39}$$

Thus, a thin sample modulates the passing beam and causes diffraction, which can be used by CDI to retrieve the sample function O.

In ptychography experiments the sample may be assumed optically thin if the criterion proposed by Tsai et al. [62] holds:

$$z_s \leq \frac{5.2(\delta x)^2}{\lambda}, \tag{2.40}$$

where z_s is the thickness of the sample in the direction of the wave propagation, δx is the desired resolution, and λ is the radiation wavelength. However, it is sometimes necessary to obtain images and volumetric information from samples whose thickness is too large for the thin sample approximation to hold [63]. In such a case, the sample can be modeled as a series of the thin planes $O_i \dots O_n$ with individual transfer functions separated by empty space with the thickness δz. In this case, the exit-wave may be calculated by consecutive application of the corresponding transfer functions and the empty space propagators as follows

$$\Psi = O_n \cdot \mathcal{P}_{\delta z}(O_{n-1} \cdot \mathcal{P}_{\delta z}(\dots \cdot \mathcal{P}_{\delta z}(O_1 P))), \tag{2.41}$$

where $\mathcal{P}_z$ is the corresponding propagator.

In all the experiments described in this work, the samples were thin enough to fulfill the equation 2.40 and were reconstructed as the 2D sample functions $O_{(x,y)}$.

2.5 Phase problem

The inability to measure the phase of the radiation with detectors is known in the CDI community as a phase problem. The electromagnetic waves of the X-ray radiation oscillate faster than the measurement time of any detector. Thus, the measured intensity is proportional to the squared modulus of the wavefield

$$I = |\Psi|^2. \tag{2.42}$$

Even in the simplest case of the far-field diffraction $I = |\mathcal{F}(\Psi)|$, the only easily-recoverable quantity is the exit-wave autocorrelation: $\Psi * \Psi^* = \mathcal{F}^{-1}(I)$ [64]. The missing phase restoration and thus solving of the phase problem can be done using phase retrieval methods [5]. These methods rely on a priori information about the sample and probe, such as the size of the support, sample-detector distance, partially-known wavefield of the probe, etc., to perform the phasing and retrieve the correct exit-wave Ψ. In the general case, CDI techniques always retrieve only the exit-wave, which in the projection approximation is a product of the object and the probe functions $\Psi = PO$ and not the object transfer function O. The object

function O can be typically found assuming a known wavefield P as $O = \Psi/P$. However, it requires a priory knowledge of the probe wavefield and is not always easily achievable in real experiments, thus requiring additional beam characterization. Ptychography differs from most CDI techniques as it is able to simultaneously reconstruct probe and sample. Therefore, it can be efficiently used for both imaging and beam characterization [25], [65].

Chapter 3

Ptychography

3.1 Principles of ptychography

Ptychography [13] is a scanning coherent diffraction imaging technique based on the measurement of multiple partially overlapping diffraction patterns during the scanning of a photon beam over a sample. The principle of ptychography is illustrated in Figure 3.1. The acquired patterns are utilized in a phase retrieval procedure to reconstruct the complex probe wavefield and, at the same time, the object complex transmission function. Scanning is performed in such a way as it achieves a controlled degree of overlap between adjacent scan points. The overlap provides redundancy in the measured data and ensures the convergence stability of the reconstruction. Moreover, the overlapping areas result in additional continuity constraints in the sample plane, preventing reconstruction ambiguities typical for other CDI techniques. Additionally, ptychography reconstructs the probe and the sample functions individually, which allows using it as a tool for beam characterization [25].

The effect of the overlap constraint on the convergence of the ptychography reconstruction is illustrated in Figure 3.2. The full-field CDI utilizes individual diffraction patterns to perform the reconstructions of the sample function. Modulus and phase of the object reconstructed in such a manner are shown in Figure 3.2 (a,b), respectively. The presence of the typical ambiguities is visible in the reconstruction of the phase of the sample in the top-right corner of Figure 3.2 (b) (shown with green squares), where the head of the bird is reconstructed twice, representing the complex-conjugate twins. The presence of the overlap constraints is manifested

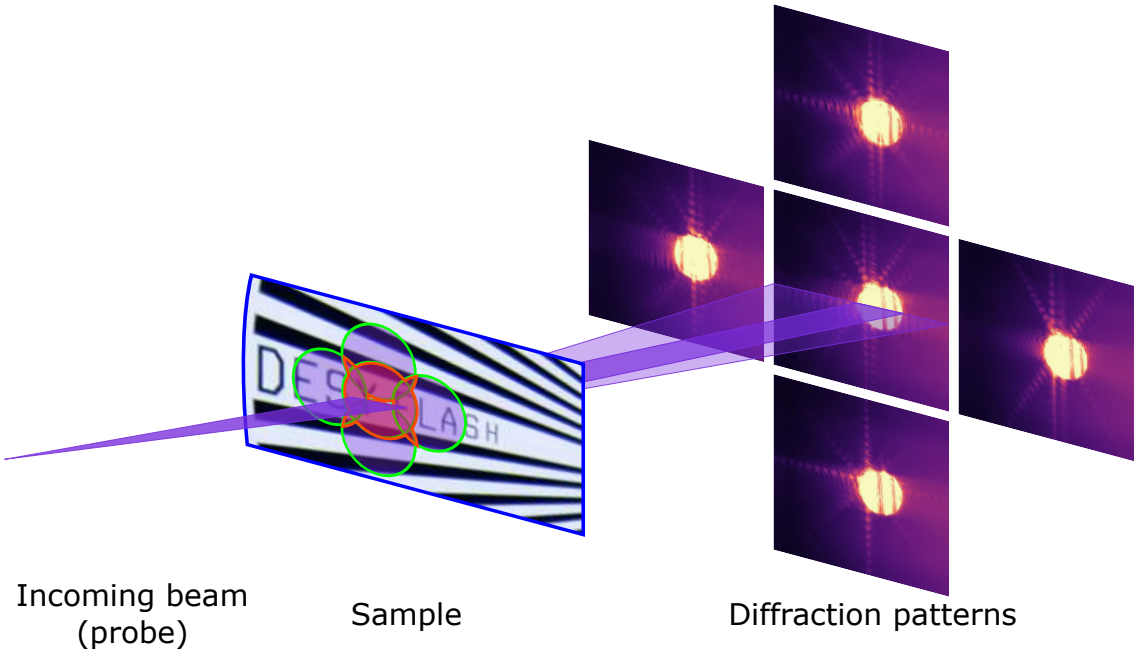

Figure 3.1: **Principle scheme of the ptychography.** Several diffraction patterns are measured from neighboring areas (shown with green circles) during the scanning of the sample. Scanning is done in such a way as to guarantee an overlap between adjacent areas (overlapped areas highlighted with orange). The overlap provides redundancy in the measured diffraction data, which enables simultaneous reconstruction of the complex transfer function of the object and the complex wavefield of the probe.

through the improvement of the resolution and the absence of the reconstruction artifacts in Figure 3.2 (c,d).

This section describes different ptychography techniques, including single-shot ptychography. Then, it presents the basic ptychographic formalism and reconstruction algorithm on the example of the extended ptychographic iterative engine (ePIE) algorithm [67]. Also, it introduces the basic ideas behind the scan position refinement and mixed-state ptychography. Additionally, it highlights the main limitations of the ptychographical algorithms based on closed-form gradients and names the improvements that should be introduced in the ptychographical forward model to enable the routine classical ptychography at SASE FELs. Details on the practical implementation of ptychography are given in [68]. An extensive overview of ptychography, in general, can be found in [69].

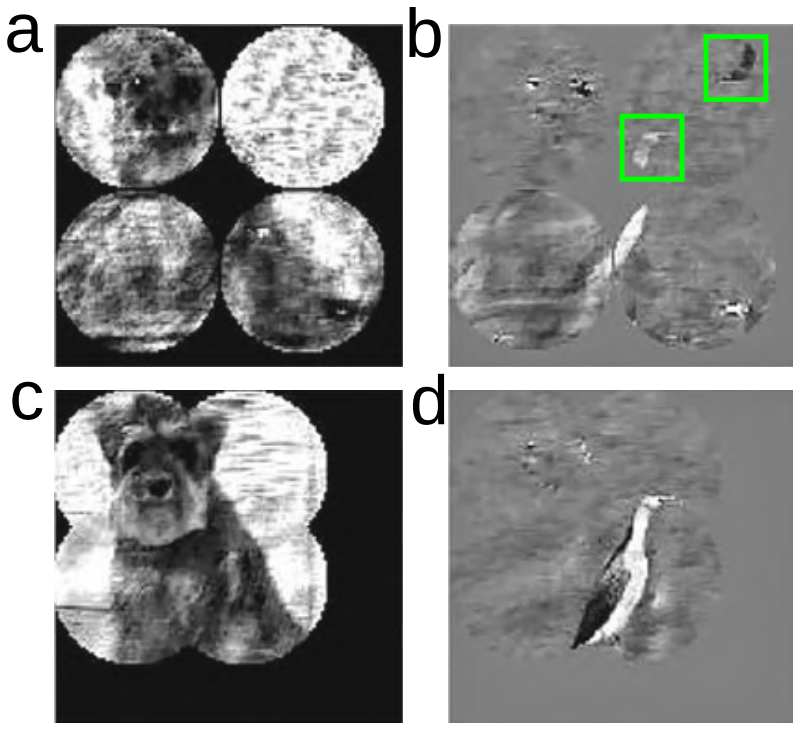

Figure 3.2: **Effect of the overlap in ptychography reconstruction.** When separate reconstructions are undertaken, each using a single diffraction pattern from four entirely different areas of a sample **(a,b)**, the usual ambiguities of full-field CDI arise. **(a)** and **(b)** represent modulus and phase of the sample transfer function, respectively. In **(b)** in the top-left circle, the bird appears twice in the lower-left and top-right parts. Two bird images (green squares in right corner)are reflected and have opposite phases, thus representing the complex-conjugate ambiguity. When the four calculations are undertaken simultaneously in a ptychograhy fashion with overlapping areas constrained to be identical (**(c)** and **(d)**, modulus and phase respectively), the reconstruction loses the ambiguities, and the resolution is significantly improved. Adapted from [66].

3.1.1 Owerview of ptychography techniques

Classical ptychography, briefly described before, gives rise to several imaging techniques, all sharing the common idea of scanning with overlap. Such scanning provides a redundant dataset for the later reconstruction and deconvolution of the probe and the sample. Various ptychographic techniques are briefly described here, and those used for the experiments described in this thesis are highlighted.

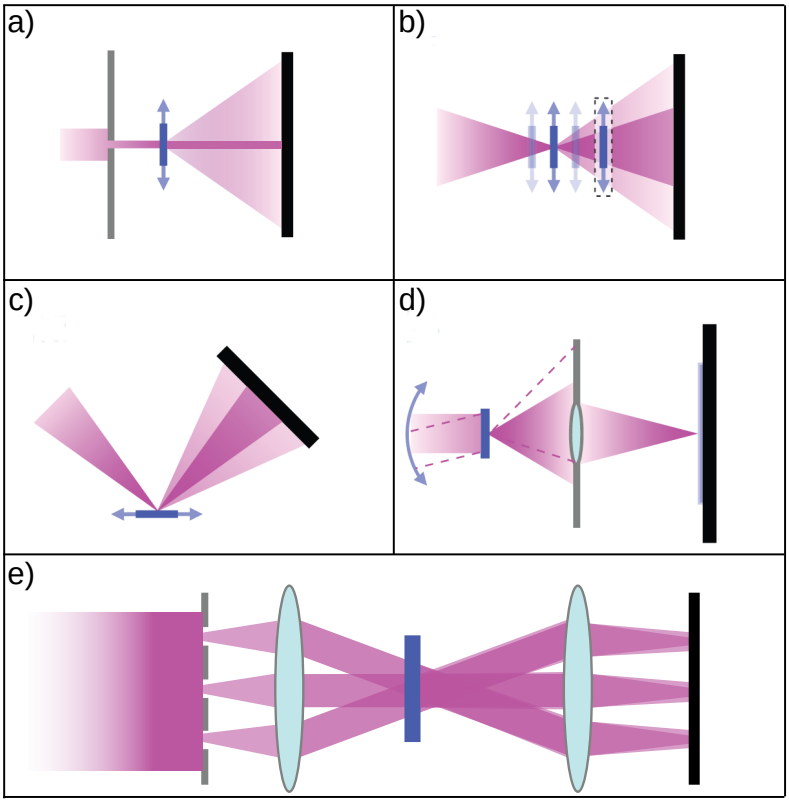

Figure 3.3: **Different ptychography setups** enable multiple kinds of ptychographical measurements. Arrows indicate scanning motion. The detector is shown in black. **(a)** In the simplest configuration, an aperture shapes the illuminating beam. **(b)** With a focused beam, the sample can be located at the focus or slightly upstream or downstream. Moving it tunes the size of the illuminating beam. For near-field ptychography, the sample is located significantly downstream of the beam focus (dashed box). **(c)** Reflection ptychography can use either specular or Bragg reflection. **(d)** In Fourier ptychography, a lens forms an image of the sample on the detector. A change of the illumination angle (dashed lines) changes the scattering angles accepted by the lens pupil. Capturing larger angles allows higher-resolution information to pass through the lens, enabling to simultaneously achieve a high resolution and large FOV with the same imaging optics. **(e)** Single-shot ptychography utilizes a 4-f system to perform single-shot illumination of the sample and combines the robustness and resolution of ptychography with the speed of full-field CDI. Adapted from [13]

The basic ptychography setup Figure 3.3 (a) assumes the use of an aperture to create a well-defined probe. However, this basic geometry has multiple disadvantages

[70]. First, the fixed size of the aperture limits the available coherent flux and prevents a flexible change of the beam size at the sample plane. Additionally, the use of the initially-plane wavefield shaped by the aperture creates a huge difference between the intensities of the scattered and un-scattered illuminations requiring a detector with a huge dynamical range or complex measurement strategies [71] to capture both the highly-scattered and the non-scattered photons.

A setup based on the use of the focused illumination Figure 3.3 (b) is the most widely used for classical ptychography experiments [68]. It allows to flexibly change the size of the illumination on the sample plane by moving the sample in and out of focus. Additionally, the presence of the known phase curvature in the probe wavefield improves the convergence of the reconstruction [72] and relaxes the requirements for the dynamical range of the detector by spreading the intensity more evenly. The limiting case of this type of setup is the near-field ptychography [73] – when the sample is placed so out-of-focus (Figure 3.3 (b) dashed rectangle) that the intensity at the detector is described by the near-field Fresnel propagation. The near-field geometry greatly increases the field of view with the price of a relatively low resolution determined by the geometric demagnification of the detector pixels.

Ptychography also can be performed utilizing a reflection geometry as shown in Figure 3.3 (c). In the case of X-rays, the crystalline object is oriented in the Bragg-reflection geometry to measure the separate Bragg peak [74]. An object is then scanned, making the smallest variation in the crystalline structure observable. This technique allows imaging of lattice dislocations and other defects and separating between the different phases of the crystalline objects. Specular reflections can also be used for ptychography measurements, giving information about surface profiles and reflective properties [75].

The Fourier ptychography setup is shown in Figure 3.3 (d). It utilizes the basic lens-based imaging setup to produce individual images with the resolution of each limited by the numerical aperture of the optical system [76]. However, in Fourier ptychography, the direction of the incoming radiation is changed, thus enabling the scanning of the object frequency spectrum in the lens pupil plane. This scanning allows the measurement of the higher frequencies of the object diffraction pattern in the lens plane, thus increasing the effective numerical aperture of the imaging system and the resolution of the final image of the sample. Thus, the Fourier ptychography

can take the images of objects with a simultaneously high resolution and large field of view. However, it is currently rarely used in the X-ray range due to the complex optical configuration [77].

A single-shot ptychography setup is shown in Figure 3.3 (e). This technique was initially proposed for visible light and utilizes a 4–f lens system to perform a single-shot illumination of an object with multiple beams simultaneously [1]. The incoming radiation is apertured by multiple pinholes, giving rise to a number of point-like sources placed in a common focal plane. This plane is placed in the back-focus of the first lens. The resulting beamlets are collimated by the first lens and cross in the common focus of the two lenses. The sample is placed in the vicinity of the cross-point such that the individual beamlets illuminate partially-overlapping areas of the sample plane, fulfilling the ptychography overlap requirement. The second lens focuses all the exit-waves on the detector, allowing the measurement of the far-field intensity distribution. This way, all the ptychographical datasets can be taken in one shot, which allows performing the time-resolved and pump-probe ptychography [78]. This technique, however, heavily relies on the use of refractive optics and was never demonstrated at X-rays before the experiment described in this thesis.

Chapter 5 describes the classical ptychography experiments performed at FLASH2 FL24 in the focused beam geometry as in Figure 3.3 (b). The novel X-ray compatible setup for single-shot ptychography and the first experiment performed at FLASH2 are described in Chapter chapter 6.

3.1.2 Overlap and sampling

Ptychography reconstructions heavily rely on the overlap between the neighboring beam positions and require some minimum density of scanning to guarantee a stable convergence of the reconstructions. Additionally, as for all CDI techniques, the sampling of the diffraction patterns in the detector plane also plays an important role and should be considered while planning the ptychography measurements.

The classical CDI sampling condition states that for full-field CDI, the phase

retrieval is possible when the following condition is fulfilled

$$\frac{\lambda z_{sd}}{2Dp_{det}} > 1, \tag{3.1}$$

where z_{sd} is the sample-detector distance, D is the beam size at the sample plane, and p_{det} is the detector pixel size. Equation 3.1 sets a limit on the maximum beam size D, and ensures that the size of the smallest registered diffraction speckle $\lambda z_{sd}/D$ exceeds two detector pixels. However, the sampling in ptychography is not only defined in the detector plane, the density of the scan in the real space (sample plane) should also be considered.

The first study of the influence of overlap degree on the convergence of the ptychography algorithms was done by Bunk et al. [79]. They first introduced the relative overlap degree parameter as

$$o = 1 - \frac{d}{D_{beam}}, \quad o \in [0...1], \tag{3.2}$$

where the probe is assumed to be a circle of a diameter D_{beam} and $d \in [0, 2r]$ is a center-to-center distance between the two neighboring areas illuminated during the scanning. The minimum threshold at which the ptychography still benefits from the overlap was estimated as 60%.

Two of these sampling conditions were combined in the work of da Silva et al. [80] leading to the unified ptychography sampling ratio given as

$$\frac{1}{\alpha_{samp}\beta_{samp}} > 1, \tag{3.3}$$

where $\alpha_{samp} = a/D$ is the ratio between the scanning step size a and the beam diameter D evaluating the sampling in real space, and $\beta_{samp} = p_{det}D_{beam}/\lambda Z_{sd}$ is the full-field CDI sampling condition evaluating the sampling of the reciprocal space. Equation 3.3 ensures the proper sampling in the joint 4D real-reciprocal space and was used as the criterion for the planning of all the experiments described in this work.

3.1.3 Position correction

Ptychography, being a scanning technique, is based on the a priori knowledge of the scan positions $\boldsymbol{r}_n$ and utilizes them for providing constraints in the sample plane. Ideally, scan positions should be determined with an accuracy better than the desired spatial resolution of the sample reconstruction, which is rarely achievable in experiments. Furthermore, uncertainties in the measurement of the scan positions are inevitable because of vibrations, thermal drifts, and scanning stage imperfections. However, dense scanning (more than 60% overlap calculated as in Equation 3.2) of the sample and resulting redundancy in the measured data enables the position correction during the regular ptychographical reconstruction.

The first position correction routine for ptychography was proposed by Guizar-Secairos, and Fienup [81]. They used the conjugate gradient method and relied on hand-derived gradients, which was quite computationally intensive. Later, several more computationally effective methods based on annealing algorithms [20] or cross-correlation [82] were proposed. Both of these techniques initially perform the reconstruction using assumed scan positions and then use additional calculations to refine the individual displacement vectors $\boldsymbol{r}_n$. The refined vectors are later used for consecutive reconstructions, thus iteratively refining the scan positions. However, these algorithms are still increasing the computational cost of the whole reconstruction by a factor of ~ 3 and relying on computations outside of the main reconstruction loop [66]. That is why they cannot be efficiently parallelized.

The ptychography reconstructions described in this work were done using a position refinement algorithm based on a differentiable affine transformation which we developed for the work published in [28]. The algorithm was implemented into the forward ptychographical model. Details on the development and performance of our position refinement approach can be found in the subsection 5.1.2.

3.1.4 Ptychography with multimodal decomposition

Although initially ptychography was developed under the assumption of a fully-coherent shot-to-shot constant beam, actual ptychography experiments may not fully meet these requirements. A probe may not be fully spatially and temporally

coherent, and thus its intensity will be effectively the incoherent sum of the multiple coherent intensities. However, this is not the only source of incoherence in an experiment. Vibrations of the sample or any parts of the optical system happening during the single image acquisition also degrade the coherence. It is worth noticing that this problem is most acute at synchrotrons and table-top sources due to the relatively low flux resulting in longer (several s) time of single image acquisition. Whereas ultra-bright pulses of the FELs allow the single image acquisition from the single fs-long shot of the FEL, thus eliminating the effects of optical system vibrations on the visibility of the diffraction fringes. Additionally, any imperfections in the detector point spread function blur the measured diffraction pattern and thus decrease the visibility of the diffraction speckles. These effects result in degradation of the reconstructed probe and sample and cannot be explained by the classical ptychography model described in subsection 3.1.5.

A solution to this problem was first proposed by Thibault and Menzel in [83]. They described ptychography in terms of measurements of state mixtures. In this formalism, all the sources of incoherence result in the simultaneous existence of the multiple probe and sample modes $P_i, i \in 1...n$ and $O_j, j \in 1...m$ respectively. These modes describe all the phenomena leading to the incoherence and interact with one another, resulting in the simultaneous existence of $n \times m$ exit-waves $\Psi_{ij} = P_i O_j$. During the measurements, each of the exit-waves produces a corresponding intensity I_{ij}. These intensities are incoherently summed by the detector resulting in the measured intensity given as

$$I = \sum_{i,j}^{n.m} I_{ij} = \sum_{i,j}^{n.m} \left| \mathcal{P}(P_i O_j) \right|^2 , \tag{3.4}$$

where $\mathcal{P}$ is the propagator from the sample plane to the detector plane.

This type of analysis allows performing the ptychography reconstructions with both, partially-coherent probes [84] and multiple states simultaneously present in a sample [85], by modification of the basic ptychography algorithm described in a subsection 3.1.5. However, this formalism is still assuming constant probe modes P_i simultaneously present at all scan positions $\boldsymbol{r}_n$ and thus cannot be straightforwardly applied to the shot-to-shot fluctuating pulses of SASE FELs. A modified version of this formalism described in section 5.1 was used in this work. This approach is

capable of achieving meaningful and consistent results at SASE FELs by representing the probe as several coherent modes with individual shot-to-shot unique modal weights.

3.1.5 Classical ptychography formalism

In classical ptychography, a sample O is scanned by a probe P in multiple positions characterized by the displacement vectors $\mathbf{r}_j, j \in 1..N$ where N is the number of scan points. Under the thin sample approximation [47], any secondary scattering and propagation of the field in the thickness of the sample are ignored. Thus, the exit-wave Ψ_j produced at the $j-$th position may be found by multiplying the object complex transfer function with the complex wavefield of the incoming probe as $\Psi_j = P \cdot O_{\mathbf{r}_j}$. The intensity produced by the exit-wave at the detector plane in the case of far-field measurements, may be calculated as

$$I_j = \left|\mathcal{F}\{\Psi_j\}\right|^2, \tag{3.5}$$

where $\mathcal{F}$ is the Fourier transform as described in section 2.2.

During the ptychography reconstruction, multiple intensity patterns are used to perform phasing and to reconstruct the complex-valued P and O functions. This task can be viewed as an optimization task of finding a pair of P and O that minimizes an error function $\mathcal{E}(I_j, \tilde{I}_j)$ evaluating the similarity between I_j (measured) and $\tilde{I}_j$ (estimated) intensity at the j-th scan position. The basic loss function $\mathcal{L}$ evaluating all the measured intensities may be formulated as a quadratic error function in the form of

$$\mathcal{L} = \sum_{j}^{N} (\sqrt{I_j} - \sqrt{\tilde{I}_j})^2. \tag{3.6}$$

A loss function of this form was proposed by Thibault and Guizar-Sicairos [86] and assumes Poisson statistics of the detector counting noise. The square root equalizes the influence of the shot noise in the bright and dark detector pixels on the value of the loss function. Further information on different noise models for ptychography and their corresponding loss functions can be found in Godard et al. [87].

The ptychography phase retrieval problem is nonlinear and does not have a simple analytical solution. Therefore, classical ptychography algorithms such as ePIE propose to solve the problem of minimization of $\mathcal{L}$ and thus reconstructing P and O through numerical gradient-based methods. For this, the gradients of the loss function with respect to the probe and sample functions $\frac{\partial \mathcal{L}}{\partial P}, \frac{\partial \mathcal{L}}{\partial O}$ should be found utilizing complex Wirtinger derivatives [88]. Omitting the intermediate derivations (examples of them may be found in [86], [89]), the gradients can be expressed as:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial O} &= \sum_j^N P_n \cdot \left[\mathcal{F}^{-1} \left\{ (1 - \frac{\sqrt{I_j}}{\sqrt{\tilde{I}_j}}) \cdot \mathcal{F}\{ \Psi_j \} \right\} \right]^* \\ \frac{\partial \mathcal{L}}{\partial P} &= \sum_j^N O_n \cdot \left[\mathcal{F}^{-1} \left\{ (1 - \sqrt{\frac{I_j}{\tilde{I}_j}}) \cdot \mathcal{F}\{ \Psi_j \} \right\} \right]^*, \end{aligned} \tag{3.7}$$

where $\cdot^*$ denotes the complex conjugate.

The ePIE algorithm utilizes the gradients from equation 3.7 to perform the minimization of $\mathcal{L}$ using the stochastic gradient descent. During a single step of the reconstruction, first, the gradients of the $\mathcal{L}$ with respect to P and O are calculated for a randomly chosen scan position and corresponding diffraction pattern I_j. Then, a step in the direction of the minimum (opposite to the gradient) of $\mathcal{L}$ is performed. The full global iteration is complete when all scan positions are visited once.

The updates for P and O at the n-th scan position can be calculated using the gradients from equation 3.7 as

$$\begin{aligned} O' &= O + \alpha_s \cdot \frac{- \left[\mathcal{F}^{-1} \left\{ (1 - \frac{\sqrt{I_j}}{\sqrt{\tilde{I}_j}}) \cdot \mathcal{F}\{ \Psi_j \} \right\} \right]}{\max |P|^2} \cdot P^* \\ P' &= P + \beta_s \cdot \frac{- \left[\mathcal{F}^{-1} \left\{ (1 - \frac{\sqrt{I_j}}{\sqrt{\tilde{I}_j}}) \cdot \mathcal{F}\{ \Psi_j \} \right\} \right]}{\max |O|^2} \cdot O^*, \end{aligned} \tag{3.8}$$

where α_s and β_s are the numerical factors controlling the gradient step size and thus the update strength. The normalization by the maximum of the squared modulus further stabilizes the algorithm and scales the update strength of the object and the probe to the same order of magnitude, since $|O| \in [0...1]$, while $|P|$ is proportional to the beam intensity measured without the sample and maybe several orders of

magnitude higher than $|O|$.

ePIE and all other ptychography algorithms based on closed-form gradients share one common drawback, which can be understood by the example of equations 3.5 - 3.8. The derivation of the closed-form gradient equations is a tedious process, especially for the more complex ptychography models. Moreover, any slight change in the ptychography reconstruction routine requires full re-derivation of the gradients and thus becomes undesirable if not impossible due to the lack of obvious closed-form solutions. This limits the possibilities of a flexible modification of the reconstruction algorithms in case of changes in the experimental setup, noise model, optimizable variables, etc. Therefore, an alternative approach to symbolic differentiation is very desirable and potentially may greatly expand existing experimental capabilities of ptychography. One alternative to symbolic differentiation is automatic differentiation (AD) – a set of techniques capable of providing the gradients of the function using its computational description [90]. AD removes the need for analytical re-derivation of the gradients and thus permits the creation of more complex ptychography algorithms. A ptychographical framework based on automatic differentiation was developed during this work and was used for the ptychographical reconstructions described here.

3.2 A brief description of automatic differentiation

Automatic differentiation (AD) is a computational approach capable of calculating the analytical values of the function derivatives when provided with the computable expression of the function itself [91]. This approach utilizes the fact that any arbitrary complex computable function may be expressed as a composition of some elementary functions with the known analytical gradients. Thus, the differentiation of such a composition can always be done through the recursive application of the chain rule to the same set of elementary functions with different input parameters. This process can be efficiently implemented in software and thus does not require the hand-derivation of the gradients.

This section briefly describes *reverse-mode* AD. Additional details on the com-

putational implementation of AD can be found in [91], while the key ideas behind the application of AD for phase retrieval, in general, are presented in [92].

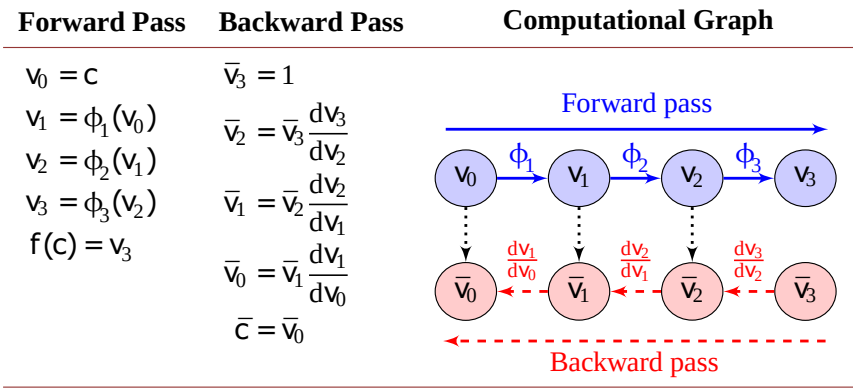

Figure 3.4: **Illustration of reverse-mode AD.** The gradient of the composite function is calculated by storing the forward computation graph and later application of the chain rule to the results of the intermediate calculations. Adapted from [93].

The main idea behind reverse-mode AD is illustrated in Figure 3.4. The function $f(c) = v_3$ describes a forward model depending on the optimizable parameter c. For the computation, any computable $f(c)$ can be represented as a composition of several elementary functions (addition, multiplication, trigonometric, logarithmic, etc.) $\phi_n(v_{n-1})$ as

$$f(c) = \phi_3(\phi_2(\phi_1(c))) = \phi_3(\phi_2(v_1)) = ... = \phi_3(v_2), \tag{3.9}$$

where v_n denote the result of the intermediate function ϕ_n. Then, the gradients of f with respect to c in point c_0 can be calculated by the application of the chain rule to the results of the intermediate functions as

$$\left.\frac{\partial f}{\partial c}\right|_{c=c_0} = \left.\frac{\partial v_3}{\partial v_2} \cdot \frac{\partial v_2}{\partial v_1} \cdot \frac{\partial v_1}{\partial c}\right|_{c=c_0}. \tag{3.10}$$

Analogous calculations are possible for any computable function representable as a composition of elementary functions. Additionally, according to the so-called 'cheap gradient principle,' the typical computational costs of the gradients of an arbitrary function of N operations are always no more than kN, $0 < k < 6$ with average values of k typically being lower than 3 [94]. This makes reverse-mode AD

a computationally efficient tool for performing AD-based iterative ptychographical reconstructions.

The main advantage of using AD for ptychographical reconstructions is that it allows to divide the monolith ptychographical algorithm into three separate components: forward model, loss function, and optimization routine. The forward model describes the experimental setup and approximates the measured intensity patterns based on optimizable input parameters. The loss function evaluates the similarity between approximated and measured intensities. The optimization routine can be freely selected from various gradient-based optimization techniques to ensure the best reconstruction results in the particular experiment. Due to this separation, AD ensures additional flexibility in the ptychographic routine. Each of its components can be changed independently and without the need for tedious manual calculations. This permits easy changes in the experimental setup and the forward model of the experiment. Eventually, AD allows one to concentrate on the physics behind the forward model describing the novel experiment without caring too much about the mathematics of its gradients.

An AD-based ptychography framework was used for all the reconstructions described in this thesis. The details about the implementation of the various components of the ptychographical reconstruction are given in chapter 4. The differentiable forward models describing the light propagation for the classical and the single-shot ptychography at SASE FELs are described in chapter 5 and chapter 6 respectively. The details on the implementation of the AD-based position refinement can be found in subsection 5.1.2, while the example of how the AD-powered ptychography allows direct reconstruction of the thickness of the sample instead of its transfer function is given in section 5.4.

Chapter 4

AD-based ptychography

Ptychography reconstruction can be seen as a problem of finding the values of the probe and the sample that minimize the loss function $\mathcal{L}(I_j, \tilde{I}_j)$ which compares the measured I_j intensity and the approximated intensity $\tilde{I}_j$ calculated by the forward model. To solve this problem, some ptychography algorithms rely on so-called projection-based methods [95]. These algorithms consecutively apply the sets of constraints expressed in the form of the projection functions [96] to probe samples and other optimizable parameters. These algorithms tend to be very computationally effective and can successfully evade the local minima during the optimization procedure. However, they tend to be hard to modify for challenging experimental conditions since the optimization of certain experimental parameters (for example, scan position correction) is hard to implement in the form of constraints and projectors. For this reason, in this work, we selected the gradient-based ptychography algorithm as a tool for creating a flexible ptychography engine capable of working at SASE FELs. The extensive review of the projection-based phase retrieval and ptychography algorithms can be found in [95], [97], [98] Other classical ptychography algorithms utilize hand-derived gradients and gradient-based optimizers to solve this problem. However, this approach makes any changes in the forward model tedious and undesirable, thus crippling the introduction of new forward models, experimental setups, and reconstruction approaches. AD evades this problem, thus permitting the development ofa novel class of ptychography reconstruction algorithms.

This chapter introduces the AD-based flexible ptychography framework developed within this thesis. First, this chapter explains the general flow of the AD-based

ptychography reconstruction and its computational details. After that, there is a detailed description of the individual components of the AD-based ptychography engine.

4.1 Computational implementation

The general flow of a single iteration of an AD-powered reconstruction is illustrated in Figure 4.1 and can be conceptually separated into three steps. First, the ap-

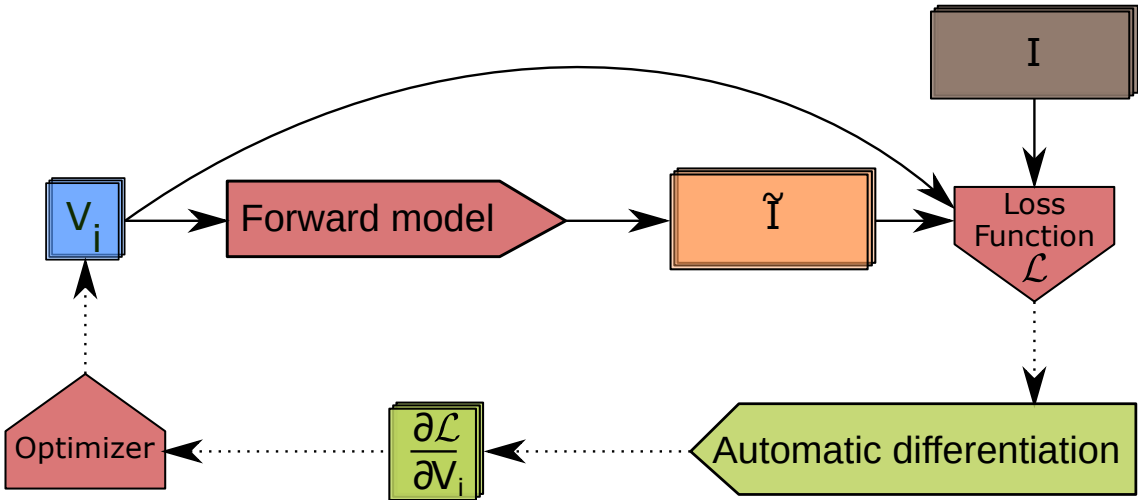

Figure 4.1: **One iteration of the AD-powered ptychography.** First, the approximated intensity $\tilde{I}$ is calculated by the forward model from the optimizable parameters V_i. Then, the loss function evaluates the quality of the approximation by comparing the measured I and approximated intensities $\tilde{I}$. The loss function is additionally regularized by calculating the various metrics from the optimizable parameters alone. The gradients of the loss function with respect to the optimizable parameters are estimated using AD and passed to the gradient-based optimizer. Finally, the optimizer changes the values of the optimizable parameters to minimize the loss function.

proximated intensity $\tilde{I}$ is calculated by the forward model from all the optimizable parameters V_i. The forward model encapsulates all assumptions about the experimental setup and performs the forward calculations differentiably, allowing the AD to track the gradients to the input parameters. Then, the quality of the approximation is evaluated by a loss function $\mathcal{L} = \mathcal{L}(I, \tilde{I}) + \sum_i \alpha_i \mathcal{R}_i(V_i)$. The loss function

is composed of two terms: the error function $\mathcal{E}(I, \tilde{I})$ evaluating the quality of the approximation, and the regularizers $\mathcal{R}_i(V_i)$ further constraining the reconstruction problem. Finally, the gradients of the loss function with respect to all the optimizable parameters $\frac{\partial \mathcal{L}}{\partial V_i}$ provided by the AD are given to an optimizer. The optimizer changes the values of the optimizable parameters to minimize the loss function and thus approaches the correct values of the sample probe and other parameters.

During a single iteration, the forward calculations may be performed for the number of measured scan positions determined by the batch size N_{batch}. The batch size can be selected arbitrarily varying from $N_{batch} = 1$ to the number of all the measured scan positions $N_{batch} = N_{measured}$. A batch size equal to one, $N_{batch} = 1$, is equivalent to the simplest realization of the ePIE algorithm [67]. All the positions are visited consecutively, with probe, sample, and other optimizable parameters being updated after each visited position. This makes the optimization process unstable and prevents effective parallelization. Increasing the batch size N_{batch} leads to the averaging of the gradients from several scan positions, thus resulting in faster and more stable convergence. Additionally, it allows for parallelizing the reconstruction process and utilizes the computational powers of the GPU more efficiently. The maximum batch size is limited by the available GPU memory since the amount of the gradient data stored grows proportionally to N_{batch}.

However, it is possible to further decrease the computational time by performing the calculations in parallel on multiple GPUs, as shown in Figure 4.2. In this approach, called *data parallelism* the values of the optimizable parameters and all the measured intensities are distributed between several GPUs such that the calculations for the individual scan positions can be done independently. The global gradient values are then obtained as the average of the local gradients calculated by the individual GPUs. The averaging procedure and further update of the optimizable values are performed by a single 'leading' GPU.

The ptychography engine described in this work is written using the PyTorch framework [99]. PyTorch allows to code the GPU-accelerated [100] AD-compatible calculations in Python [101]. Additionally, it provides an implementation of state-of-the-art gradient-based optimizers, thus enabling a fast and easy ptychography software development.

All the reconstructions described in this work were performed on the DESY

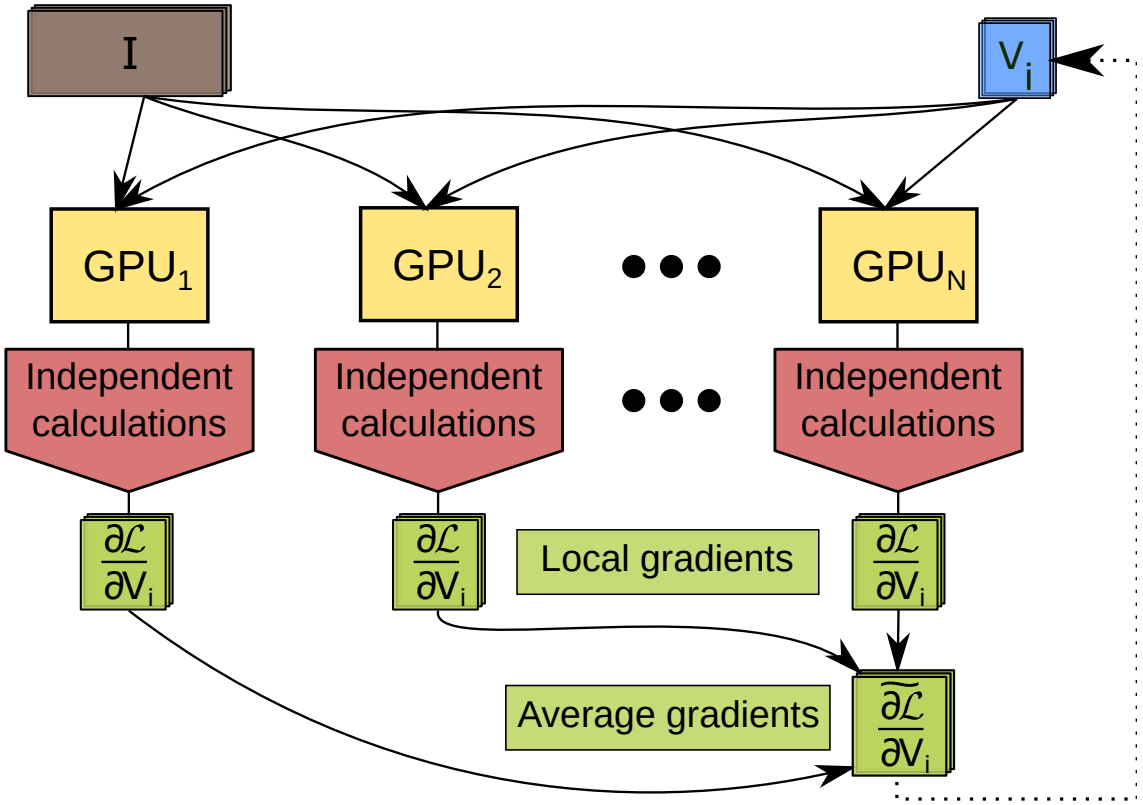

Figure 4.2: **One iteration of the multi-GPU reconstruction.** The measured intensities I are split equally between several GPUs. The current values of the optimizable parameters V_i are copied across several GPUs, with all calculations performed independently and in parallel. The local gradient values are averaged on a single 'leading' GPU and are used to update the optimizable parameters.

Maxwell cluster [102] and utilized up to the 4 GPUs in the data-parallel mode to maximize computational efficiency. The maximum possible batch size N_{batch} was used for each reconstruction to speed up the reconstruction convergence.

4.2 Forward model and optimizable parameters

The forward model aims to approximate the measured intensity distribution from the optimizable parameters in a differentiable way. It does so by describing the light propagation, its interaction with the sample, and other processes such as scanning

of the sample. All the optimizable parameters included in the forward model are described here.

- **Sample:** The sample can be represented either by its complex transfer function such that the effective modulation of the incoming wavefield Ψ_{in} is expressed as

$$\Psi_{out} = \Psi_{in} \cdot \left(\Re(O_{(x,y)}) + i\Im(O_{(x,y)})\right), \tag{4.1}$$

  where $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts of the object transfer function respectively. However, this representation might be impractical for strong phase samples, since a phase variation of more than 2π will result in phase wrapping and would require an additional unwrapping step after the reconstruction. In this case an alternative representation of the sample transfer function based on the equation 2.38 is more beneficial and allows the direct reconstruction of the effective refraction index components $\beta_{eff_{(x,y)}} = \beta z_{s_{(x,y)}}$, and $\delta_{eff_{(x,y)}} = \delta z_{s_{(x,y)}}$, where $z_{s_{(x,y)}}$ is the sample thickness distribution. In this representation, the probe modulation is calculated as

$$\Psi_{out} = \Psi_{in} \cdot \left(\exp\left(-k\beta_{eff(x,y)}\right) \cdot \exp\left(-ik\delta_{eff(x,y)}\right)\right). \tag{4.2}$$

  This representation allows for reconstructing the sample phase without the need for additional phase unwrapping. Also, the refractive index of most of the samples is positive away from absorption edges which can be used as an additional constraint in the process of reconstruction. Furthermore, in the case of a homogeneous sample with an a priori known chemical composition, it is possible to use known β and δ to reconstruct the local sample thickness $z_{s_{(x,y)}}$ directly. This reduces the number of unknowns by half and improves the reconstruction convergence. The complex transfer function representation (equation 4.1) was used for the reconstruction of the strong amplitude Siemens star test object (Sections 5.2,6), while the *thickness* representation was used for the studies of the plastic samples morphology described in Section 5.4.

- **Scan positions:** The scan positions correction is an integral part of the ptychography reconstruction since the maximum achievable resolution is directly limited by the accuracy of the measured scan positions [103]. The position

correction is included in the forward model by describing the individual sample positions $\boldsymbol{r}_j$ with a set of affine transformation matrices A_{θ_j}. Each of the A_{θ_j} describes a single sample position $\boldsymbol{r}_j$. Then, the virtual 'scanning' of the sample is done utilizing the differentiable grid sampler as described in [28], [104]. This allows AD to trace the gradient to both, the object function O and the individual scan positions $\boldsymbol{r}_j$. The individual A_θ is represented as

$$A_\theta = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & h_x & 0 \\ h_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}, \tag{4.3}$$

where t_x and t_y describe horizontal and vertical translations, γ is the angle of rotation with respect to the coordinate axis, h_x, h_y are the shears for x and y axes, and s_x, s_y are the scaling factors for x and y axes. All of these may be optimized individually, however, for the experiments described in this work, only the lateral shifts were of interest. Thus, the AD-based position correction was performed by optimizing only the t_x, t_y values. A detailed description of this approach can be found in Section 5.1.2.

- **Probe:** To accommodate for the partial coherence effects and the shot-to-shot fluctuations of the SASE FEL beam, the probe in the forward model is represented as

$$P_j = \alpha_{ij} M_i, \tag{4.4}$$

where α_{ij} are unique modal weights for the i-th mode at j-th position, and M_i is the set of probe modes individually propagated to the detector plane. The approximated intensity is then calculated as an incoherent sum of the intensities of individual modes as in Equation 3.4:

$$\tilde{I}_j = \sum_i \left| \mathcal{P} \left\{ \alpha_{ij} M_i O \right\} \right|^2, \tag{4.5}$$

where $\mathcal{P}$ is an appropriate propagator.

The number of modes and the weights dimensionality may be changed according to the coherent properties of the radiation used for the CDI experiment. For example, several coherent modes, as well as shot-to-shot unique coeffi-

cients α_{ij}, are required in the case of the classical ptychography at SASE FELs (described in Chapter 5) to describe simultaneously a partially coherent and shot-to-shot varying probe. Whilst, for the single-shot ptychography (described in Chapter 6), it was enough to use only the 1D set of the modal weights α_i, since all the beamlets were produced by the same partially coherent probe.

- **probe defocus:** The ptychography experiments described in this work were performed in the focused beam geometry (3.3 (b)), and therefore, the probe wavefront was dominated by the defocus term. Thus, the individual mode wavefield can be represented as

$$M = \Omega_f \Psi_p, \tag{4.6}$$

where $\Omega_f = \exp\left(\frac{i\pi(x^2+y^2)}{\lambda f}\right)$ is a pure defocus term driven by the optimizable defocus parameter f, and Ψ_p is the residual wavefield. The separate optimization of the defocus parameter speeds up the convergence of the reconstructions and makes it possible to compensate for focusing imperfections. Potentially, other aberrations known to dominate the wavefront of the probe may be added to the forward model in a similar manner [92].

- **Propagation distance:** In the case of propagation in the near-field geometry (equation 2.12), the propagation distance z can be optimized. It is possible because both the Fresnel TF and IR propagators explicitly depend on z. This allows correcting for uncertainties in the setup dimensions. For example, such optimization was performed for the grating-sample distance during the single-shot ptychography reconstruction described in Chapter 6.

- **Diffraction grating efficiency:** In single-shot ptychography neighboring areas of the sample are simultaneously illuminated by multiple beamlets produced by splitting the FEL pulse with a 2D diffraction grating. Thus, the individual beamlets initially have different intensities due to the different diffraction efficiencies for the corresponding grating orders. The efficiency can be estimated from the grating design but should be experimentally refined to improve the resolution of the reconstructions. It is realized by adding the op-

timizable scaling factors β_j for the intensity of each of the individual diffraction orders. Further details of the single-shot ptychography forward model can be found in Chapter 6.

4.3 Loss function

The loss function evaluates the quality of the reconstruction by calculating the error between the measured I and approximated $\tilde{I}$ intensities as

$$\mathcal{L} = \mathcal{E}(I, \tilde{I}). \tag{4.7}$$

However, this form of the loss function in the presence of noise in the measured data may lead to overfitting [105] and an unstable convergence. This can be solved by adding the various a priory assumptions regarding the optimizable variables, such as the smoothness of sparsity of the sample function $O_(x, y)$ to the loss function. In practice, it is done by adding the regularization terms $\mathcal{R}_i$ leading to the modified form of the loss function:

$$\mathcal{L} = \mathcal{E}(I, \tilde{I}) + \sum_i \gamma_i \mathcal{R}_i, \tag{4.8}$$

where γ_i are the weights of the regularizers R_i. This section lists error metrics and regularizers used for the reconstructions described in this work.

4.3.1 Error metrics

- **Least square error** (LSE) evaluates the Euclidean distance between the measured I and approximated $\tilde{I}$ intensities. However, to equalize the influence of the high and low values of I, the LSE is calculated for $\sqrt{I}$ as proposed by Godard et al. [87], since I obeys Poisson statistics. The LSE is calculated as follows:

$$\mathcal{E}_{LSE}(I, \tilde{I}) = \frac{1}{N} \sum_i^N \left| \sqrt{I_i} - \sqrt{\tilde{I}_i} \right|^2 \tag{4.9}$$

- **Poisson negative log likelihood** (PNLL) might be a more efficient error metric in the case of relatively low photon counts and a Poisson noise dominated measured intensity. The PNLL introduced in [87] can be calculated as follows:

$$\mathcal{E}_{PNLL}(I, \tilde{I}) = \frac{1}{N} \sum_{i}^{N} \tilde{I} - I \log(\tilde{I}) \tag{4.10}$$

In theory, the use of PNLL results in a higher achievable resolution in low photon flux measurements at the price of slower convergence. However, all the experiments described in this work were performed with sufficiently high photon flux resulting in the same resolution with both, LSE and PNLL error metrics.

4.3.2 Regularizers

- **The l_1 norm** is calculated for the complex, refractive or *thickness* object representation as

$$\begin{aligned} l_1(O) =& \frac{1}{N} \sum_{i} \alpha_1 |O_i| + \alpha_2 (\arg(O_i)), \\ l_1(O) =& \frac{1}{N} \sum_{i} \alpha_1 |\delta_i| + \alpha_2 |\beta_i|, \\ l_1(O) =& \frac{1}{N} \sum_{i} z_{s_i}, \end{aligned} \tag{4.11}$$

  respectively, where N is the total number of the object pixels, $|\cdot|$ and $\arg(\cdot)$ denote the modulus and the argument of the object function respectively, α_1, α_2 are the empirically selected amplitude and phase specific weights introduced to adapt the regularizer for mostly phase or mostly amplitude objects, and z_{s_i} is the 2D distribution of the object thickness. The l_1 norm promotes sparsity in the reconstructed object or probe and reduces reconstruction artifacts appearing as individual overly intense or transparent areas.

- **The Total variation denoising** (TVD) term [106] for the complex, refractive

or thickness object is calculated as follows:

$$\begin{aligned}\mathcal{TV}(O) =& \frac{1}{N}\sum_i \alpha_1(|\nabla_x|O_i|\ | + |\nabla_y|O_i|\ |) + \alpha_2(|\nabla_x \arg(O_i)| + |\nabla_y \arg(O_i)\ |),\\ \mathcal{TV}(O) =& \frac{1}{N}\sum_i \alpha_1(|\nabla_x\delta_i| + |\nabla_y\delta_i|) + \alpha_2(|\nabla_x\beta_i| + |\nabla_y\beta_i|),\\ \mathcal{TV}(O) =& \frac{1}{N}\sum_i |\nabla_x z_{s_i}|,\end{aligned} \tag{4.12}$$

where $\nabla_x\cdot$ and $\nabla_y\cdot$ represent the gradients in the x and y directions calculated as the finite differences. TVD promotes smooth solutions, reduces the noise, and prevents discontinuities in the reconstruction.

- **The orthonormality constraint** ensures the othonormality of the reconstructed probe modes M_i, which, as has been shown in [107] ,improves the convergence of the probe mode reconstructions. It can be formulated as

$$\mathcal{ORT}(M) = \sqrt{\sum_{i,j}^{K} \left|M_b^{ia}M_{ja}^b - \mathbb{I}_K\right|^2}, \tag{4.13}$$

where $M_b^{ia}M_{ja}^b = G_j^i,\ j,i \in 1...K$ is the Einstein notation for a $K \times K$ matrix of all possible mutual dot products of individual modes, and $\mathbb{I}_K$ is the identity matrix of a rank K. This regularizer forces the reconstructed modes to be orthonormal without the need for out-of-loop orthonormalization.

The aforementioned regularizers were extensively applied for all the reconstructions described in this work. The selection of the regularization strength coefficient γ_i was performed empirically to provide the best reconstruction results.

4.4 Optimizer and hyperparameters selection

The optimizer utilizes the gradients of the loss function provided by AD to change the optimizable variables minimizing the loss function and thus reconstructing the correct sample and probe values. While there are plenty of gradient-based optimizers

[108] available, most of them utilize the same principles to perform the optimization. The basic gradient descent (GD) algorithm utilizes the gradients $\frac{\partial \mathcal{L}}{\partial v}$ of the loss function $\mathcal{L}$ with respect to the optimizable variable v to change the optimizable variable at each step as

$$v' = v - l\frac{\partial \mathcal{L}}{\partial v}, \tag{4.14}$$

where l is the learning rate. The learning rate determines the step size and is an *hyperparameter* of the reconstruction [109]. This means that it is directly determining the success of the reconstuction while cannot be determined from the measured values I_j and needs to be set based on some other assumptions. However, the basic GD optimizer tends to behave poorly in case of complex loss function landscapes and tends to overshoot the minima due to the fixed step size l. This problem can be mitigated by the *momentum* gradient descend (MGD) [110] which considers a history of the previous gradients when calculating an update as

$$\begin{aligned} \kappa' &= \alpha\kappa + l\frac{\partial \mathcal{L}}{\partial v}, \\ v' &= v - \kappa', \end{aligned} \tag{4.15}$$

where $\alpha\kappa$ is the momentum term proportional to the previous update value κ.

The more advanced Adam optimizer [111] was used for the reconstructions described in this work. This algorithm uses the weighted mean and variance of the past values of gradients similarly to the MGD and shows a better performance than most of the other gradient-based optimization approaches [108]. Additionally, it is less sensitive to the absolute scale of the gradients and thus can better work with non-normalized input data, which simplifies its use for ptychography. Further details about the Adam implementation can be found in [111].

The Adam optimizer allows the selection of individual learning rates for each of the optimizable variables. The learning rates determine the update size at each step of the reconstruction, are essentially hyperparameters, and thus should be selected empirically [109]. The correct selection of the learning rates for each optimizable parameter is important for a successful reconstruction. Low values of learning rates will lead to slow convergence of the optimizable parameters. In contrast, high values will lead to unstable operation of the optimization scheme and divergence of the optimizable parameters. With a fixed batch size, it becomes possible to determine

the optimal learning rates l_n for each optimized parameter. While several different methods of learning rate estimation exist, in this work, the method proposed by Smith et al. [112] was used. This method proposes performing the optimization with learning rates changing from very low values to very high ones, $l \in 10^{-15}...10^{5}$, and then selecting the value at which the loss function decreases most rapidly by the lowest value of its gradient. This procedure was repeated for all hyperparameters resulting in a set of learning rates providing the best reconstruction results. The regularizer weights were selected empirically to obtain a smooth and sparse solution without visible artifacts.

4.5 Conclusion and outlook

This chapter described the AD-powered ptychography framework capable of performing ptychography reconstructions by minimizing the loss function and evaluating the error between the measured and the reconstructed intensities. The minimization is performed by a gradient-based optimizer utilizing the gradients of the loss function obtained by a reverse-mode AD. This approach fully evades the need for any manual gradient derivation, thus greatly simplifying the development and modification of the ptychography algorithm.

The individual components of the framework may be flexibly changed without the need for any additional analytical calculations. This permits easy adaptations to various experimental configurations and simplifies ptychography in challenging experimental conditions such as simultaneously spatially-fluctuating ad partially-spatially-coherent beams of the SASE FELs. Furthermore, the numerical calculations inside the framework are performed using a state-of-the-art PyTorch library. This allows for greatly increasing the speed of the phase retrieval. Additionally, the built-in support of the multi-GPU calculations enables faster reconstruction from a large number of measurements by utilizing the data-parallel mode.

In the future, additional forward models can be implemented to expand the described framework to Bragg-ptychography [74], and ptycho-tomography [113]. Additionally, the introduction of efficient second-order gradient-based optimizers may dramatically improve the convergence speed.

The next Chapters show practical applications of the described framework. Chapter 5 presents several classical ptychography experiments performed at the FLASH2, while Chapter 6 describes the first single-shot ptychography experiment performed at a soft X-ray FEL.

Chapter 5

Realisation of classical ptychography at a SASE FEL

SASE FELs produce simultaneously ultra-bright and ultra-short light pulses, which makes them ideal photon sources for performing coherent diffraction imaging. However, SASE FEL pulses are only partially spatially-coherent [24] and are subject to significant spatial fluctuations. These properties make the ptychography at SASE FELs a very challenging task usually solved by selecting only the most coherent part of the beam [25], which leads to a decrease in the usable photon flux.

A forward model presented in this chapter is capable of working with a simultaneously partially spatially-coherent and spatially-fluctuating probe. This model was developed for and applied in the proof-of-concept ptychography experiment performed at FLASH2. The proposed model was used for the reconstruction of a high-resolution image of the sample as well as the main probe modes and shot-to-shot unique modal weights. The resolution and the throughput of the ptychography FLASH2 were further increased in the next measurements performed with the PERCIVAL detector. The PERCIVAL detector was used for the first time at a soft X-ray FEL during this experiment. The large detector chip size, in combination with the high dynamical range, allowed to improve the achievable resolution by 30%. In turn, the high detector frame rate enabled faster measurements resulting in a total measurement time of 2 *min*.

Finally, this chapter presents the results of ptychographic imaging of thin plastic films. In this experiment, ptychography was used to study the morphology of thin

plasma-treated plastic samples due to its ability to image large sample areas with high resolution. The AD-based reconstruction algorithm was used to directly obtain the thickness of the samples by reformulation of the sample transfer function in terms of the effective refractive index. The obtained thickness distribution was used to analyze the effects of the different plasma-treatment regimes on surface properties of the samples.

The contents of this chapter are published in [28], [29].

5.1 Forward model for description of SASE FEL pulses

In AD-powered ptychography reconstruction, the role of the forward model is to describe the light propagation and interaction with the sample to approximate the measured intensity. In turn, convergence and resolution of the reconstruction depend on the quality of the approximation and thus on the ability of the forward model to describe all the parameters of the experiment.

This section presents a forward model capable of describing the shot-to-shot fluctuating and partially-spatially coherent SASE FEL pulses. First, it introduces the formalism describing the coherent properties of the probe. Then it presents the position correction algorithm that accommodates the pointing instabilities and imperfectly known scan positions. Finally, it gives an overview of the entire computational graph.

5.1.1 Partial coherence treatment

At FLASH2, light pulses are highly but not fully spatially coherent [24]. Additionally, due to the stochastic nature of the SASE process, the individual pulses experience spatial shot-to-shot fluctuations, which also should be considered by the forward model. The partial spatial coherence is typically described by the expression of the probe wavefield in terms of mutually incoherent modes on an orthogonal basis [83]. Under this formalism, a probe P is expressed as a set of K mutually in-

coherent orthonormal modes M_i with corresponding modal weights α_i. This allows calculating the resulting intensity in the detector plane as

$$I = \sum_{i}^{K} \alpha_i \left| \mathcal{P}\{M_i \times O\} \right|^2, \tag{5.1}$$

where $\mathcal{P}$ is the appropriate propagator. However, this expression still assumes constant illumination at all scan positions and should be adapted to take into account the shot-to-shot fluctuations of the beam.

In turn, a typical approach for ptychography with a fluctuating but spatially fully-coherent probe is orthogonal probe relaxation (OPR) [114], [115]. This method assumes that probes at different scan positions are still sufficiently similar despite some spatial fluctuations. Or, mathematically speaking, they all can be expressed in a common basis with a relatively small number of basis vectors. If this assumption holds, then the problem of reconstructing a unique probe for each scan position can be simplified by reconstructing only a small set of common basis vectors and the individual weight coefficients for each of the scan positions. In practice, a common basis is typically found by calculation of the truncated singular value decomposition (TSVD) [116] of the matrix of all unique probes after each full iteration of the reconstruction. Within the OPR, the measured intensity at the j-th position can be calculated as follows:

$$I_j = \left| \mathcal{P}\{O_{\mathbf{r}_j} \cdot P_j\} \right|^2 = \left| \mathcal{P}\{O_{\mathbf{r}_j} \cdot \sum_{i}^{K} \alpha_{ij} M_i|\} \right|^2, \tag{5.2}$$

where M_i are vectors from the orthonormal basis of rank K obtained with TSVD, and α_{ij} are corresponding weights of the i-th basis vector. The OPR assumes all probes to be fully coherent and thus cannot be applied to partially coherent illumination.

In this work, a model that simultaneously considers partial spatial coherence and spatial fluctuations of the probe was used to describe the SASE FEL radiation. The model considers each FEL pulse as partially coherent and containing multiple mutually-incoherent modes with unique modal weights. However, the number of unique modes is assumed to be limited. The correctness of this assumption has been experimentally proven by coherence measurements performed at FLASH2 [24].

Under this assumption, each measured probe can be expressed as an incoherent sum of orthonormal modes, and each measured intensity pattern is described by an incoherent sum as:

$$I_j = \sum_{i=1}^{K} \left| \mathcal{P} \left\{ \alpha_{ij} M_i \cdot O_{\mathbf{r_j}} \right\} \right|^2, = \sum_{i=1}^{K} \sqrt{\alpha_{ij}} \left| \mathcal{P} \left\{ M_i \cdot O_{\mathbf{r_j}} \right\} \right|^2 \tag{5.3}$$

where a_{ij} is a set of $K \times N$ modal weights unique for each individual mode $i \in 1...K$ and FEL pulse $j \in 1...N$. This model is capable of describing the simultaneously partially spatially-coherent and spatially-fluctuating beams of the SASE FEL. Under this assumption, all fluctuations of the beam may be explained either as a change in corresponding modal weights or as a shift of the probe as a whole by changing the corresponding translational vector $\mathbf{r}_j$, which is found by the position refinement algorithm.

5.1.2 Position correction by automatic differentiation

Ptychography reconstruction relies on a mutual overlap of neighboring probe positions, directly linking the achievable resolution with scanning precision. This makes the position correction routines essential to the high-resolution reconstruction. AD-powered ptychography is not an exception and requires a position correction routine, ideally implemented directly to the forward model for additional flexibility and computational efficiency. For this work, an AD-based position correction algorithm was developed. It is based on the affine transformation and differentiable sampling described in [104]. This approach was used for position correction during all reconstructions described in this work.

The model used for the AD-based position refinement is illustrated in Figure 5.1. At each scan position j exit wave Ψ_j is calculated as the product of the probe complex wavefield P (Figure 5.1 (d)) and the corresponding region of the object O^* shifted by the displacement vector $\boldsymbol{r_j}$ (Figure 5.1 (c)). Both P and Ψ_j have the same dimensionality of $N_{pixels} \times N_{pixels}$ as the measured diffraction pattern I_j and are fixed to the detector coordinate frame (gray frame in Figure 5.1 (b,c)) and placed in a coordinate grid G^t. Whilst the sample O has the dimensionality of $H \times W$

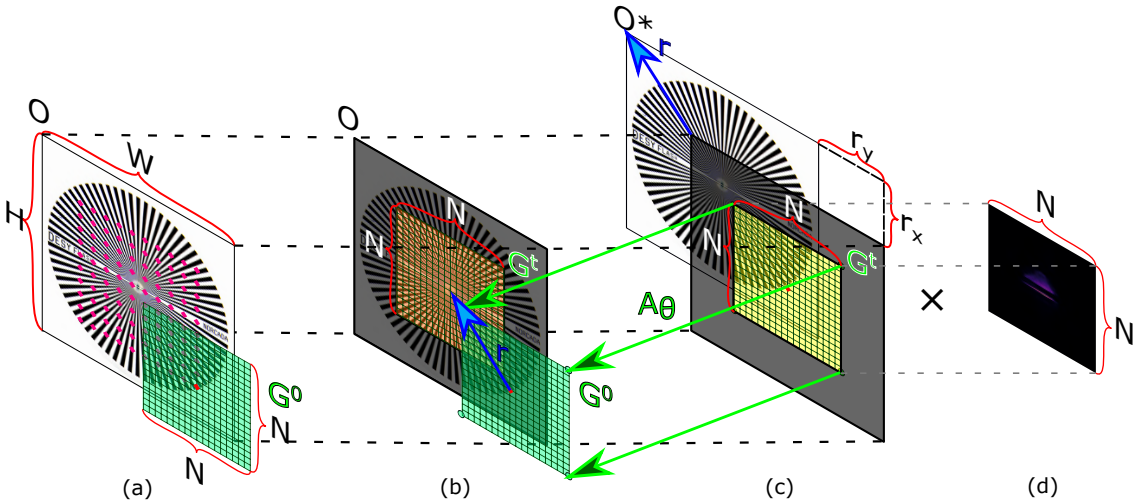

Figure 5.1: **Application of the affine transformation for scan position refinement. (a)** Sample with estimated scan positions. The red dot is the centre of the G^0 coordinates with the $N \times N$ coordinate grid (green). **(b)** The centered reference coordinate system G^t (orange grid) has a shift of $\boldsymbol{r}$ (blue arrow) relative to the grid G^0. **(c)** The sample O^* is shifted by $\boldsymbol{r} = (r_x, r_y)$ to align the part of the sample centered around the current scan point. **(d)** Exit wave calculated as the multiplication of the probe function with the central part of the computational array. Reference coordinate system and object coordinate system are connected by the affine transformation $\mathcal{A}_\theta$. Its parameters are iteratively optimized to get the best estimate of $\boldsymbol{r}_j$ for each individual scan position. Adapted from [28].

where $H, W > N$ are height and width of the reconstructed object area (Figure 5.1 (a)). Therefore, only the $N \times N$ area of the sample centered around a particular scan position (shown with green in Figure 5.1 (a)) is illuminated at the j-th scan position r_j and contributes to the j-th measured intensity I_j. This area is placed in a coordinate grid G_j^0 unique for each j-th scan position (green grid in Figure 5.1 (a)). Thus, the AD-compatible description of the object scanning requires first, to find the relations between all local coordinate grids G_j^0 and the reference coordinate grid of the detector G^t. And then, to define a differentiable way to calculate the values of the shifted objects O_j^* from the initial object O.

AD-based position refinement describes the relations between the two coordinate

grids $G^0 = (x^0, y^0)$ and $G^t = (x^t, y^t)$ as

$$G^0 = \mathcal{A}_\theta(G^t),$$
$$\begin{pmatrix} x^0 \\ y^0 \\ 1 \end{pmatrix} = \mathcal{A}_\theta \begin{pmatrix} x^t \\ y^t \\ 1 \end{pmatrix} = \begin{bmatrix} A_{00} & A_{01} & A_{02} \\ A_{10} & A_{11} & A_{12} \\ A_{20} & A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} x^t \\ y^t \\ 1 \end{pmatrix}, \tag{5.4}$$

where $\mathcal{A}_\theta$ is the affine transformation operator expressed as

$$A_\theta = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & h_x & 0 \\ h_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}, \tag{5.5}$$

where t_x and t_y describe horizontal and vertical translations, γ is the angle of rotation with respect to the coordinate axis, h_x, h_y are the shears for x and y axes, and s_x, s_y are the scaling factors for x and y axes. This formalism allows describing any planar movement of the sample and performing the position correction by finding the values of $\mathcal{A}_\theta$ which minimize the loss function. Additionally, in the absence of noise, the values of $\mathcal{A}_\theta$ can be refined up to a numerical precision enabling sub-pixel positioning.

Affine transformation can be used for the AD-based position correction as soon as the gradients of the loss function with respect to $\mathcal{A}_\theta$ and O^0 are defined. In general, the calculation of the transformed object O^* from the initial sampling grid (x_i^0, y_i^0) and initial object O^0 can be described as follows [104]:

$$O^*(i) = \sum_{n=1}^{H} \sum_{m=1}^{W} O_{nm}^0 \mathcal{K}(x_i^0 - m; \Phi_x) \mathcal{K}(y_i^0 - n; \Phi_y), \ \forall i \in [1...HW], \tag{5.6}$$

where Φ_x, Φ_y are the parameters of the sampling kernel $\mathcal{K}$ which is used to interpolate pixel values. The bilinear sampling kernel was used in this work because of its low numerical complexity turning equation 5.6 into:

$$O^*(i) = \sum_{n=1}^{H} \sum_{m=1}^{W} O_{nm} max(0, 1 - \left|x_i^0 - m\right|) max(0, 1 - \left|y_i^0 - n\right|). \tag{5.7}$$

The gradients with respect to $G^0 = (x_i^0, y_i^0)$ and O are given as:

$$\frac{\partial O^*(i)}{\partial O_{nm}} = \sum_{n=1}^{H}\sum_{m=1}^{W} max(0, 1 - \left|x_i^0 - m\right|)max(0, 1 - \left|y_i^0 - n\right|), \tag{5.8}$$

$$\frac{\partial O^*(i)}{\partial x_i^0} = \sum_{n=1}^{H}\sum_{m=1}^{W} O_{nm} max(0, 1 - \left|y_i^0 - n\right|) \begin{cases} 0 & \text{if } |m - x_i^0| \geq 1 \\ 1 & \text{if } m \geq x_i^0 \\ -1 & \text{if } m < x_i^0 \end{cases}, \tag{5.9}$$

and the same as in Eq. 5.9 for $\frac{\partial O^*(i)}{\partial y_i^0}$. Thus, the differentiable sampling yields the gradients of the loss function with respect to the sample function O and the coordinate grid $G^0 = (x_i^0, y_i^0)$ allowing AD to further trace the gradients to $\mathcal{A}_\theta$ and thus to perform the scan position refinement. This formalism incorporates sample movements as well as the FEL beam jitter into the forward model and permits scan position refinement through the optimization of the individual components of $\mathcal{A}_\theta$.

5.1.3 Forward model

A resulting forward model for classical ptychography at a SASE FEL is shown in Figure 5.2. The model combines the multi-modal expression of the probe with the affine-transformation-based scan position refinement for a description of the partially spatially-coherent and spatially-fluctuating SASE FEL beams. Thus, the forward pass of the model calculates the intensity at the j-th scan position as

$$I_j = \sum_i \left|\mathcal{P}\left\{\alpha_{ij} M_i \cdot S \cdot A_{\theta_j}\left\{O\right\}\right\}\right|^2, \tag{5.10}$$

where A_{θ_j} represents an affine transform driven by its parameter matrix θ_j, S is the support of the probe function, $\mathcal{P}$ is the propagator determined by the experiment configuration and $\alpha_{ij} M_i$ is the set of modes M_i with unique modal weights α_{ij} for each of the i-th mode at j-th scan position. The loss function is composed of the error term and regularizers resulting in the following expression

$$\mathcal{L} = \mathcal{E}_{LSE}(I, \tilde{I}) + \gamma_{l_1} l_1(O) + \gamma_{TV}\mathcal{TV}(O) + \gamma_{\mathcal{ORT}}\mathcal{ORT}(M), \tag{5.11}$$

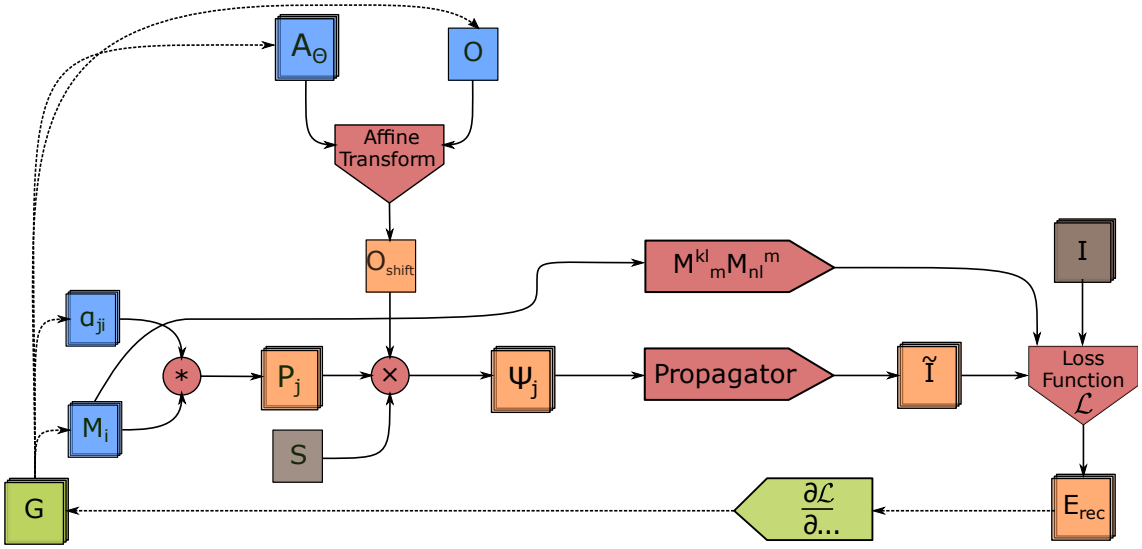

Figure 5.2: **Computational graph of the forward model for classical ptychography at SASE FELs.** Input data (measured intensities - I and support S) are colored in brown, operators are colored in red, optimizable variables (modes - M_i , modal weights - α_{ji}, affine transformation matrix - $\mathcal{A}_\theta$, and object function - O)are colored in blue, intermediate variables are colored in orange, backward propagation of error represented with dotted line and green color. Adapted from [28].

The detailed expressions of the $\mathcal{E}_{LSE}$, $l_1(O)$, $\mathcal{TV}(O)$ and $\mathcal{ORT}(M)$ can be found in Section 4.3.

This model was utilized to perform the classical ptychography experiments at FLASH2 described in this chapter.

5.2 Proof of concept experiment

The forward model was applied during the first classical ptychography experiment at FLASH. The main experimental parameters are summarized in table 5.1. The measurements were performed at beamline FL24 at FLASH2 at a wavelength of $\lambda = 13.5\ nm$ which corresponds to photon energy $E_{ph} = 91.8\ eV$. The pulses had an average photon pulse energy $E_{pulse} = 90\ \mu J$.

The experimental setup is illustrated in Figure 5.3. The FLASH2 beam was focused using bendable Kirkpatrick-Baez (KB) mirrors to focus approximately 2.6

Parameter	**Value**
Wavelength	13.5 nm (91.8 eV)
Pulse energy	90 μJ
Probe size	300 nm
Scan size	11×11
Overlap	90%
Images per position	10
Focus sample distance	120 mm
Sample detector distance	2000 mm
Detector size	2048×2048 $pixels$
Detector pixel	13.5×13.5 μm
Cropped ROI	1024×1024 $pixels$
Numerical resolution	900 nm

Table 5.1: Parameters of classical ptychography proof of concept experiment.

m after the center of the second mirror [33], [51], [117]. At the focal plane, the KB optics produces a focus of ≈10×10 μm^2 FWHM, which was tuned and characterized using a Shack–Hartmann wavefront sensor [118]. The sample was placed approximately 12 cm behind the focus leading to a beam size of approximately 300×300 μm^2 at the sample plane. An aperture of 700 μm diameter was placed 1 cm before the sample to filter out the photon pulses, whose pointing was extremely far from the current scan position. The photon beam was attenuated with the gas attenuator, and metal foil filters were used to prevent the sample damage, as well as suppress the higher-order harmonics of the FEL radiation [32]. Attenuation was also necessary to avoid detector saturation since no central beamstop was used during the measurements.

A 4.5×4.5 mm^2 Siemens star pattern (shown in Figure 5.3 (b)) on a 150 nm thick SiN membrane was used as the sample. The opaque part of the Siemens star was made of 190 nm thick Au and 10 nm thick Cr layers and is thin enough to fulfill the projection approximation for $\lambda = 13.5nm$. At this wavelength, the Siemens star may be considered a binary black-white object. The sample has features of 0.5 μm and 300 μm size in the central(stripes) and border(letters) regions, respectively, as shown in the inset in Fig. 5.3 (a).

The sample was mounted on a 2-axis translation stage with scanning axes perpendicular to the optical axis. Two regions of the sample were measured during

the experiment: *letters* region with *DESY FL* near the border of the sample and the central *stripes* region. The diffraction patterns were measured from 11×11 scan positions on a rectangular grid. Both regions were scanned with a $20\mu m$ step size which is equivalent to 90% overlap. Such a high overlap degree was used to ensure a high redundancy in the measured data and thus the convergence of the reconstruction. The scanned areas (letters and stripes) are shown in the insets in Fig. 5.3. Ten diffraction images were measured for each of the scan positions, with the most intense pattern selected as the representative for further analysis. This, together with the use of a restricting aperture, allowed to limit the influence of the limited pointing stability by selecting the pulses hitting the center of the restricting aperture. However, no significant aperturing was observed during the measurements showing that the real pointing stability was significantly better than the size of the aperture. An ANDOR iKon-L SO CCD camera with a pixel size of 13.5 μm and 2048×2048 pixels was placed 2 m behind the sample to measure the diffraction patterns. The relatively low camera frame rate of 1.1 $sec/frame$ resulted in a total pure measurement time of 23 $min/scan$, which, together with the motor scanning time, resulted in a total time of 25 $min/scan$. Based on the experimental geometry and wavelength, the Fresnel propagator of equation 2.14 was used in the reconstruction to describe the light propagation.

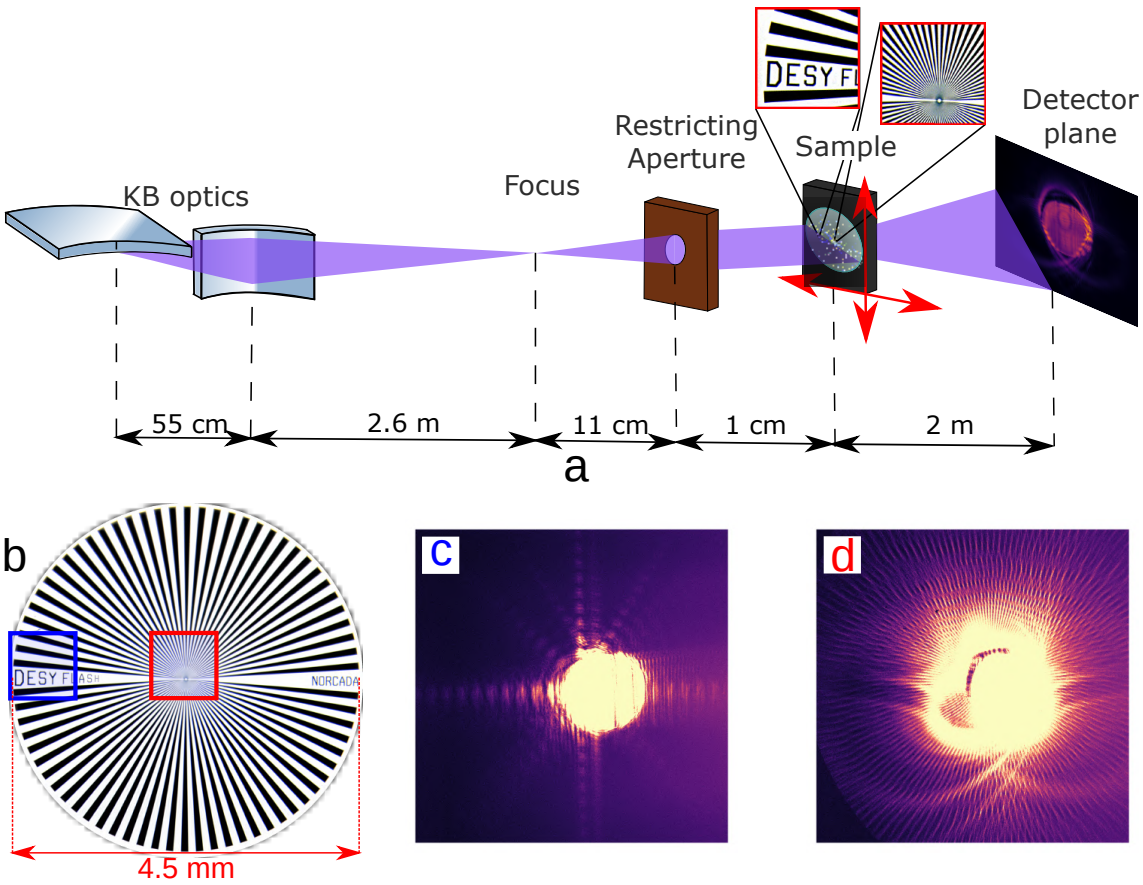

Figure 5.3: **Classical ptychography at FLASH2. (a)** Outline of the experimental setup. The FEL beam was focused using KB mirrors, resulting in a divergent probe geometry. A restricting aperture several times larger than the beam cross-section was used to limit spatial fluctuations. The sample was placed approximately $12\ cm$ after the focus position. An ANDOR iKon-L SO CCD was placed in the intermediate field $2\ m$ downstream of the sample. **(b)** Image of the sample with scanned areas shown with colored rectangles. **(c),(d)** Typical diffraction patterns measured at the letters and central area, respectively. Intensity values are capped to show the full dynamic range. Adapted from [28].

5.2.1 Reconstruction and results

The reconstruction was performed utilizing the AD-powered ptychography engine with the forward model described in Section 5.1. The calculations were performed utilizing four Quadro M6000 Nvidia GPUs with 12 GB memory each. The parallelization allowed a batch size of 120 and 40 scan positions for the single-mode (fully coherent case) and multi-modal reconstruction, respectively. The central 1024×1024 pixels of the diffraction data were selected for reconstruction to ease the computational costs. Background noise was subtracted by using an average of 100 dark frames. The initial probe assumption was obtained by backpropagating the square root of the averaged intensities measured without a sample in the beam to the sample plane as

$$P_0 = \mathcal{P}^{-1}\left\{\sqrt{\bar{I}_j}\right\}. \tag{5.12}$$

The sample was initiated with random module values ranging between 1 and 0 and a planar phase. While the optimizer learning rates and other hyperparameters were selected as described in 4.4.

As a first step, reconstruction with an AD-based ptychography model without position correction and assuming full spatial coherence was performed. In theory, this approach should demonstrate a performance compatible with the classical ptychographic engines such as mPIE [67]. The same fully coherent illumination was assumed for all scan positions, and no position refinement was performed. The reconstruction procedure consisted of 200 iterations of an object-only optimization followed by 350 iterations of object and probe joint optimization. After that, the reconstruction error converged to stable values and did not improve further. The reconstruction of the *stripes* region did not converge to any meaningful result, and the reconstructed *letters* region was significantly distorted due to the imperfectly known scan positions. The obtained probe and sample functions are shown in Fig. 5.4.

The results illustrate that basic AD-based models can perform satisfactorily in ptychography experiments for stable illumination conditions and perfectly known scan positions. However, they lack an AD-based position correction routine to correct the imperfectly known scan positions and cannot produce satisfactory results at SASE FELs.

Then, a reconstruction using a fully coherent probe and a simultaneous scan

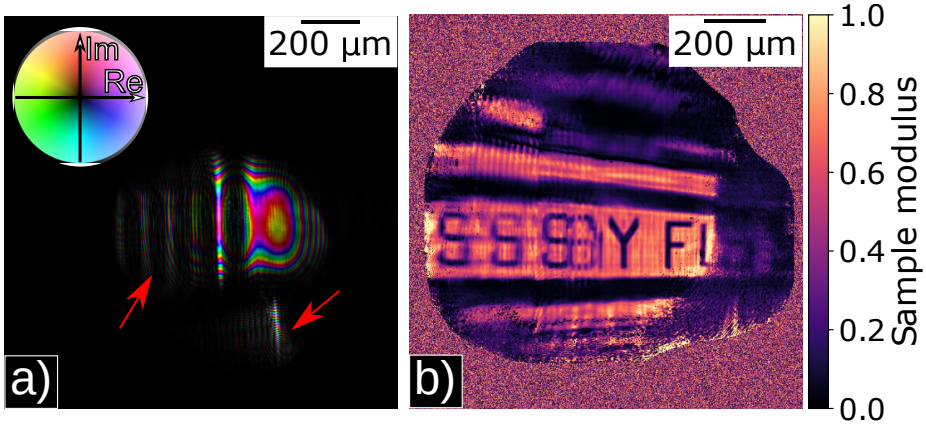

Figure 5.4: **Results of the reconstruction in assumption of a fully coherent average beam without the position correction. (a)** reconstructed sample amplitude from the letters area. **(b)** reconstructed probe function, complex values represented by domain coloring with saturation representing the modulus and hue representing the phase according to the colorwheel. Reconstruction artifacts resulting from imperfectly known positions shown by red arrows. Adapted from [28].

position refinement was conducted to evaluate the performance of the position correction procedure. The optimization procedure consisted of a total of 1700 iterations. During the first 200 iterations, only the sample and scan positions were optimized. Then, the joint probe sample and scan position optimization with additional TVD and $l1$ regularization was performed for 1000 more iterations. Additionally, during the first 700 iterations, the probe support was included in the forward model by zeroing the probe values outside the restricting aperture. This step is not critical for the algorithm convergence, but, as shown in [67], it can speed up the reconstruction by further centering the probe.

The reconstructed sample functions for the *stripes* and the *letter* regions obtained after this step are shown in Fig. 5.5 (a) and (b), respectively. The scan position refinement retrieved some features missing in the first reconstruction, such as the upper lines in Fig. 5.5 (b), and removed striped artifacts present in the probe and the sample reconstructions using raw positions (red arrows in Figure 5.4 (b)). In this step, the *stripes* region of the sample (Figure 5.5 (a)) converged to a meaningful result. However, it was reconstructed only in the very center with multiple artifacts. The reconstructed probe function is shown in Figure 5.5 (c) and can be seen as an

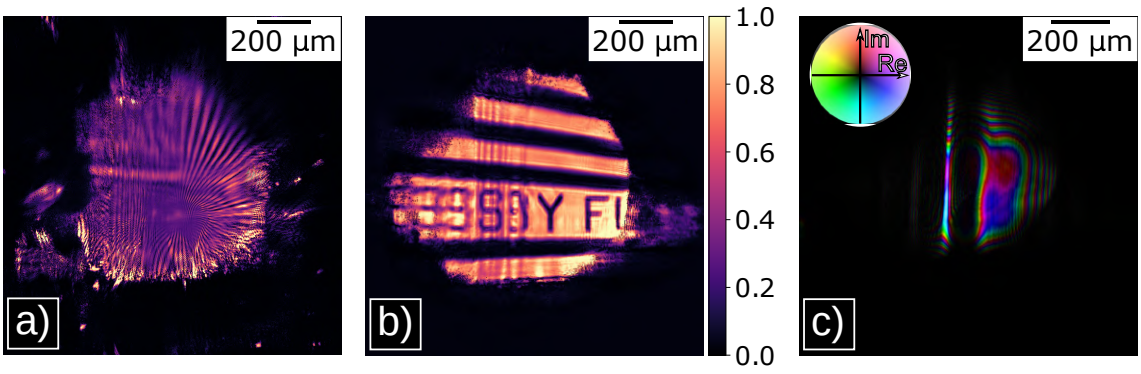

Figure 5.5: **Results of the reconstruction in assumption of a fully coherent average beam with position correction. (a)** Reconstructed modulus of the stripes-area of the sample. **(b)** Reconstructed modulus of the letters-area of the sample. **(c)** Reconstructed probe assuming full coherence and the same illumination for all of the scan points; represented with domain colouring according to the colour wheel. Adapted from [28].

averaged coherent beam illuminating the sample during the ptychographic scanning.

The stability of the AD-based position correction was evaluated by performing several independent reconstructions. The obtained scan position maps are shown in the Figure 5.6 together with the initially assumed scan positions. The AD-based scan position refinement algorithm is stable enough to reconstruct most of the scan positions within sub-pixel distance from their previous values during the trial-to-trial examination; thus, it can be concluded that it works reliably.

At the next stage of the reconstruction, the obtained sample and the average probe function were used as the initial data for the multi-modal reconstruction. The main probe modes were initialized with the orthogonal vectors obtained from the SVD of the probe wavefield reconstructed at the previous stage. The number of modes was increased until the occupancy of the least significant one became less than 0.5% in the final reconstruction. The sample, main modes, and unique modal weights for each scan position were reconstructed for an additional 200 iterations.

The reconstructed *letters* region of the sample is shown in Figure 5.7 (a). The left part of the sample, which was distorted when using the fully coherent reconstruction (Figure 5.5 (b)), can now be clearly distinguished using the multi-modal reconstruction. Additionally, the upper left part of the sample, which was not reconstructed previously due to a dim illumination, is now visible. Finally, the multi-modal recon-

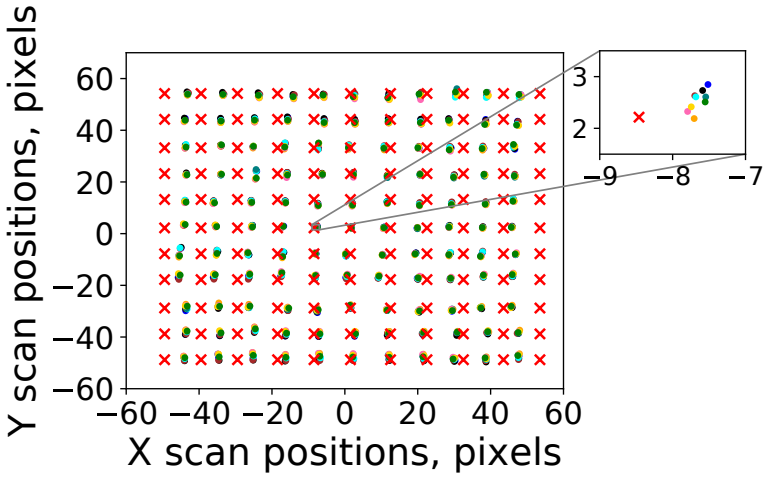

Figure 5.6: **Results of the position correction.** Scan position maps (red crosses) and those obtained from the AD-based position refinement algorithm (circles, different colors correspond to different independent reconstructions). Most of the retrieved scan positions show a sub-pixel stability. Adapted from [28].

struction procedure with position correction was repeated for the central region of the sample, with the modes obtained for the letter-region reconstruction used as the initial assumption for the probe. The reconstruction of the stripes-region is shown in Figure 5.7 (b). In contrast to the fully coherent reconstruction, the multi-modal reconstruction was able to retrieve a much larger part of the sample with better resolution. The image obtained is partially blurred as the size of the stripes in the center is less than a micrometer, thus being smaller than the numerical resolution.

The resolution of various modalities of reconstruction, as well as the dynamics of the values of the loss functions, are shown in Figure 5.8. The resolution of the reconstruction was evaluated using the Fourier ring correlation (FRC) [119] calculated for the two sample functions independently reconstructed from the two different sets of measured diffraction patterns. This way, the resolution of the reconstruction was estimated as 9.65 μm,3.23 μm, and 2.68 μm for the reconstruction from the unrefined positions, with the position correction and in multi-modal approximation, respectively. Thus, the proposed partially coherent forward model better describes the SASE FEL radiation resulting in a higher resolution of reconstructions. This is also proven by the values of the loss function achieved during the reconstruction (Figure 5.8 (b)). The reconstruction within the proposed partially coherent forward

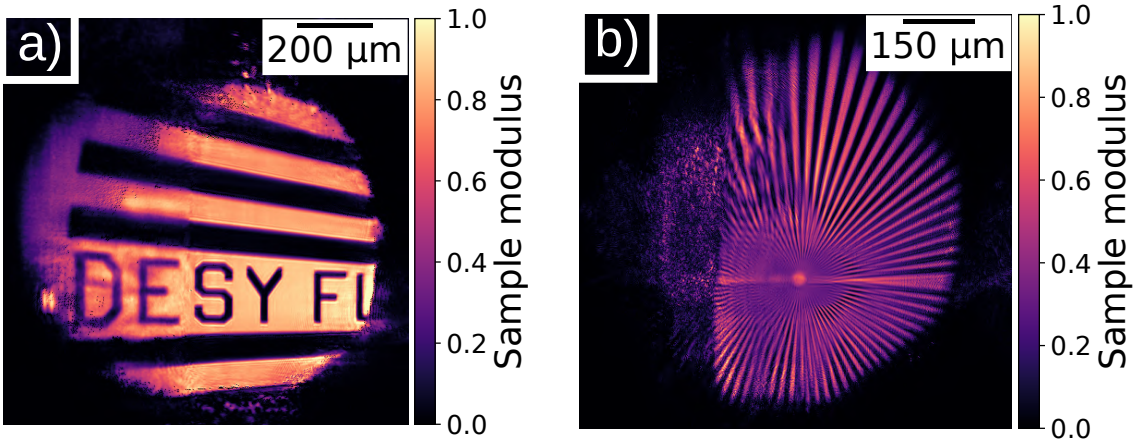

Figure 5.7: **Results of the reconstruction in assumption of a partially-coherent and spatially-fluctuating beam.** Sample modulus in the letters **(a)** and stripes **(b)** areas reconstructed in the multi-modal approximation. Adapted from [28].

model (Figure 5.8 (b) green curve) results in loss function values which are 1.5 orders of magnitude lower than for the other reconstruction methods(Figure 5.8 (b) orange and blue curves).

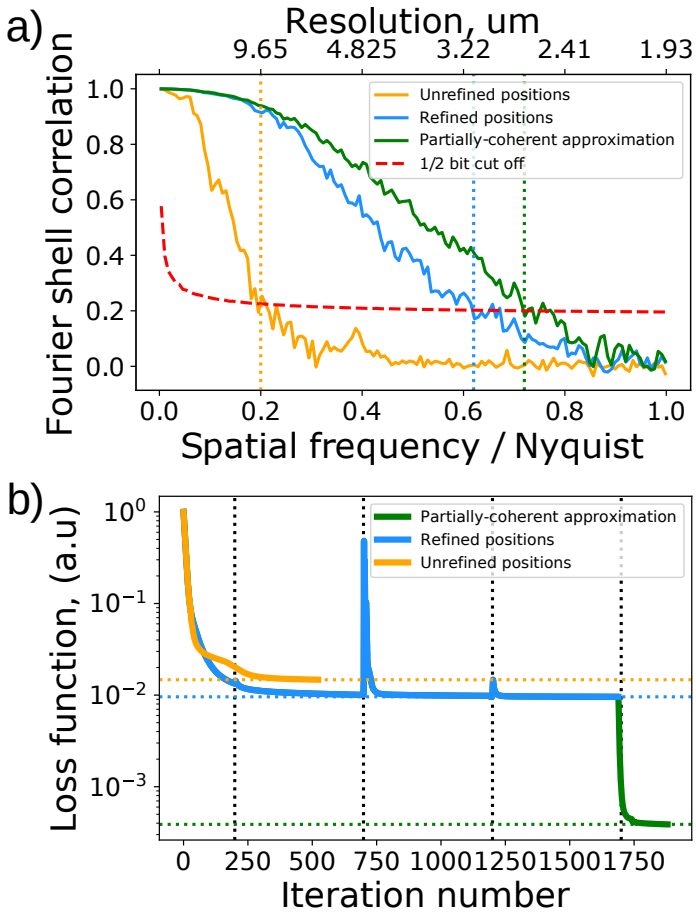

Figure 5.8: **Comparison of the FRC curves (a) and loss function values (b) of different reconstruction regimes.** For both **(a)** and **(b)**: AD-based reconstruction - orange curves; Average beam with position refinement - blue curves; Multi-modal - green curves. **(a)** resolution estimated using FRC. The red dashed line represents the 0.5 bit threshold. **(b)** Evolution of the loss functions during the reconstruction process. Dashed vertical lines mark: the iterations at which the probe optimization was started – 200, support was removed – 700, TV denoising of the sample was started – 1200, and the reconstruction in the partially coherent approximation was started - 1700. Adapted from [28].

5.2.2 Beam characterization

Another significant benefit of the proposed forward model is its ability to reconstruct the main modes of the FEL radiation and the shot-to-shot unique modal weights. The modal expansion can be used to reconstruct and analyze each FEL shot individually. This enables the use of ptychography for shot-so-shot wavefield characterization and provides insides into the coherence properties of the SASE FEL radiation.

The main probe modes M_i reconstructed within the partially-coherent assumption are shown in Figure 5.9. The average modal weights $\overline{\alpha}_i$ obtained by averaging

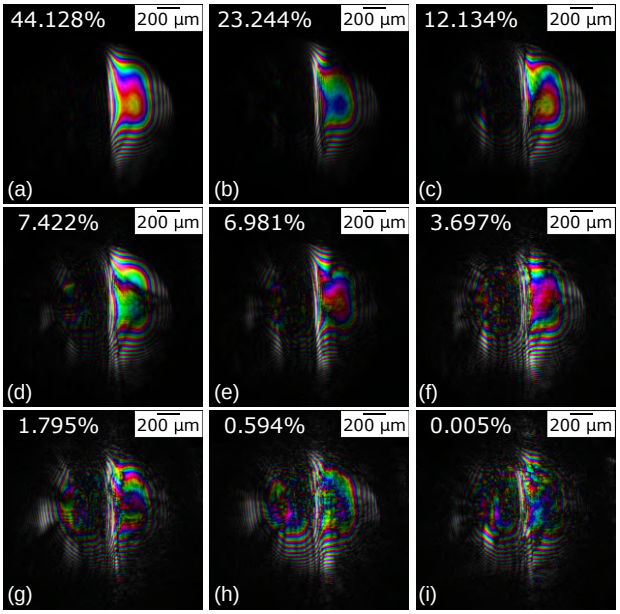

Figure 5.9: **Probe modes at the object plane reconstructed under partial coherence approximation and their average weights.** Amplitude mapped in brightness and phase in hue. Average modal weights shown in upper-left corners. Adapted from [28].

the single-shot modal weights α_{ij} across all the j scan positions are shown in the top left corners in Figure 5.9. It is possible to determine an average global degree of transverse coherence of the FLASH2 radiation considering the retrieved modes and their average weights α_{ij} from Fig. 5.9. If all the modes originate from the partial coherence of the FLASH2 source its average degree of transverse coherence can be

estimated from the average modal weights according to equation 2.34 as

$$\overline{\xi} = \frac{\sum \overline{\beta}_i^{\,2}}{\left(\sum \overline{\beta}_i\right)^2} = 54\% \;;\; \overline{\beta}_i = \overline{\alpha}_i{}^2. \tag{5.13}$$

This result is in agreement with coherence measurements performed at FLASH2 [24] and FLASH1 [54].

Additionally, it is possible to use the retrieved modal weights α_{ij} to estimate the shot-to-shot fluctuations of the FLASH pulses and their spatial coherence since the proposed forward model provides unique modal weights α_{ij} for each individual shot. The stability of the reconstruction with respect to the coherence properties was evaluated by performing several independent multi-modal reconstructions from the same diffraction patterns measured in the *letters* region. The single-shot global degree of spatial coherence was determined for each of the scan positions in each of the reconstructions according to equation 5.13. The obtained values are shown in Figure 5.10 and show that the reconstruction steadily converges to the same values

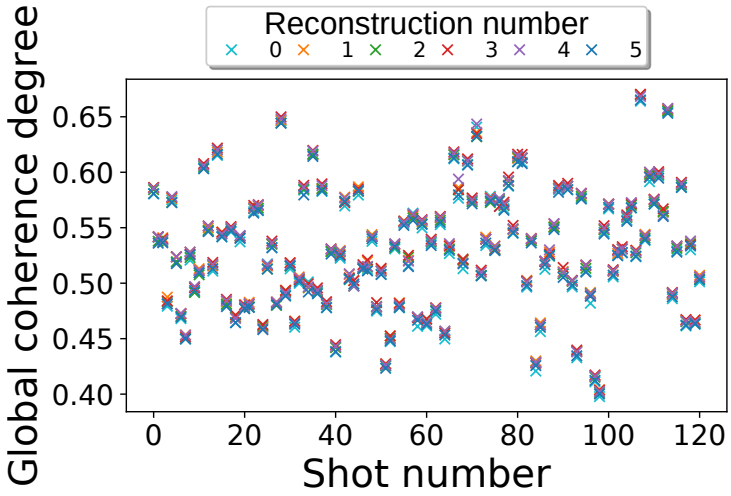

Figure 5.10: **Global degree of spatial coherence calculated for each reconstructed single-shot probe.** Results obtained from several independent reconstructions are plotted on top of each other and distinguished by color. The overlap of the obtained values indicates the high stability of the reconstruction. Adapted from [28].

of single-shot global coherence degree for independent reconstructions. This proves the ability of the proposed forward model to reconstruct the coherence properties of

the individual SASE FEL pulses and illustrates how the AD-powered ptychography measurements can be used for single-shot coherence characterization.

The reconstructed probe modes can also be utilized to characterize the performance of the beamline focusing optics for the partially coherent illumination. This can be done by studying the propagation of each individual mode. The focusing of the KB mirrors was studied by numerical propagation of the main mode in the vicinity of the focal plane. The sections of the propagated wavefield are shown in Figure 5.11 (a,b) The blue and orange dashed lines indicate the sample and focus positions, respectively. The vertical and horizontal focal planes were estimated by the maximum sharpness of the propagated beam [120] at 12.3 cm and 13.5 cm upstream of the sample. The main mode propagated to the plane of the second KB mirror is shown in Figure 5.11 (c).

The reconstructed modes and their single-shot weights may also be used for the reconstruction of the individual FLASH pulses to study the effects of the shot-to-shot spatial fluctuations. Examples of the reconstructed pulses are shown in Figure 5.12. Intensity distributions of two probes reconstructed from different scan positions and the average intensity at the sample plane are shown in Figure 5.12 (a-c). The intensity distribution of the same probes but at the focus position is shown in Figure 5.12 (a-c). The intensity distribution in the focus of the KB was reconstructed with a numerical resolution of $0.8\mu m$.

In conclusion, this section demonstrates that the AD-based ptychography with the forward model described in Section 5.2 is capable of performing ptychography reconstructions from data measured at SASE FELs. The model can describe the simultaneously spatially-fluctuating and spatially-partially-coherent light pulses by using the affine transformation-based position refinement and multi-modal representation of the probe with shot-to-shot unique modal weights. This allows for reconstructing the sample with a better resolution than classical ptychography algorithms without the need for beam aperturing or spatial fluctuation suppression. Additionally, the shot-to-shot unique modal weights enable the reconstruction of the unique wavefields for every scan position. This makes possible the shot-to-shot characterization of both the spatial and the coherent properties of the SASE FEL radiation and evaluation of the focusing optics performance. It is worth mentioning that the described model is not limited to SASE FELs and can be applied to other

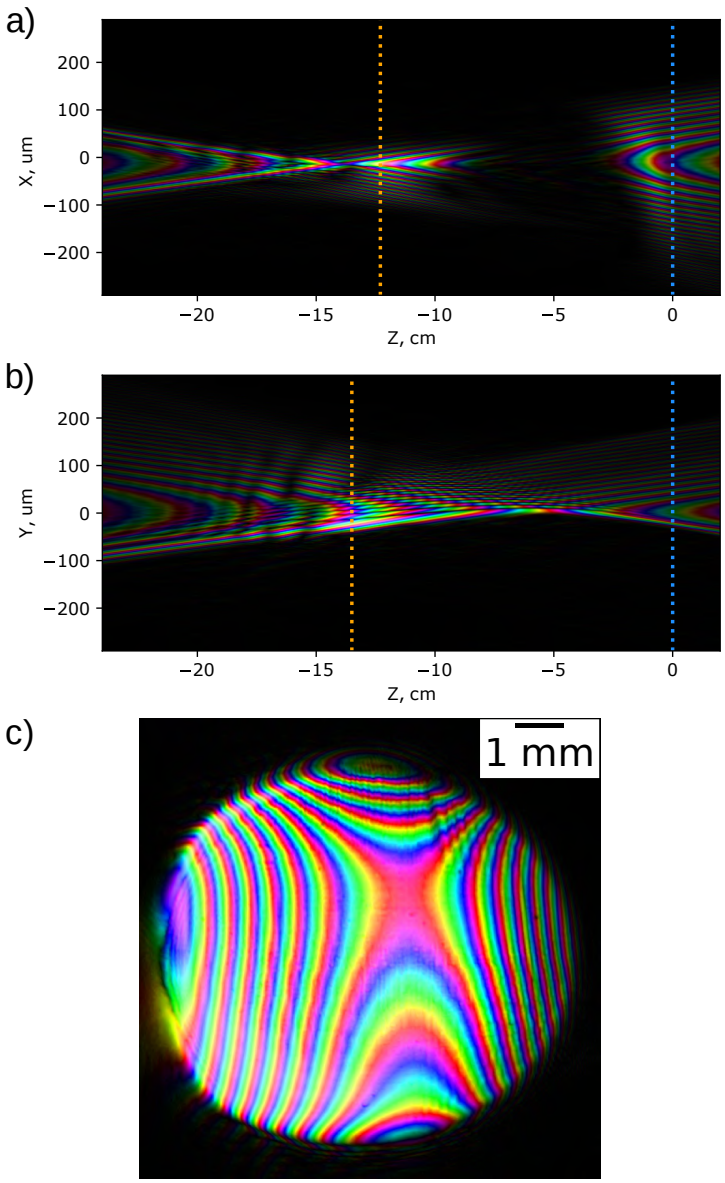

Figure 5.11: **Spatial characterization of the main reconstructed mode.** Horizontal **(a)** and vertical **(b)** caustics of the propagation of the reconstructed main mode. Sample and focal planes are shown with blue and orange lines, respectively. The horizontal and vertical focal planes are different due to the astigmatism that the KB system had in this special setup. **(c)** The main mode backpropagated to the second KB mirror plane with tilts and defocus removed. Complex values are shown with domain coloring, with modulus mapped to saturation and phase mapped to hue. Adapted from [28].

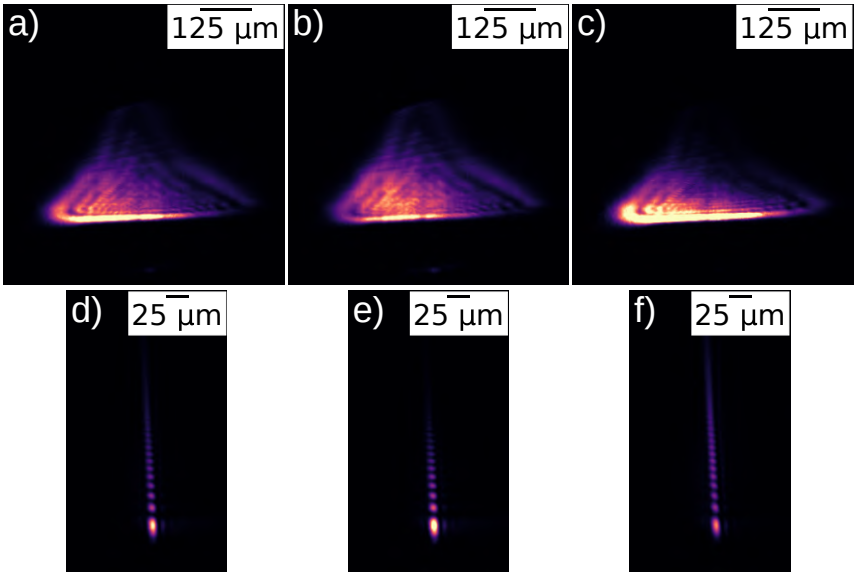

Figure 5.12: **Comparison of average and single-shot intensity distributions. (a)** Intensity of the average reconstructed probe. **(b, c)** Intensities of two reconstructed probes from different single shots.**(d)** Intensity of the back-propagation of **(a)** to the focal plane. **(e, f)** Intensities of the back-propagation of **(b, c)** to the focal plane. Color scales are shared among **(a – c)** and **(d - f)** to demonstrate intensity variations. Adapted from [28].

fluctuating sources such as HHG based ones [121].

5.3 Improving the resolution with the PERCIVAL detector

In the previous section, the first ptychography experiment at FLASH2 has been presented. It illustrates how advanced computational techniques simplify ptychographic reconstruction and bring it closer to routine use on SASE FELs. However, purely computational improvements alone are not enough to improve this technique for its routine use at SASE FELs.

Standard Andor iKon L detector was used to measure the data leading to the reconstruction of the object, main probe modes, and shot-to-shot modal weights. However, the low frame rate (1.1 sec) of the detector prevented the full use of the high repetition rate of FLASH. At the same time, the relatively low dynamical range of 4.5 orders required significant attenuation of the FEL beam reducing the amount of registered highly scattered photons. In this section, the results of the first experiment at FLASH utilizing the novel PERCIVAL detector [122] are reported.

5.3.1 Parameters of the PERCIVAL detector

The photon flux of a single FEL pulse is high enough to successfully measure the single diffraction pattern for ptychography reconstruction. This can significantly decrease the data acquisition time even with several diffraction patterns measured per position. However, efficient use of the high repetition rates requires the detector to have a high frame rate compatible to the pulse repetition rate of the FEL itself.

In addition, a high-resolution ptychograpjhy imaging requires a very high dynamic range. The detector must simultaneously measure high flux in the central region to reliably extract large-scale sample features and single highly scattered photons to ensure a high-resolution reconstruction.

All these requirements are realized in the PERCIVAL detector [122], which was utilized for the measurements described in this section. The parameters of the PERCIVAL detector are summarized in table 5.2, while the details of its design and performance can be found in [123]. The PERCIVAL detector has a high frame rate of up to 300 Hz, a sensor size of 1484 $\times$ 1408 pixels with a pixel size of 27 $\times$

Parameter	**Value**
Pixel size	$27 \times 27\ \mu m$
Sensor size	$1484 \times 1408\ pix$
Scan size	11×11
Frame rate	$300 Hz$
Pixel Well Depth	$3.5 Me^-$
Pixel noise	$16 e^-$

Table 5.2: Characteristics of the PERCIVAL detector.

$27\ \mu m^2$, and a high dynamical range, thus allowing to fully utilize the potential of the ptychography imaging at SASE FELs. The PERCIVAL detector was used for the classical ptychography measurements described in this section as well as for the single-shot ptychography measurements described in Chapter 6.

5.3.2 Experiment

The experiment was performed at beamline FL24 at FLASH2 at a wavelength of 13.5 nm (91.8 eV). The key parameters of the experiment are summarized in table 5.3. The overall experiment was performed similarly to the one described in Section 5.2 (Figure 5.3). However, several setup parameters were changed to improve the achievable resolution, and better utilize the large chip size of the PERCIVAL detector (the experimental setup is shown in Figure 5.13 (A) with the changed parameters highlighted in turquoise). First, a focusing distance of the KB was shortened to 2 m to maximize a divergence of the beam and thus increase the footprint of an unscattered beam at the detector plane, reducing the required dynamical range. Additionally, the sample was placed closer (5 cm downstream) to the focal plane, thus reducing the probe size at the sample down to $\sim 50\ \mu\ m$. This resulted in the increased oversampling in the sample plane (can be calculated as in Equation 5.6). Accordingly, the size of the beam-defining aperture was reduced to $50\ \mu\ m$ to provide the support and limit potential beam jitter. Finally, the detector was moved closer to the sample (1.35 m downstream) to maximize the highest measurable scattering angle and thus increase the maximum achievable resolution.

The measurement procedure was similar to the one described in Section 5.2.

Parameter	**Value**
Wavelength	13.5 nm (91.8 eV)
Pulse energy	60 μJ
Probe size	50 um
Scan size	11×11
Overlap	70%
Images per position	10
Focus sample distance	50 mm
Sample detector distance	1350 mm
Detector size	1484×1408 $pixels$
Detector pixel	27×27 μm^2
Cropped ROI	1200×1200 $pixels$
Numerical resolution	562 nm

Table 5.3: Parameters of classical ptychography experiment performed with the PERCIVAL detector.

The Siemens star sample was scanned in an 11×11 positions rectangular grid (green rectangle in Figure 5.13 (B) with 10 images measured at each scan position. The high frame rate of the PERCIVAL detector allowed for measuring the diffraction images with a FLASH2 repetition rate of 10 Hz. This allowed achieving the measurement time of 2 minutes which, together with the motor scanning time, resulted in a total time of 6 $min/scan$.

Also, several diffraction patterns were later measured with the same beam-defining aperture by the Andor iKon M detector to compare the Andor and the PERCIVAL measurement capabilities. The representative diffraction patterns measured with the PERCIVAL and Andor are shown in Figure 5.13 (C, D), respectively. It is clearly seen that the PERCIVAL detector is capable of measuring higher scattering angles due to its bigger chip size and dynamical range.

5.3.3 Reconstruction and results

The ptychography reconstruction from the data measured with the PERCIVAL detector was performed by the AD-based ptychography algorithm utilizing the forward model described in Section 5.1. The reconstruction lasted for 1000 iterations lead-

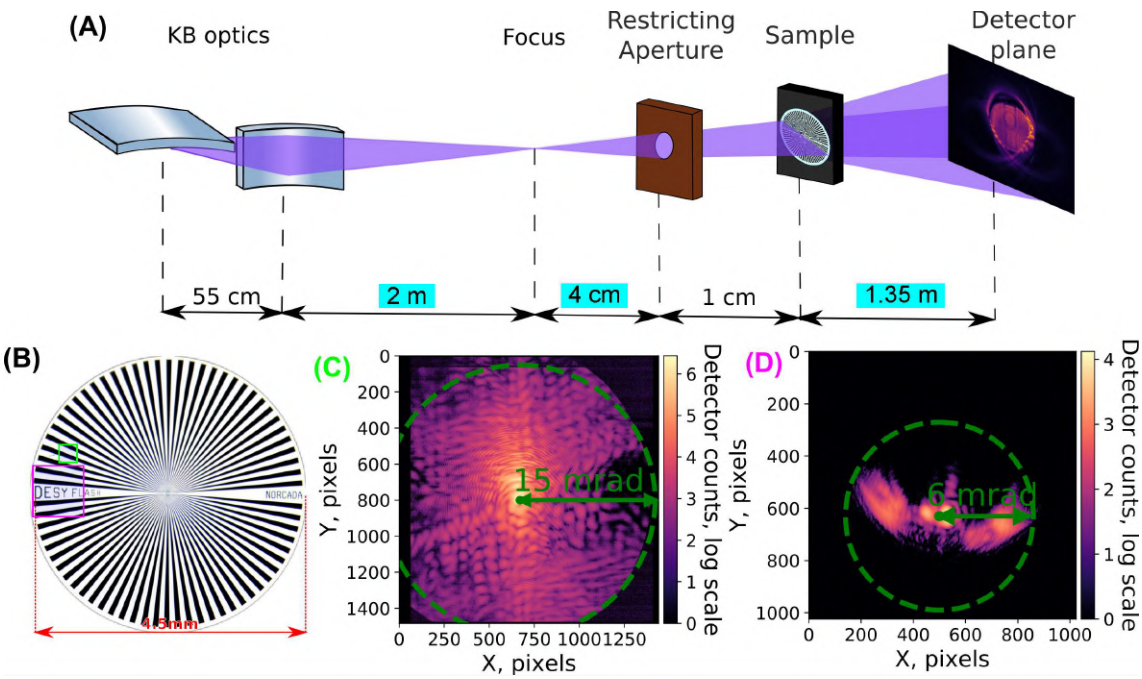

Figure 5.13: **Ptychography with the PERCIVAL detector. (A)** Outline of the experimental setup. The overall setup is similar to the one used in the Section 5.2 with the key differerences highlighted in turqoise. **(B)** image of the sample with scanned areas shown with a green rectangle. **(C),(D)** Typical diffraction patterns measured with the PERCIVAL and Andor iKon M detectors. The highest measured scattering angle is shown with dashed circles.

ing to the reconstruction shown in Figure 5.14(A). To estimate the improvements achieved with the PERCIVAL detector, these results were compared with the reconstruction obtained from the data measured by the Andor iKon L detector during the experiment described in Section 5.2 shown in Figure 5.14 (B).

The resolution of the retrieved objects was estimated using the FRC as shown in Figure 5.15 to compare the quality of the reconstructions. The resolution in Figure 5.15 compared only in terms of the spatial frequencies since the different experimental setups resulted in the different numerical resolutions making their direct comparison not meaningful. The FRC curves, shown in Figure5.15, demonstrate a considerable improvement of the resolution achieved by utilizing the PERCIVAL detector.

In conclusion, the results presented in this section demonstrate how the higher

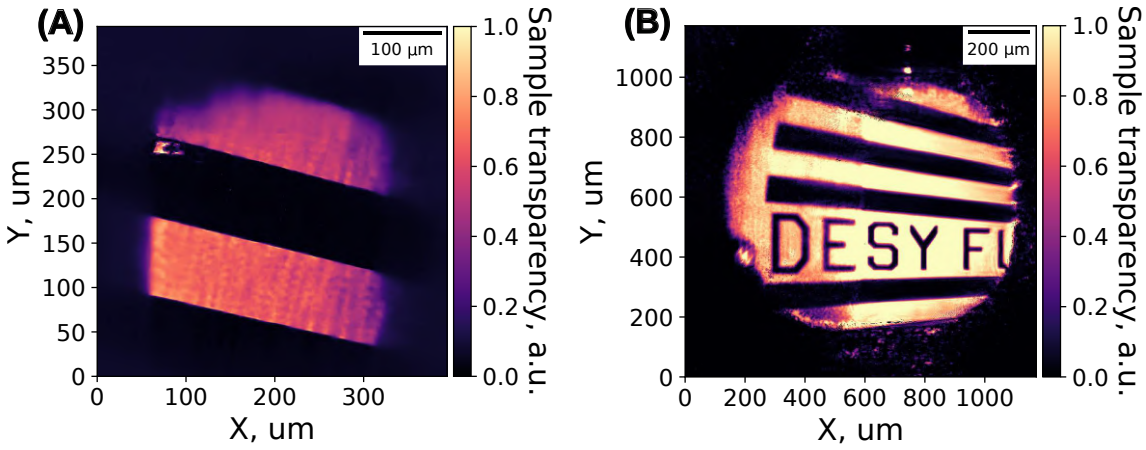

Figure 5.14: **Results of the ptychography reconstruction.** Sample modulus reconstructed from the data measured with the **(A)** PERCIVAL and **(B)** Andor iKon M detectors.

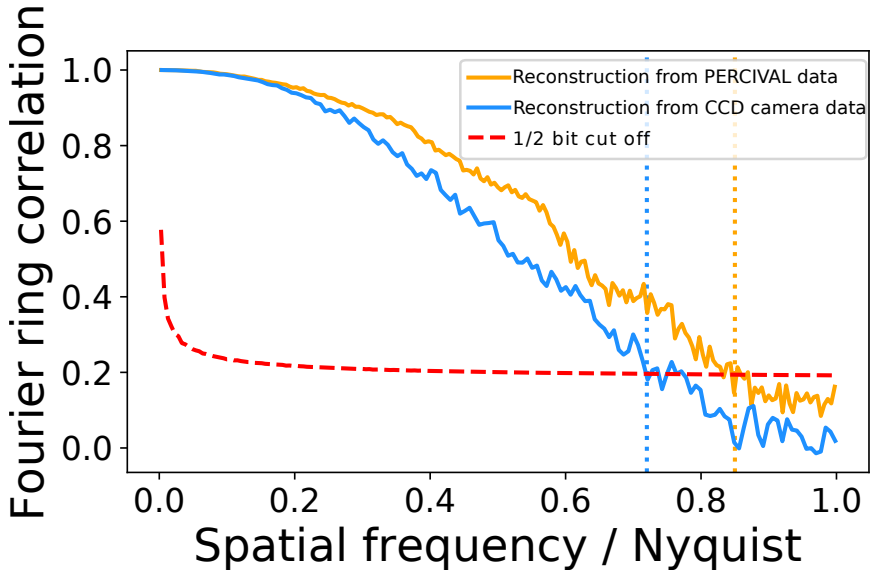

Figure 5.15: **Comparison of the reconstruction resolution.** The orange and blue FRC curves are calculated for the reconstructed object modulus shown in Figure 5.14 (A) and (B) rescpectively. The red dashed line represents the 0.5 bit threshold.

dynamical range and larger detector chip size enable the registration of photons scattered to higher angles, thus improving the achievable resolution of reconstruction. Additionally, the higher frame rate enabled faster measurements and better utilization of the FLASH pulses. Both of these improvements simplify the measure-

ments and facilitate the routine use of classical ptychography at SASE FELs for high-throughput high-resolution imaging.

5.4 Imaging of plasma-treated plastic films

In this experiment, ptychography was used to study the effects of the different plasma treatment regimes [124] on the morphology of thin plastic (poly (ethylene terephthalate) (PET)) films. PET plastic is one of the most widely used in industry due to its hardness, thermal stability, chemical resistance, and formability [125]. However, its hydrophobic properties are disadvantages in applications related to adhesion [125]. This problem is typically solved by various treatment strategies aimed to produce a surface modification of the plastic.

In this experiment, the effects of two different plasma treatment strategies (corona and FLAIR® [126]) were studied and compared with the untreated plastic sample. The setup used for the treatment of the PET sample is shown in Figure 5.16. During the corona treatment procedure (Figure 5.16(A)), the plastic film is placed on a grounded roller supporting the moving film at a distance of $1 - 2$ mm from a high-voltage electrode [127].

Either the roller or the electrode or both are covered with a dielectric coating. A high voltage of the order of 10 kV applied to the electrode creates a plasma in the air gap between the electrode and the roller. The polymer film passing through the air gap is treated by the plasma on the side facing the electrode. The power density of the plasma generated between the high-voltage electrode and the roller in an industrial corona treater is limited to a few W/cm^2. In turn, the FLAIR® treatment (Figure 5.16(B)) scheme is structured similarly to the corona treater but operates with power densities of up to several 100 W/cm^2. In this power density range, the plasma species in the air plasma differ significantly from those in a corona treatment plasma, and different effects on the morphology and chemical composition of the polymer film can be expected [126], [128]. In this study, PET film samples were treated with corona and FLAIR® plasma at a power density of 5 W/cm^2 and 200 W/cm^2, respectively. In both cases, the total energy per area of the plasma-treated film was kept the same at 180 kJ/m^2. This was achieved by varying the

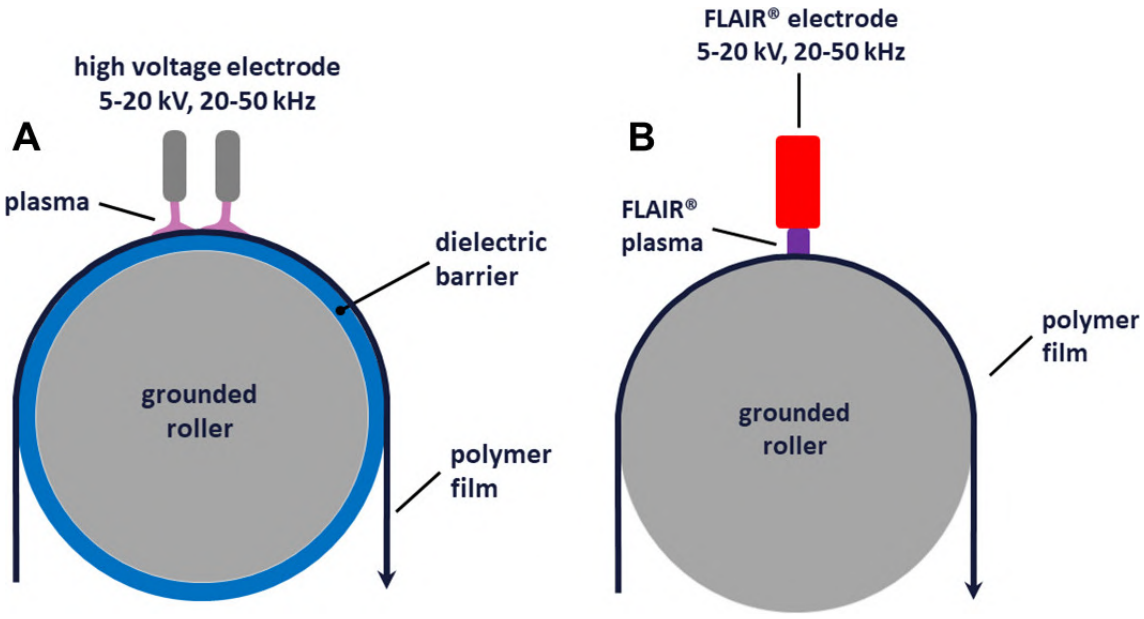

Figure 5.16: **Plasma treatment setups** for corona **(A)** and FLAIR® [126] **(B)** treatment of PET films. In both setups the polymer film is moved between the grounded roller and the high-voltage electrode through plasma produced by the discharge. Corona plasma **(A)**has lower density ($\sim 5\ W/cm^2$) than the FLAIR® one **(B)** ($\sim 200\ W/cm^2$) resulting in the different morphological modifications of the sample. Adapted from [29].

speed of the film and thus the treatment time: 3.6 s in the case of corona and 0.09 s in the case of FLAIR®.

The efficiency of these methods can be monitored by various methods such as atomic force microscopy (AFM) and scanning electron microscopy (SEM). However, both of these methods have a relatively limited field of view, thus making studies of a large surface area a very tedious task. In turn, the X-ray ptychography in the intermediate or far-field regime is capable of obtaining high-resolution (down to nm) images of the extended areas (up to several mm^2 depending on the time of the scan) of the sample.

This section describes the first application of X-ray ptychography for imaging of the industrially relevant samples performed at FLASH. The technique was used to study the morphological changes on PET films treated with two different types of individual plasma surface treatment. The 2D thickness distribution was reconstructed directly through the reconstruction of the refractive index of the sample,

allowing studies of the sample surface morphology. This section gives details on the *thickness* representation of the sample and describes the ptychography experiment performed at FLASH. Finally, it presents the details of the reconstruction process and analyzes the obtained results.

5.4.1 Direct reconstruction of the sample thickness

During the reconstruction of the plastic samples, it was assumed that the plasma treatment greatly changes the sample morphology and almost does not affect the chemical composition of the sample. Thus the average chemical composition of sample bulk can be assumed to be homogeneous and unchanged by the treatment. Therefore the effective refraction index of a sample depends only on the sample thickness variations. Additionally, the sample thickness z_s fulfilled the projection approximation [62] formulated as

$$z_s \leq \frac{5.2(\delta x)^2}{\lambda}, \tag{5.14}$$

where δx is the maximum achievable resolution of the reconstruction and λ is the illumination wavelength. All these factors allow expressing the sample transfer function according to equation 4.2 as

$$O = \left(\exp\left(-k\beta_{PET} z_{(x,y)}\right) \cdot \exp\left(-ik\delta_{PET} z_{(x,y)}\right)\right), \tag{5.15}$$

where β_{PET} and δ_{PET} are the component of the refraction index of PET at the particular wavelength, and the $z_{(x,y)}$ is the 2D distribution of the sample thickness. This *thickness* representation reduces the number of unknowns in the sample function and also introduces the positivity constraint on $z_{(x,y)}$ improving the convergence of the reconstruction. Additionally, it results in the direct reconstruction of the sample thickness allowing its quantitative characterization. It is worth mentioning that AD allows the use of the *thickness* representation without the need to re-derive any gradients allowing the fast adaptation of the reconstruction algorithm for the needs of the particular experiment.

5.4.2 Experiment

The experiment was performed at beamline FL24 at FLASH2 at a wavelength of 2.66 nm (466 eV) (third harmonic of 8 nm). The main experimental parameters are summarized in table 5.4, and the experimental setup is shown in Figure 5.17.

Parameter	**Value**
Wavelength	2.6 nm
Pulse energy	90 μJ
Probe size	50 um
Scan size	20×20
Overlap	90%
Images per position	10
Focus sample distance	120 mm
Sample-detector distance	660 mm
Detector size	1024×1024 $pixels$
Detector pixel	13×13 μm
Cropped ROI	800×800 $pixels$
Numerical resolution	160 nm

Table 5.4: Parameters of classical ptychography plastic measurements.

The experiment was performed similarly to the one discussed in Sections 5.2,5.2. The KB mirrors were set to focus the beam 2.1 m after the second mirror to increase the beam divergence. The sample was placed 12 cm downstream of the focal plane to increase the beam diameter on the sample to 50 μm and enlarge the resulting FOV of ptychography reconstruction. An ANDOR iKON-M detector (1024×1024 pixels, $13 \times 13 \mu m^2$ each) was placed 65 cm downstream of the sample to record the diffraction patterns.

The experiment consisted of the measurement of three different plastic samples (treated by corona discharge, treated by FLAIR®, and control untreated one) and the test Siemens star object, which was used to obtain the preliminary probe function through the ptychography beam characterization. During the measurements, the samples were scanned along a 20×20 rectangular scan position grid with 10 images measured at each scan position. A large number of scan points, together with the beam diameter of 50μ m at the sample plane, resulted in an increase in the reconstruction FOV to 145×145 μm^2. During the measurements, the FEL beam was attenuated utilizing the gas attenuator to minimize the heat load on the sample.

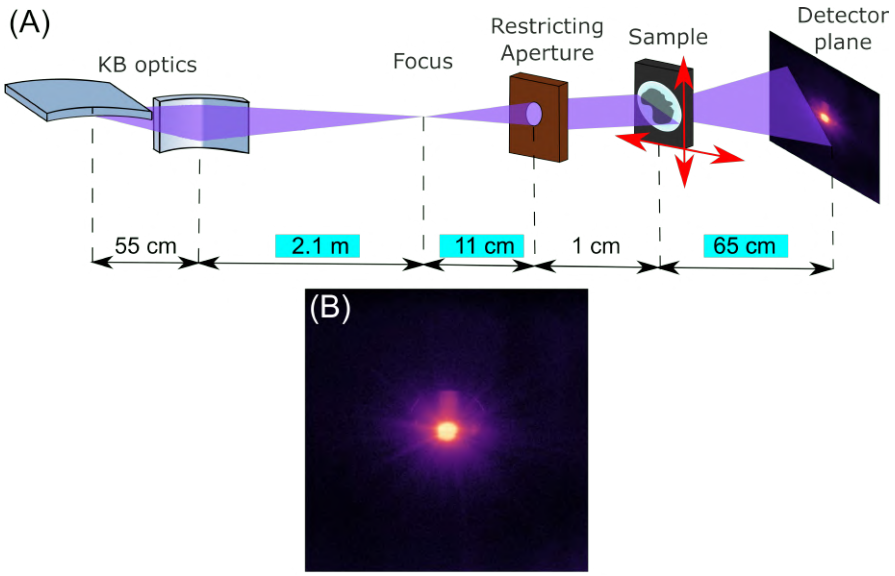

Figure 5.17: **Experiment setup used for ptychographical measurement of plastic film morphology. (A)** Outline of the experimental setup. The overall setup is similar to the one used in the Sections 5.2,5.3 with the key differerences highlighted in turqoise.**(B)** Typical diffraction pattern measured from the plastic sample. Adapted from [29].

The radiation of the fundamental wavelength was suppressed with the metallic foil (Zr 200 nm and Al 400 nm) filters.

5.4.3 Reconstruction and results

The reconstruction was performed by the AD-powered ptychography engine described in Section 5.1. During the first step, a reconstruction of the Siemens star sample was performed. The reconstruction took 500 iterations utilizing 4 GPUs NVIDIA P100 in parallel with an iteration time of 1 *sec*. The initial probe was obtained by the back-propagation of the average measured intensity pattern. The probe modes were initialized as the orthonormal vectors obtained from the SVD of the probe matrix.

The reconstructed main probe modes and their average modal weights were used to initialize the reconstructions of the plastic samples. The sample *thickness* representation (equation 4.2) was used to directly reconstruct the thickness of the plastic

samples. Each of the reconstructions took 300 iterations, with the probe modes kept constant during the first 100 iterations. The resolution of the reconstruction of plastic samples was estimated as 600 nm by the FRC.

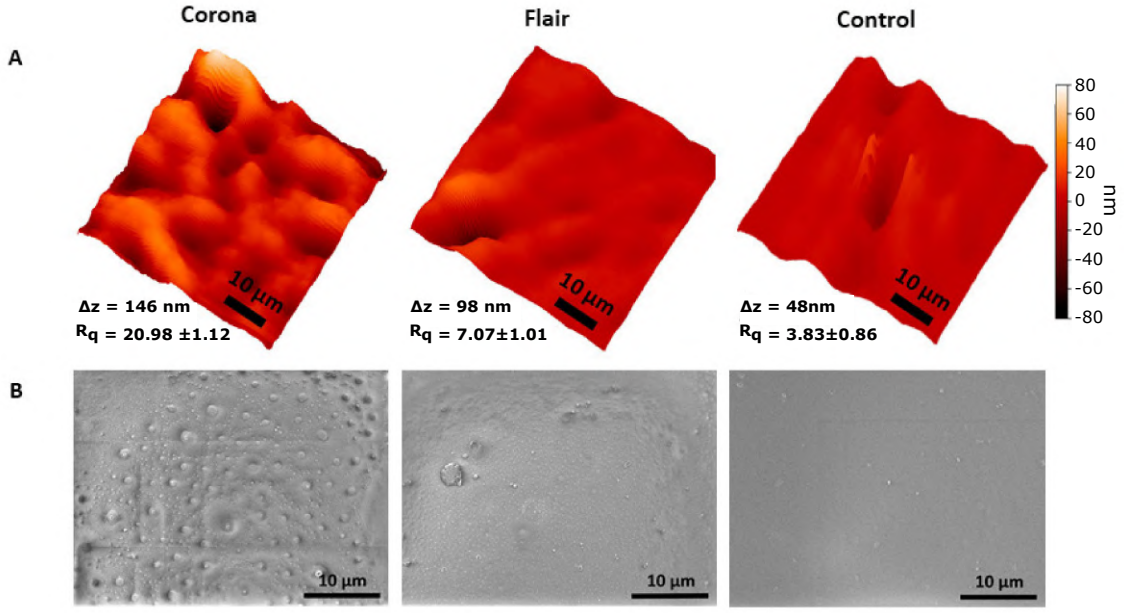

Figure 5.18: **Comparison of images of plastic films. (A)** Ptychography and SEM **(B)** images of plasma-treated and control PET films. Δz denotes the maximum peak-to-valley difference. R_q denotes the root-mean-square (RMS) roughness [129]. Adapted from [29].

$50 \times 50 \mu m^2$ regions of reconstructed 2D thickness distribution of the plastic samples are shown in Figure 5.18 (A). The changes in the surface morphology of the different samples are clearly seen. The obtained results are in good agreement with the one provided by the SEM (shown in Figure 5.18 (B)). This indicates the good performance of the ptychography for imaging the surface morphology of plastic samples.

Both techniques showed nearly constant morphological changes after corona treatment (root-mean-square (RMS) roughness [129] $R_q = 20.98 \pm 1.12$ peak-to-valley difference $\Delta z = 146\ nm$). The FLAIR® treatment resulted in smoother surface features ($R_q = 7 \pm 1.01$, $\Delta z = 98\ nm$). The contol (untreated) sample showed the smoothest surface ($R_q = 3.83 \pm 0.86$, $\Delta z = 48\ nm$) as expected. Additionally the high throughput of the ptychography at FEL allowed to image 100s of

μm^2 of sample surface in a matter of minutes potentially enabling the imaging of much bigger (10s mm^2) samples.

To summarize, the results presented in this section describe the use of AD-based ptychography at a SASE FEL for direct reconstruction of the sample thickness. This method allows the study of the surface morphology of extended samples (up to mm^2) with high resolution (down to nm). This distinguishes ptychography from other imaging methods such as atomic force microscopy [130]. Ptychography is able to image a large surface area and thus provide more reliable quantitative information on the morphology of the surface.

Moreover, the direct reconstruction of the 2D distribution of the index of refraction of the sample is not limited to chemically homogeneous samples. It potentially can be used for quantitative analysis of the chemical composition of extended samples. In the future, this method can be combined with reflective ptychography [75] allowing separate quantitative studies of the chemical composition and morphology for chemically diverse samples. Additionally, this method is not limited to FELs and can be utilized at any coherent photon source.

5.5 Conclusion and outlook

This chapter describes the forward model for an AD-based ptychography engine capable of performing classical ptychography at SASE FELs. The model can work with simultaneously partially spatially-coherent and spatially-fluctuating SASE FEL pulses by describing them in the basis of incoherent probe modes with shot-to-shot unique modal weights. This capability is demonstrated by performing the first classical ptychography experiment at FLASH2. During this experiment, the transfer function of the sample was successfully reconstructed, despite the spatial fluctuations and partial coherence of the probe. Additionally, the reconstructed main probe modes and shot-to-shot unique modal weights allowed to perform a beam characterization without aperturing the beam or selecting its most coherent part.

The measurements performed with the PERCIVAL detector demonstrated both faster measurements and improvement of the achievable resolution. Such improve-

ments will help exploit the high-flux SASE FELs pulses better, potentially enabling high-resolution imaging of large extended samples in just a few minutes.

Finally, the ptychographic measurement of plasma-treated plastic films was the first applied ptychography experiment at FLASH. In this experiment, the AD-powered ptychographical framework showed its ability to directly reconstruct the sample thickness utilizing its refractive representation in the forward model. This change required no manual gradient derivation due to the flexibility provided by the AD.

To summarize, this chapter describes how the novel AD-based ptychography model enables routine ptychographical measurements at SASE FEL. This technique can be used both for the high-resolution imaging of the sample and for the shot-to-shot beam characterization. The FLASH facility would greatly benefit from such beam-characterization capabilities by its use for optimization of the beamline optics performance and investigating the machine properties. Moreover, this model is not limited only to the SASE FELs and can be used at other photon sources, including synchrotrons, seeded FELs, and HHG-based ones.

However, classical ptychography is still a scanning technique. This prevents its use for single-shot imaging and thus full exploitation of the ultra-short FEL pulses for time-resolved imaging. The next chapter presents a forward model and an experimental setup enabling single-shot ptychography measurements at FELs.

Chapter 6

Single-shot ptychography at SASE FELs

Ptychography relies on the dense scanning of a sample to obtain its high-resolution reconstruction. The scanning procedure forms the core of the experimental technique and cannot be entirely avoided either by using faster detectors or the brightest and shortest FEL light pulses. The femtosecond length of the FEL pulses can enable imaging of dynamical processes with exceptional temporal resolution. However, classical ptychography relying on scanning cannot capture the dynamics of systems. Therefore, an X-ray compatible implementation of single-shot ptychography can significantly enhance the potential of imaging at FEL sources enabling diffraction-limited time-resolved imaging of extended samples and dynamics of complex matters at X-ray wavelengths.

Single-shot ptychography was first proposed by Sidorenko et al. [1] and was further expanded to Fourier [131], multi-slice [132] and multi-wavelength [133] regimes in the visible range. A single-sot ptychography setup proposed in [1] is shown in Figure 6.1. The setup, based on a 4-f lens arrangement, is capable of performing simultaneous illumination of a sample by multiple beamlets in one shot by utilizing a pinhole array placed in the front focal plane of the first lens as a light source. The sample is placed close to the common focal plane of the two lenses. The overlap of the beamlets produced by the individual pinholes can be adapted by adjusting the focus-sample distance. However, this concept cannot be directly transferred to the X-ray regime due to the lack of efficient high numerical aperture refractive optics at

these wavelengths.

This Chapter describes an alternative X-ray compatible setup for single-shot ptychography inspired by the work of Pan et al. [134] (Section 6.1.1) and the differentiable forward model (Section 6.1.2.) used in the AD-based ptychography framework for the single-shot ptychography reconstructions. Section 6.3.3 presents the results of the first single-shot ptychography experiment in the X-ray range performed at FLASH. Finally, Section 6.3.4 concludes this chapter with discussion of the possible improvements in single-shot ptychography and its future perspectives at FELs. The text of this Chapter closely follows the work available as a preprint at [30].

6.1 X-ray compatible single-shot ptychography

Single-shot ptychography removes the need for scanning by recording the entire set of diffracted intensities simultaneously. It utilizes a single light pulse instead of con-

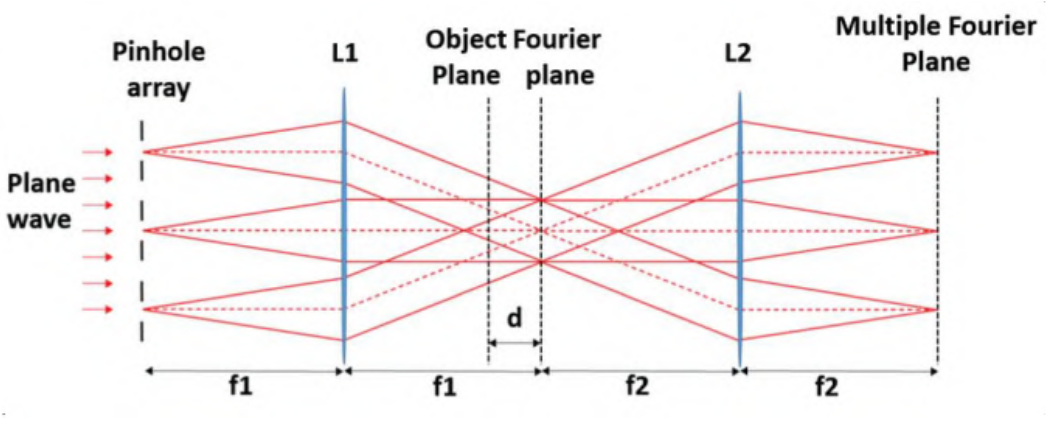

Figure 6.1: **Single-shot ptychography setup based on a 4-f lens system.** Lens $L1$, with focal distance $f1$, focuses the light beams produced by an array of pinholes. The object is located in the vicinity of the common focal plane of $L1$ and $L2$. The distance d between the focal plane ad the object plane determines the overlap in the sample plane and the total scanned area of the sample. Lens $L2$, with focal distance $f2$, focuses the light diffracted from the object to the detector resulting in a far-field geometry. Adapted from [1].

secutively measuring diffraction patterns from different sample positions. This can be realized by splitting the probe P into several beamlets B_j that mutually overlap on the sample surface. For successful phase retrieval in single-shot ptychography, two key prerequisites need to be met for the measured intensity distribution.

First, a functional dependence has to exist between the probe and the beamlets: $B_j = g_j(P)$, which links the initial probe P with the j-th beamlet B_j and all the beamlets together. Due to this dependency, the forward model can still be formulated with respect to the probe P. Thus, there is no need to optimize each of the beamlets B_j individually, which dramatically reduces the number of optimizable variables. Additionally, the neighboring beamlets B_j should have a certain degree of overlap in a sample plane, preventing the reconstruction ambiguities and ensuring the convergence by the ptychographical oversampling [80], [135].

As a second prerequisite, the diffracted intensity of each of the exit-waves of the beamlets at the detector plane ideally should be localized and separable from its neighbors. This minimizes inter-beamlet interference. It allows the use of the forward model ignoring interference of the neighboring beamlets. Thus, each localized intensity $I_j \in I$ can be attributed to a unique beamlet. When both conditions are fulfilled, the measured intensity can be calculated as follows:

$$I_{meas} = \sum_{j}^{N} I_j = \sum_{j}^{N} \left| \mathcal{P} \left\{ B_j \cdot O_{\mathbf{r}_j} \right\} \right|^2 = \sum_{j}^{N} \left| \mathcal{P} \left\{ g_j(P) \cdot O_{\mathbf{r}_j} \right\} \right|^2. \tag{6.1}$$

During the reconstruction, the entire measured intensity pattern I_{meas} is separated into the individual sub-patterns I_j attributed to the particular beamlet B_j hitting the sample at position $O_{\mathbf{r}_j}$. This allows forming the "classical-like" ptychography dataset consisting of diffraction patterns measured from the partially overlapping positions at the sample plane. Then, in the reconstruction process, the probe P and the sample O are retrieved by minimizing the difference between the measured I_j and the approximated intensities $\tilde{I}_j$ for each of the individual sub-patterns.

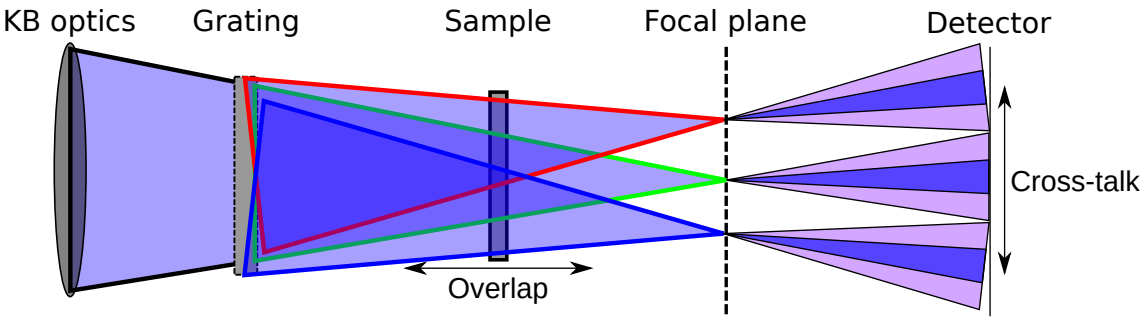

Figure 6.2: **X-ray compatible single-shot ptychography setup** based on a beam splitting grating. The FEL beam is focused downstream of the sample by the KB mirrors. The beam-splitting grating is placed in front of the sample and splits the beam into several diffraction orders (shown in red, green, and blue). The sample, placed in the vicinity of the grating, is illuminated by the overlapping beamlets produced by the grating. The degree of overlap and the field of view in the sample plane can be controlled by changing the grating-sample distance. The detector is placed downstream of the focal plane. The focal length of the KB mirrors is selected to simplify the separation of the beamlets (pink) and minimize the possible cross-talk between the radiation scattered by the sample (light-pink). The setup does not use refractive optics, which allows its effective use in the X-ray range. Adapted from [30].

6.1.1 Experimental setup

Both key criteria are realized in the proposed X-ray compatible setup shown in Figure 6.2. An FEL probe is focused downstream of the sample by KB mirrors. The probe is split by a 2D diffraction grating placed in the vicinity of the sample. The grating splits the incoming FEL beam into a number of beamlets (shown in red, green, and blue in Figure 6.2) partially overlapping in the sample plane. The degree of overlap of the beamlets and the FOV (i.e., the total area of illumination on the sample) is determined by the specific design of the grating and the grating-sample distance. In turn, the inter-beamlet interference and the footprint of the individual beamlets in the detector plane are minimized by varying the focal length of the KB. The setup does not use any refractive optic and thus can be used with X-ray wavelengths. A detailed description of the grating parameters can be found in Section 6.2.

The resolution of the single-shot ptychography imaging with the proposed setup

is limited by the highest scattering angle registered for a particular beamlet. The simple 2D diffraction grating results in an illumination arranged in a $x \times x$ square. In this case, the maximum achievable resolution can be estimated as $\frac{\lambda z x}{N p}$, where λ is the wavelength, z is the sample-detector distance, N is the number of pixels of the detector, and p is the detector pixel size [47]. The number of measurable localized diffraction patterns is set by the combination of the beam-splitting grating and the detector parameters, which defines the trade-off between the FOV, the maximum achievable resolution, and the ptychographic oversampling [80], [135].

6.1.2 Forward model

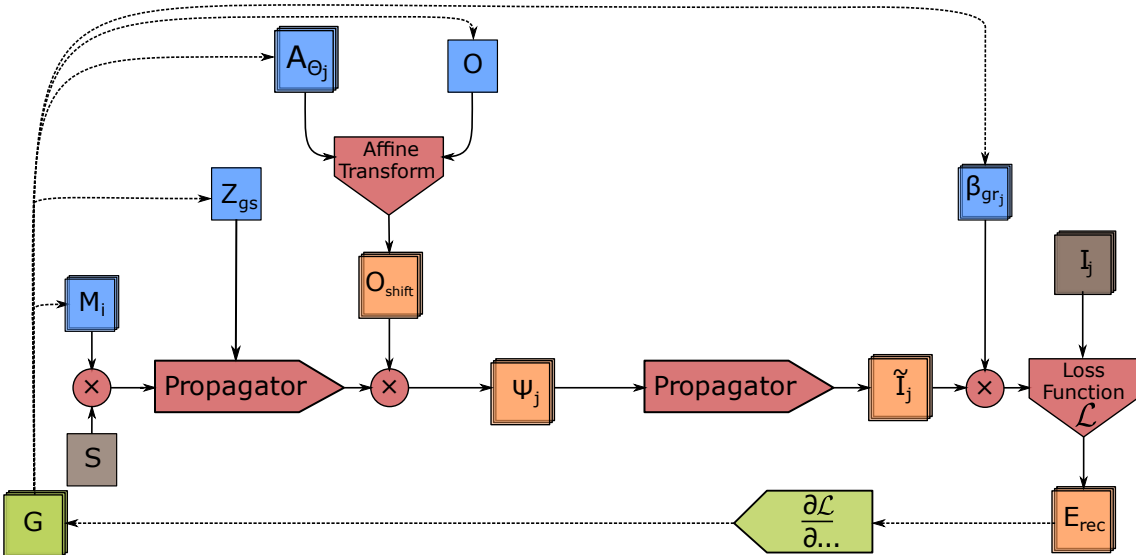

Figure 6.3: **Computational graph of the forward model of single-shot ptychography.** Input data (measured intensities - I_j and numerical support S) colored in brown, operators colored in red, optimizable variables (probe modes - M_i , grating-sample distance z_{gs}, affine transformation parameters - A_{Θ_j}, object function - O and diffraction orders efficiency - β_{gr_j}) colored in blue, intermediate variables colored in orange, backward propagation of error marked with dotted line and green color.

The forward model describing the proposed setup is shown in Figure 6.3. The

model approximates the intensity I_j produced by j-th beamlet B_j as follows:

$$\tilde{I_j} = \beta_{gr_j} \sum_i \left| \mathcal{P}_2 \left\{ \mathcal{A}_{\theta_j}(O) \cdot \mathcal{P}_1 \left\{ M_i \cdot S \right\} \right\} \right|^2, \tag{6.2}$$

where $\mathcal{P}_1$ is the Fresnel TF propagator describing the propagation of the probe from the grating plane to the object plane, $\mathcal{P}_2$ is the two-step propagator in the intermediate field describing the propagation of the exit-wave from the sample plane to the detector plane, S is the support representing the active area of the grating, M_i are the probe modes used to describe the partially coherent radiation $\mathcal{A}_{\theta_j}$ represents an affine transform describing the sample positions $O_{\mathbf{r}_j}$ as described in Section 4.3, and β_{gr_j} is the scaling coefficient representing grating efficiency for j-th diffraction order. $\mathcal{P}_1$ and $\mathcal{P}_2$ is implemented as in equations 2.10 and $\mathcal{P}_2$ is implemented as

$$\begin{aligned} \mathcal{P}_2(\Psi_{\rho_s}) = \Psi_{\rho_d} = -\frac{i}{\lambda z_{id}} \exp(i\frac{k}{2z_{id}}\rho_d^2)\cdot \\ \mathcal{F}\left\{ \exp(i\frac{k}{2z_{id}}\rho_i^2) \cdot \left[-\frac{i}{\lambda z_{si}} \exp(i\frac{k}{2z_{si}}\rho_i^2) \cdot \mathcal{F}\left\{ \exp(i\frac{k}{2z_{si}}\rho_s^2)\Psi_{\rho_s} \right\} \right] \right\}, \end{aligned} \tag{6.3}$$

where Ψ_{ρ_g},Ψ_{ρ_s},Ψ_{ρ_d} are complex wavefields in the grating, sample, and detector planes respectively, z_{gs}, z_{si}, and z_{id} are grating to sample, sample to intermediate plane, and intermediate plane to detector distances respectively, $k = \frac{2\pi}{\lambda}$ is the wave number, $\rho = (x, y)$, ρ_g, ρ_s, ρ_i and ρ_d denote the transverse coordinates at the grating, sample, intermediate, and detector planes respectively. The use of the two-step propagator for calculation of the intensity distribution in the detector plane allowed to have a higher numerical resolution of 2.1 μm and 1.9 μm for the Andor and PERCIVAL detectors, respectively.

Contrary to the forward model of the classical ptychography described in Section 5.1, the forward model of the single-shot ptychography optimizes the probe modes M_i but not the modal weights. All the beamlets are produced from the same probe and thus share their coherent properties. Also, the single-shot model utilizes a TF propagator (Equation 2.13) to describe the light propagation from the grating to the sample plane. This is done to directly optimize the grating-sample distance influencing the beamlet shape at the sample plane. This optimization can be done during the ptychography reconstruction improving the convergence and the

achievable resolution. Additionally, the single-shot forward model includes the optimizable intensity scaling coefficients to accommodate for the varying efficiency of the different diffraction orders.

As a result, the following parameters are optimized during the reconstruction: object function O, probe modes M_i, scan coordinates $\boldsymbol{r}_j$ expressed through the affine parameters θ_j, grating-sample distance z_{gs} included into the TF propagator $\mathcal{P}_1$, and order-specific diffraction grating efficiency α_j.

6.2 Experiment

The proposed experimental setup and the forward model were tested in a proof of principle single-shot ptychography experiment performed at the FLASH2 beamline FL24 [33], [51] at wavelength $\lambda = 13.5\ nm$ corresponding to the photon energy $E_{ph} = 91.8\ eV$ with the average pulse energy of 85 μJ. The experimental setup is shown in Figure 6.4. Metal foil filters in combination with a gas attenuator [33], [136] were used to attenuate the beam to prevent sample damage and detector saturation and to suppress higher FEL harmonics. The FEL beam was focused using a pair of bendable KB mirrors [137] to a focal spot position of 9 cm behind the sample. A 2D beam-splitting transmission-grating (produced by Norcada inc) with a period of 2.930 μm, an opening size of 0.732 μm, and an active area of $200 \times 200\ \mu m^2$ was placed 210 cm after the center of the last KB mirror. The grating parameters were optimized to equalize the resulting diffraction efficiency of the grating orders to reduce the required dynamical range of the detector. This optimization of the grating allowed to simultaneously measure diffraction from $(0, 0)$ up to $(2, 3)$ grating order depending on the detector chip size. The grating was made from a 200 nm thick gold layer electroplated on a 50 nm silicon nitride membrane. The active area of the grating was smaller ($\sim 1/4$) than the beam cross-section at the grating plane to select the most intense and spatially coherent part. The sample was placed 9 mm downstream of the grating, which led to an average mutual overlap of 80% between neighboring beamlets in the sample plane. The diffraction patterns were measured using two detectors: an Andor iKON-M SO CCD camera (1024×1024 pixels, $13 \times 13 \mu m^2$ each) and the PERCIVAL detector (1440×1484 pixels, $27 \times 27 \mu m^2$

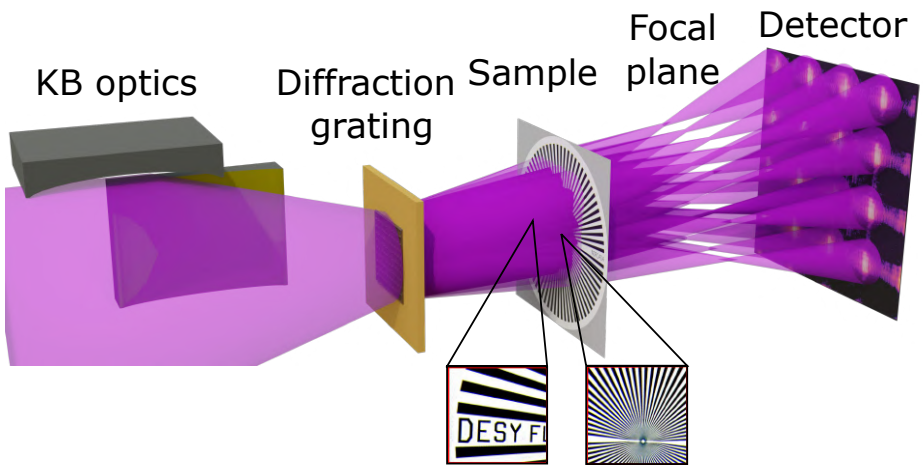

Figure 6.4: **Single-shot experimental setup.** The FEL beam was focused 9 cm after the sample using bendable KB mirrors. The beam-splitting 2D diffraction grating was placed 210 cm after the center of the last KB mirror. The sample was placed approximately 0.9 cm behind the grating and thus illuminated by the overlapping beamlets produced by the grating. The two areas of the sample *letters* and *stripes* are shown in the insets. The Andor iKon-M SO CCD was placed in the intermediate field 66 cm downstream of the sample. Alternatively, the PERCIVAL detector was placed in the intermediate field 135 cm downstream of the sample. Adapted from [30].

each) [122]. The sample-to-detector distances were selected to fit at least 4×4 beamlets onto the detector chip. It was 66 cm for the Andor and 135 cm for the PERCIVAL, respectively. The diffraction data were measured from the *letters* and *stripes* areas by Andor and PERCIVAL detectors, respectively.

6.3 Reconstruction and results

6.3.1 Raw data segmentation

In single-shot ptychography, measured intensities must be properly tessellated and then segmented into separate diffraction patterns, each attributed to a single beamlet. Later, the patterns are placed in equal-sized computational frames used for the numerical light propagation. This way, it is possible to obtain a "classical-alike" dataset, which can be used for the reconstruction. A variable angle between the

grating and detector axes which is inevitable during the experiments prevents the use of any tessellation strategy based on a fixed grating-detector geometry. This is the reason the segmentation routine based on Voronoi tessellation as proposed by Barolak et al. [133] was used in this work.

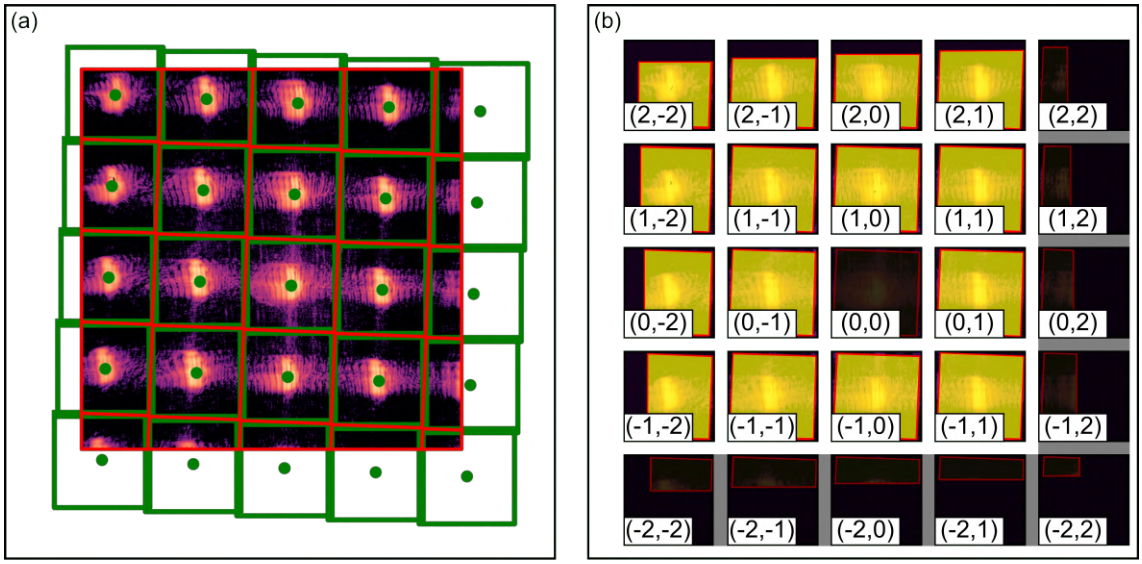

Figure 6.5: **Splitting of the single-shot ptychography data by Voronoi tesselation (A)** Voronoi tessellation of the measured data. Centers of masses of sub-patterns corresponding to each grating order are shown with green dots. Green polygons represent individual Voronoi cells fitting in the area of the detector chip. Green squares show the smallest square simultaneously circumscribed about and concentric with the largest Voronoi cell. **(B)** Split data used for the reconstruction. The area highlighted in blue is inside the corresponding Voronoi cell and is constrained during the reconstruction. Each individual square is equivalent to a measurement from an individual beamlet and represents the individual computational frame used for the propagation of the respective diffraction order. The coordinates of the individual diffraction order at the sample plane can be estimated using the grating-sample distance and grating parameters. All the data pieces with less than half of the pixels measured were ignored during the reconstruction (right- and lower- most squares, dark in **(B)**). The (0, 0) order was also ignored during the reconstruction due to high levels of parasitic scattering from the beam passing through the frame of the grating. Adapted from [30].

This algorithm finds so-called Voronoi cells v_j for a given set of points p_j. Each of the Voronoi cells is defined as a set of detector pixels that are closer to one

particular point than to any other. The process of data segmentation is illustrated in Figure 6.5. The orientation of the grating axes with respect to the detector pixel axes can be found by utilizing the centers of mass of the intensity patterns. These patterns are produced by the individual diffraction orders (shown by the green dots in Figure 6.5). The horizontal $(0, x)$ and vertical $(x, 0)$ diffraction orders form a basis of the grating coordinate system. This basis can be utilized to calculate the coordinates of all the diffraction orders at the detector plane. In turn, these coordinates were used as the input points p_j for the Voronoi tessellation. This results in the Voronoi cells shown with the red polygons in Figure 6.5 (a, b). Each of these cells surrounds a part of the detector in which all the intensity was assumed to be produced by a single beamlet corresponding to the particular diffraction order. The scan coordinates at the sample plane $\boldsymbol{r}_j$ corresponding to each order were estimated from the diffraction grating parameters and the grating-sample distance. In turn, the size of the computational frames for the light propagation was found as the smallest square simultaneously circumscribed about and concentric with the largest Voronoi cell (shown with green squares in Figure 6.5 (a)). In some of the frames, only part of the pixels (fitting inside the particular Voronoi cell shown with the red polygons in Figure 6.5) could be reliably attributed to the corresponding diffraction order. In this case, the area outside the Voronoi cells (shown in black in Figure 6.5) (b)) was left unconstrained during the reconstruction.

6.3.2 Reconstruction

The segmented data was used as the input for the single-shot ptychography reconstruction. The reconstruction was performed using an AD-based ptychography framework with the forward model described in Section 6.1.2. The order-specific diffraction grating efficiencies β_{gr_j} and the main orthogonal modes M_i were initialized using the average grating diffraction pattern measured without the sample. The j-th diffraction efficiency β_{gr_j} was estimated as

$$\beta_{gr_j} = \frac{I_j}{I_0}, \tag{6.4}$$

where I_j is the total intensity diffracted to the j-th order, and I_0 is the intensity of the $(0, 0)$ order. The total intensity I_j attributed to the $j - th$ diffraction order was found as the sum of all intensities measured by the pixels belonging to the corresponding Voronoi cell v_j obtained from the splitting of the raw data.

The approximate complex wavefields of the diffraction orders were calculated at the detector plane utilizing the known defocus produced by the KB. The back-propagated wavefields were orthogonalized using a singular value decomposition and used as the initial values for the M_i. At the same time, the initial coordinates of the beamlets at the sample plane r_j were estimated from the known grating parameters and the grating-sample distance z_{gs}.

The reconstruction took 6×10^3 iterations of the gradient-based optimization utilizing the ADAM optimizer. Due to the small size of the dataset, the computations for all the beamlets fit in the memory of the single NVIDIA Tesla P100 GPU. The average iteration of the reconstruction took 20 ms, resulting in a total reconstruction time of 2 min. The regularization strength coefficients γ_i and the individual learning rates for the optimizable parameters were selected using the method described in Section 4.4 to provide fast and stable convergence.

6.3.3 Results

The raw data with the averaged dark background subtracted are shown in Figure 6.6. Figure 6.6 (a) corresponds to the full image measured by the Andor detector in the *stripes* region of the sample, while Figure 6.6 (b) corresponds to the central part of the image measured from the *letters* region with the PERCIVAL detector. In both images, the intensities produced by the different beamlets are separated with a limited degree of cross-talk present. Reconstructed sample transfer functions and main probe modes are shown in Figure 6.6 (c-f).

The FLASH beam was apertured by the opaque frame of the diffraction grating resulting in the selection of roughly 1/4 of the beam area, increasing the spatial coherence of the probe. However, the multi-modal expression of the probe was used during the reconstruction for its ability to accommodate cross-talk intensity, which in this case was effectively treated as noise [138]. The occupancies of the

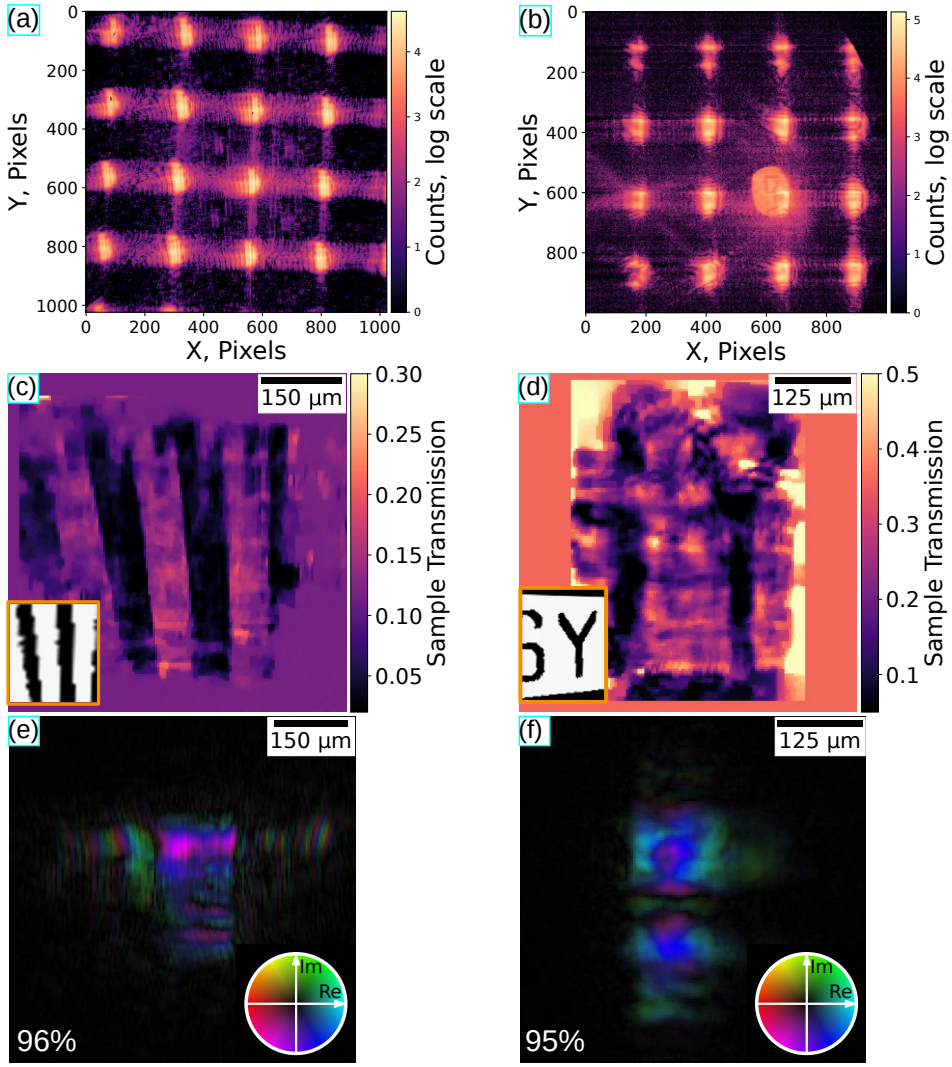

Figure 6.6: **Raw data and results of single-shot ptychography. (a)**, **(b)** Raw data used for the reconstructions of the **(c)** - *stripes* and **(d)** - *letters* regions of the sample. **(a)** and **(b)** were measured using the Andor and the PERCIVAL detector respectively. **(c)**, **(d)** Reconstructed sample transmission for **(c)** - *stripes* and **(d)** - *letters* regions of the Siemsens star sample. **(e)**,**(f)** Reconstructed complex wavefields of the main probe mode corresponding to **(b)** and **(c)** respectively. Percentages in the lower left corner represent the occupancy of the main mode. Adapted from [30].

main reconstructed modes are shown in Figure 6.6 (e,f) lower left corners and are high enough to consider the probe highly coherent. As expected, the reconstruction results were not drastically changed using the partially coherent approximation.

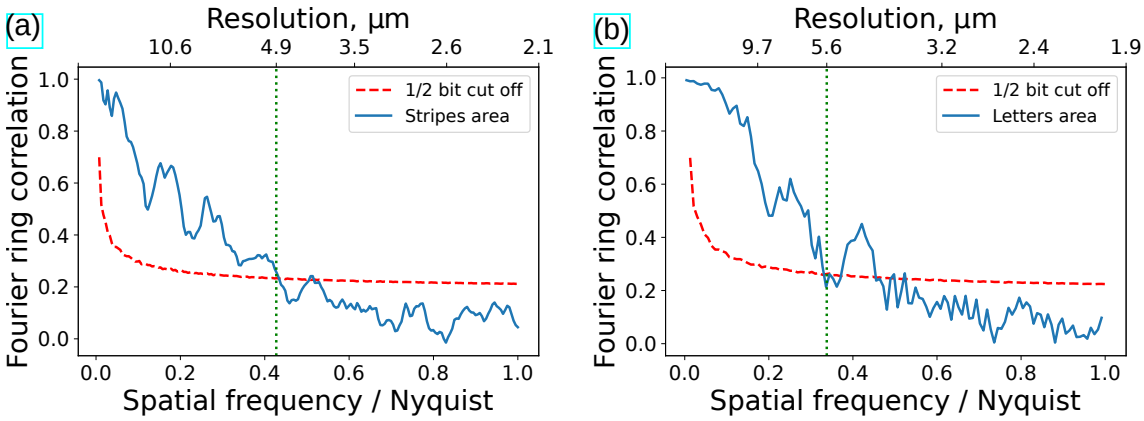

Figure 6.7: **Resolution of the single-shot ptychography reconstruction. (a)**, **(b)** Estimated resolutions of reconstructions using Fourier ring correlation, corresponding to Figure 6.6 **(c)** and **(d)** respectively. Adapted from [30].

The resolution of the reconstructions was evaluated using FRC as shown in Figure 6.7. The resolution at the 0.5 bit cut-off (red dashed lines in Figure 6.7 (a, b)) was found to be approximately 4.9 μm and 5.6 μm for the *stripes* and the *letters* regions respectively.

6.3.4 Conclusion

A novel experimental scheme and differentiable forward model for single-shot ptychography imaging at SASE FELs were realized. They utilize an experimental setup based on a combination of the beam-splitting diffraction grating and the focusing KB optics. The setup was tested during a proof-of-concept experiment at FLASH2, in which images of the two sample regions were reconstructed. The reconstruction was performed by an AD-based ptychography engine utilizing a differentiable forward model specifically developed for the single-shot ptychography.

Several challenges arose during the experiments. First, the use of an absorption diffraction grating resulted in a high thermal load on the grating and its damage during the experiment due to the absorbance of the radiation in the non-transparent areas. This also limited the diffraction efficiency and the total photon flux on the sample. In the future, the use of the phase grating resulting in lower absorbance and

thermal load may solve this problem. However, due to manufacturing restrictions, it may not be easily achieved for XUV wavelengths.

Secondly, the mesh scan pattern produced by the 2D diffraction grating is highly symmetric and may cause raster-grid pathology [139]. Design of a more elaborate fan-out [140] grating tailored for the particular setup and wavelength combination will result in a more asymmetric scan pattern [141]. Additionally, it will improve the separation and minimize the cross-talk between the neighboring beamlets in the detector plane. The use of a beam splitting device with higher diffraction angles may also simplify the raw data separation and improve the achievable resolution by increasing the maximum measured scattering angle of the sample. This, additionally, would require a modification of the forward model to adapt it for a non-paraxial beam propagation [142].

Moreover, the reconstruction was initiated using an approximated probe wavefield and idealized grating transfer function. The initial probe guess was obtained from the averaged intensity measured without the sample and back-propagated to the sample plane. The grating transmission function was assumed to affect only the intensity of the particular beamlet under the paraxial approximation. However, the quality of the reconstruction may be improved by performing a preliminary beam and grating characterization. This would allow using the retrieved probe wavefield and the grating transfer function during the reconstruction of the sample. It would potentially improve convergence as well as the achievable resolution. This could be achieved by additional classical ptychography measurements done on the beam-splitting grating itself with the same setup. However, these measurements would require an additional high-precision motorized stage for the grating to scan it with nanometer resolution.

Finally, the interference effects between the neighboring beamlets resulted in a cross-talk when splitting the data for the reconstruction. The inter-beamlet interference cannot be described by the current forward model and thus was treated as noise during the reconstruction. This problem may be solved by the formulation of an alternative forward model capable of simulating the interference between neighboring beams. The overall diffraction pattern can be expressed as the intensity of the coherent sum of separately propagated complex wavefields interfering at the detector plane. This formalism can integrate the inter-beamlet interference fringes

into the forward model and increase the resolution of the reconstruction. However, it requires an optimization of the inter-beam phase difference, which might prevent the convergence of the reconstruction.

Overall, further development of the single-shot ptychography will increase the achievable resolution and robustness of the technique. An improved experimental setup and an introduction of more sophisticated forward models will extend the use of single-shot ptychography at FELs to sliced 3D ptychography imaging of extended samples. An improved setup would allow much higher sample throughput than scanning methods. In combination with a pump-probe concept, the technique would allow fully utilizing the potential of the femtosecond-long FEL pulses for ultrafast time-resolved imaging of dynamical processes.

Chapter 7

Conclusion and Outlook

An AD-based ptychography framework was developed to perform ptychography experiments at soft X-ray SASE FELs.

The first goal of this work was to develop a ptychography framework that is both flexible and computationally efficient. A GPU-supported AD engine, as a back-end of the framework, allows conceptually splitting the ptychography reconstruction into three separate parts: the forward model, the loss function, and the gradient-based optimizer. Each of these parts can be modified independently and without the need for manual re-derivation of the gradients. This enables fast and easy implementation of the forward models describing various experimental configurations in challenging experimental conditions, such as the simultaneously spatially fluctuating and partially-spatially-coherent beams of SASE FELs. Additionally, the use of the state-of-the-art PyTorch library with the built-in support of the multi-GPU calculation enables fast reconstruction from a large number of measurements. Additional forward models can be implemented in the future to expand the described framework to Bragg-ptychography, ptycho-tomography, and reflective ptychography. Additionally, the introduction of efficient second-order gradient-based optimizers will dramatically improve the convergence speed.

The second goal of this work was the development of a forward model capable of describing simultaneously partially spatially-coherent and spatially-fluctuating light pulses to facilitate a routine, high throughput ptychography at SASE FELs. This problem was solved by the description of the individual pulse wavefields in the basis of incoherent probe modes with shot-to-shot unique modal weights. Additionally, an

AD-based position correction routine was developed to accommodate spatial fluctuations of the probe and possible scan position imprecisions. This model was tested in a proof-of-concept experiment at FLASH2. In this experiment, the transfer function of the sample was successfully reconstructed, despite the spatial fluctuations and partial coherence of the probe. Additionally, the reconstructed main probe modes and shot-to-shot unique modal weights allowed performing a beam characterization without aperturing the beam or selecting its most coherent part.

The influence of better instrumentation on the achievable resolution and throughput of the ptychographic measurements was investigated by performing the measurements with a PERCIVAL detector. This experiment demonstrated 30% improvement of the achievable resolution and 10 times improvement in the measurement time compared to a standard CCD detector, Andor iKon-M. Such improvements will help to exploit the high-flux SASE FEL pulses more efficiently, potentially enabling high-resolution imaging of the large extended samples in just a few minutes.

Finally, the soft X-ray ptychography was applied for imaging of plasma-treated plastic films as the first practical ptychography experiment on an industrially highly relevant sample. In this experiment, the forward model was modified to directly reconstruct the sample thickness by utilizing the refractive representation of the sample. This modification required no manual gradient derivation due to the flexibility provided by the AD. The reconstruction provided a high-resolution image of the 2D distribution of the sample thickness. These results allowed estimating the morphology of the sample surface and the effects of the different plasma treatment techniques on the plastic surface structure.

In the future, several improvements may be made to the experimental setup. First, the use of a more precise and faster motorized stage for scanning the sample would result in a shorter scanning time and higher achievable resolution. Additionally, it would enable ptychographic measurements with a μm-sized beam produced by the KB mirrors at FL24, increasing the photon flux on a sample and improving the resolution. Finally, the use of a diffuser placed upstream of the sample will add complexity to the probe wavefield, improving the achievable resolution even further [143].

The final goal of this work was to bring single-shot ptychography to X-ray FELs. This technique has the potential to enable high-resolution time-resolved imaging of

extended samples. An X-ray-compatible single-shot setup was designed, and the corresponding forward model for the AD-based ptychography framework was developed. The proposed setup and forward model were successfully tested during the proof-of-concept experiment at FLASH2. The experiment resulted in the successful reconstruction of a test sample and the probe wavefield using data measured from a single pulse of the FEL.

Several improvements may be made to this experiment. First, a more elaborate beam-splitter should be designed with better diffraction efficiency, beam separation, and higher thermal stability in mind. This would improve the numerical resolution by minimizing the inter-beamlet cross-talk and better utilization of the detector sensor area. Additionally, a more elaborate measurement sequence, including a preliminary characterization of the grating transfer function and the complex wavefield of the probe, could be realized. This would allow taking the imperfections of the grating into account to initialize the reconstruction with a better estimation of the probe, potentially improving the convergence and the achievable resolution of the reconstruction. Finally, the design of a more complex forward model might allow describing the inter-beamlet interference, which would dramatically improve the numerical resolution of the technique. This can be done by expressing the measured intensity as the intensity of the coherent sum of the complex wavefields of the individual beamlets. Additionally, the accommodation of non-paraxial effects in the beamlet propagation would allow the use of a larger amount of beamlets and thus an increased field of view.

Further development of the technique will increase its resolution and robustness. More sophisticated forward models, combined with an improved setup, will enable single-shot ptychography at soft X-ray FELs for single-shot reflective ptychography, sliced 3D ptychography imaging of extended samples, and time-resolved pump-probe ptychography.

Overall, this work demonstrates how advanced computational approaches and detector development enable routine use of classical ptychography at FELs for high-throughput high-resolution imaging and shot-to-shot beam characterization. Additionally, it presents the X-ray- and FEL- compatible realization of single-shot ptychography is potentially capable of high-resolution and time-resolved imaging of the dynamical samples.

Acknowledgments

First, I want to thank Elke Plönjes-Palm, Christian Schroer, and Martin Beye for giving me the opportunity to be a part of a FLASH team and for supervising my PhD studies.

I want to express my gratitude to Masoud Mehrjoo, Mabel Ruiz-Lopez, and Elke Plönjes-Palm. You were always available for consultations and helped me with all my struggles during the planning of experiments and data analysis.

I want to thank Barbara Keitel and Svea Kreis for helping me with the practical realization of all experimental setups. Also, I want to express my gratitude to Stefan Düsterer, Sören Grunewald, Markus Degenhardt, Erland Müller, Marc Temme, Rui Pan, Seung-gi Gang, Hannah Gerth, and all people from FLASH who helped me with and participated in beamtimes. Additionally, I want to thank Cornelia Wunderer, Jonathan Correa, Alessandro Marras, and the whole PERCIVAL group for providing the PERCIVAL detector and helping with measurements, and all the DESY MAXWELL cluster team for providing the computational resources that I have used extensively during this work.

Special thanks go to all my Hamburg friends: Robin, Nina, Dan, Pragya, Ivette, Stephan, and Katja for all the fun activities we had together.

Last but definitely not least, I want to thank my family and especially my dear girlfriend Masha, who tolerated my passion for work and researches and always supported me with her love in all aspects of my life.

Bibliography

[1] P. Sidorenko and O. Cohen, "Single-shot ptychography," *Optica*, vol. 3, no. 1, pp. 9–14, 2016.

[2] E. Hecht, *Optics*. Pearson Education India, 2012.

[3] I. Snigireva and A. Snigireva, "Microscopy techniques — x-ray microscopy," in *Encyclopedia of Analytical Science (Second Edition)*, P. Worsfold, A. Townshend, and C. Poole, Eds., Second Edition, Oxford: Elsevier, 2005, pp. 151–158, ISBN: 978-0-12-369397-6.

[4] H. N. Chapman and K. A. Nugent, "Coherent lensless x-ray imaging," *Nature photonics*, vol. 4, no. 12, pp. 833–839, 2010.

[5] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: A contemporary overview," *IEEE signal processing magazine*, vol. 32, no. 3, pp. 87–109, 2015.

[6] K. Balewski *et al.*, "Commissioning of petra iii," in *Proceedings of 1st International Particle Accelerator Conference: IPAC*, vol. 10, 2010.

[7] T. Popmintchev, M.-C. Chen, D. Popmintchev, *et al.*, "Bright coherent ultrahigh harmonics in the kev x-ray regime from mid-infrared femtosecond lasers," *science*, vol. 336, no. 6086, pp. 1287–1291, 2012.

[8] P. Emma, R. Akre, J. Arthur, *et al.*, "First lasing and operation of an ångstrom-wavelength free-electron laser," *nature photonics*, vol. 4, no. 9, pp. 641–647, 2010.

[9] J. Miao, T. Ishikawa, I. K. Robinson, and M. M. Murnane, "Beyond crystallography: Diffractive imaging using coherent x-ray light sources," *Science*, vol. 348, no. 6234, pp. 530–535, 2015.

[10] H. Öztürk, X. Huang, H. Yan, I. Robinson, I. C. Noyan, and Y. Chu, "Performance evaluation of bragg coherent diffraction imaging," *New Journal of Physics*, vol. 19, no. 10, p. 103 001, 2017.

[11] S. Marathe, S. Kim, S. Kim, *et al.*, "Coherent diffraction surface imaging in reflection geometry," *Optics express*, vol. 18, no. 7, pp. 7253–7262, 2010.

[12] S. Eisebitt, J. Lüning, W. Schlotter, *et al.*, "Lensless imaging of magnetic nanostructures by x-ray spectro-holography," *Nature*, vol. 432, no. 7019, pp. 885–888, 2004.

[13] M. Guizar-Sicairos and P. Thibault, "Ptychography: A solution to the phase problem," *Physics Today*, vol. 74, no. 9, pp. 42–48, 2021.

[14] H. N. Chapman, A. Barty, M. J. Bogan, *et al.*, "Femtosecond diffractive imaging with a soft-x-ray free-electron laser," *Nature Physics*, vol. 2, no. 12, pp. 839–843, 2006.

[15] H. N. Chapman, C. Caleman, and N. Timneanu, "Diffraction before destruction," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 369, no. 1647, p. 20 130 313, 2014.

[16] G. Van Der Schot, M. Svenda, F. R. Maia, *et al.*, "Imaging single cells in a beam of live cyanobacteria with an x-ray laser," *Nature communications*, vol. 6, no. 1, pp. 1–9, 2015.

[17] J. Geilhufe, B. Pfau, C. M. Günther, M. Schneider, and S. Eisebitt, "Achieving diffraction-limited resolution in soft-x-ray fourier-transform holography," *Ultramicroscopy*, vol. 214, p. 113 005, 2020.

[18] W. Schlotter, R. Rick, K. Chen, *et al.*, "Multiple reference fourier transform holography with soft x rays," *Applied Physics Letters*, vol. 89, no. 16, p. 163 112, 2006.

[19] T. Gorkhover, A. Ulmer, K. Ferguson, *et al.*, "Femtosecond x-ray fourier holography imaging of free-flying nanoparticles," *Nature Photonics*, vol. 12, no. 3, pp. 150–153, 2018.

[20] A. Maiden, M. Humphry, M. Sarahan, B. Kraus, and J. Rodenburg, "An annealing algorithm to correct positioning errors in ptychography," *Ultramicroscopy*, vol. 120, pp. 64–72, 2012.

[21] M. Odstrcil, P. Baksh, S. Boden, *et al.*, "Ptychographic coherent diffractive imaging with orthogonal probe relaxation," *Optics express*, vol. 24, no. 8, pp. 8360–8369, 2016.

[22] A. Schropp and C. G. Schroer, "Dose requirements for resolving a given feature in an object by coherent x-ray diffraction imaging," *New Journal of Physics*, vol. 12, no. 3, p. 035 016, 2010.

[23] E. J. Jaeschke, S. Khan, J. R. Schneider, and J. B. Hastings, *Synchrotron light sources and free-electron lasers: accelerator physics, instrumentation and science applications*. Springer, 2016.

[24] T. Wodzinski, M. Mehrjoo, M. Ruiz-Lopez, *et al.*, "Single-shot transverse coherence measurements with young's double pinholes at flash2," *Journal of Physics Communications*, vol. 4, no. 7, p. 075 014, 2020.

[25] A. Schropp, R. Hoppe, V. Meier, *et al.*, "Full spatial characterization of a nanofocused x-ray free-electron laser beam by ptychographic imaging," *Scientific reports*, vol. 3, no. 1, pp. 1–5, 2013.

[26] S. Sala, B. J. Daurer, M. Odstrcil, *et al.*, "Pulse-to-pulse wavefront sensing at free-electron lasers using ptychography," *Journal of applied crystallography*, vol. 53, no. 4, 2020.

[27] L. Giannessi and C. Masciovecchio, "Fermi: Present and future challenges," *Applied Sciences*, vol. 7, no. 6, p. 640, 2017.

[28] K. Kharitonov, M. Mehrjoo, M. Ruiz-Lopez, *et al.*, "Flexible ptychography platform to expand the potential of imaging at free electron lasers," *Optics express*, vol. 29, no. 14, pp. 22 345–22 365, 2021.

[29] M. Ravandeh, M. Mehrjoo, K. Kharitonov, *et al.*, "X-ray ptychographic imaging and spectroscopic studies of plasma-treated plastic films," *Polymers*, vol. 14, no. 13, p. 2528, 2022.

[30] K. Kharitonov, M. Mehrjoo, M. Ruiz-Lopez, *et al.*, "Single-shot ptychography at a soft x-ray free-electron laser, preprint (version 1)," *available at Research Square [https://doi.org/10.21203/rs.3.rs-1629463/v1]*, 2022.

[31] I. Robinson, G. Gruebel, and S. Mochrie, "Focus on x-ray beams with high coherence," *New Journal of Physics*, vol. 12, no. 3, p. 035 002, 2010.

[32] S. Schreiber, B. Faatz, J. Feldhaus, K. Honkavaara, R. Treusch, M. Vogt, *et al.*, "Status of the flash facility," *MOPD01, these proceedings*, 2012.

[33] E. Plönjes, B. Faatz, M. Kuhlmann, and R. Treusch, "Flash2: Operation, beamlines, and photon diagnostics," in *AIP Conference Proceedings*, AIP Publishing LLC, vol. 1741, 2016, p. 020 008.

[34] E. Saldin, E. Schneidmiller, and M. V. Yurkov, *The physics of free electron lasers*. Springer Science & Business Media, 1999.

[35] D. Attwood and A. Sakdinawat, "X-ray and euv free electron lasers," in *X-Rays and Extreme Ultraviolet Radiation: Principles and Applications*, 2nd ed. Cambridge University Press, 2017, pp. 227–278.

[36] R. Röhlsberger, J. Evers, and S. Shwartz, *Synchrotron light sources and free-electron lasers*, 2020.

[37] K. Tiedtke, A. Azima, N. Von Bargen, *et al.*, "The soft x-ray free-electron laser flash at desy: Beamlines, diagnostics and end-stations," *New journal of physics*, vol. 11, no. 2, p. 023 029, 2009.

[38] B. Faatz, M. Braune, O. Hensler, *et al.*, "The flash facility: Advanced options for flash2 and future perspectives," *Applied Sciences*, vol. 7, no. 11, p. 1114, 2017.

[39] M. Beye, S. Klumpp, and W. Wurth, "Flash2020+ upgrade of flash conceptual design report," *DESY Hamburg, March*, 2020.

[40] E. Allaria, N. Baboi, K. Baev, *et al.*, "Flash2020+ plans for a new coherent source at desy," in *12th Int. Particle Accelerator Conf.(IPAC'21), Campinas, Brazil*, 2021.

[41] S. Reiche *et al.*, "Overview of seeding methods for fels," *Proc. IPAC*, pp. 2063–2067, 2013.

[42] E. A. Schneidmiller, B. Faatz, M. Kuhlmann, *et al.*, "First operation of a harmonic lasing self-seeded free electron laser," *Physical Review Accelerators and Beams*, vol. 20, no. 2, p. 020 705, 2017.

[43] O. Y. Gorobtsov, G. Mercurio, F. Capotondi, *et al.*, "Seeded x-ray free-electron laser generating radiation with laser statistical properties," *Nature communications*, vol. 9, no. 1, pp. 1–6, 2018.

[44] J. W. Goodman, "Introduction to fourier optics, roberts & co," *Publishers, Englewood, Colorado*, 2005.

[45] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[46] D. G. Voelz and M. C. Roggemann, "Digital simulation of scalar optical diffraction: Revisiting chirp function sampling criteria and consequences," *Applied optics*, vol. 48, no. 32, pp. 6132–6142, 2009.

[47] D. Paganin *et al.*, *Coherent X-ray optics*, 6. Oxford University Press on Demand, 2006.

[48] E. Wolf, "New theory of partial coherence in the space–frequency domain. part i: Spectra and cross spectra of steady-state sources," *JOSA*, vol. 72, no. 3, pp. 343–351, 1982.

[49] F. Wittwer, "Development and study of refractive phase retrieval and x-ray multibeam ptychography," Ph.D. dissertation, Staats-und Universitätsbibliothek Hamburg Carl von Ossietzky, 2020.

[50] L. Raimondi, C. Svetina, N. Mahne, *et al.*, "Status of the kb bendable optics at fermi@ elettra fel," in *Adaptive X-Ray Optics III*, International Society for Optics and Photonics, vol. 9208, 2014, p. 920 804.

[51] M. Ruiz-Lopez, "Status of flash2 (conference presentation)," in *X-Ray Free-Electron Lasers: Advances in Source Development and Instrumentation V*, International Society for Optics and Photonics, vol. 11038, 2019, 110380F.

[52] C. Jacobsen, J. Deng, and Y. Nashed, "Strategies for high-throughput focused-beam ptychography," *Journal of Synchrotron Radiation*, vol. 24, no. 5, pp. 1078–1081, 2017.

[53] I. A. Vartanyants and A. Singer, "Coherence properties of hard x-ray synchrotron sources and x-ray free-electron lasers," *New Journal of Physics*, vol. 12, no. 3, p. 035 004, 2010.

[54] A. Singer, I. Vartanyants, M. Kuhlmann, S. Duesterer, R. Treusch, and J. Feldhaus, "Transverse-coherence properties of the free-electron-laser flash at desy," *Physical Review Letters*, vol. 101, no. 25, p. 254 801, 2008.

[55] A. Singer, F. Sorgenfrei, A. P. Mancuso, *et al.*, "Spatial and temporal coherence properties of single free-electron laser pulses," *Optics express*, vol. 20, no. 16, pp. 17 480–17 495, 2012.

[56] D. Mai, J. Hallmann, T. Reusch, *et al.*, "Single pulse coherence measurements in the water window at the free-electron laser flash," *Optics express*, vol. 21, no. 11, pp. 13 005–13 017, 2013.

[57] T. Mey, B. Schäfer, K. Mann, B. Keitel, M. Kuhlmann, and E. Plönjes, "Wigner distribution measurements of the spatial coherence properties of the free-electron laser flash," *Optics express*, vol. 22, no. 13, pp. 16 571–16 584, 2014.

[58] A. Singer, U. Lorenz, F. Sorgenfrei, *et al.*, "Hanbury brown–twiss interferometry at a free-electron laser," *Physical review letters*, vol. 111, no. 3, p. 034 802, 2013.

[59] O. Y. Gorobtsov, G. Mercurio, G. Brenner, *et al.*, "Statistical properties of a free-electron laser revealed by hanbury brown–twiss interferometry," *Physical Review A*, vol. 95, no. 2, p. 023 843, 2017.

[60] E. A. Schneidmiller and M. Yurkov, "Coherence properties of the radiation from flash," *Journal of Modern Optics*, vol. 63, no. 4, pp. 293–308, 2016.

[61] M. Rose, I. Vartaniants, W. Wurth, and E. Weckert, "Coherent x-ray diffractive imaging of biological samples in 2d and 3d with synchrotron and xfel radiation," Ph.D. dissertation, Deutsches Elektronen-Synchrotron, 2018.

[62] E. H. Tsai, I. Usov, A. Diaz, A. Menzel, and M. Guizar-Sicairos, "X-ray ptychography with extended depth of field," *Optics express*, vol. 24, no. 25, pp. 29 089–29 108, 2016.

[63] M. Kahnt, L. Grote, D. Brückner, *et al.*, "Multi-slice ptychography enables high-resolution measurements in extended chemical reactors," *Scientific reports*, vol. 11, no. 1, pp. 1–11, 2021.

[64] S. Marchesini, H. He, H. N. Chapman, *et al.*, "X-ray image reconstruction from a diffraction pattern alone," *Physical Review B*, vol. 68, no. 14, p. 140 101, 2003.

[65] B. J. Daurer, S. Sala, M. F. Hantke, *et al.*, "Ptychographic wavefront characterization for single-particle imaging at x-ray lasers," *Optica*, vol. 8, no. 4, pp. 551–562, 2021.

[66] P. W. Hawkes and J. C. Spence, *Springer handbook of microscopy*. Springer, 2019.

[67] A. Maiden, D. Johnson, and P. Li, "Further improvements to the ptychographical iterative engine," *Optica*, vol. 4, no. 7, pp. 736–745, 2017.

[68] J. C. da Silva, C. Guilloud, O. Hignette, *et al.*, "Overcoming the challenges of high-energy x-ray ptychography," *Journal of synchrotron radiation*, vol. 26, no. 5, pp. 1751–1762, 2019.

[69] J. Rodenburg and A. Maiden, "Ptychography," *Springer Handbook of Microscopy*, pp. 819–904, 2019.

[70] P. Li and A. M. Maiden, "Ten implementations of ptychography," *Journal of microscopy*, vol. 269, no. 3, pp. 187–194, 2018.

[71] M. Rose, D. Dzhigaev, T. Senkbeil, *et al.*, "High-dynamic-range water window ptychography," in *Journal of Physics: Conference Series*, IOP Publishing, vol. 849, 2017, p. 012 027.

[72] K. Nugent, A. Peele, H. Quiney, and H. Chapman, "Diffraction with wavefront curvature: A path to unique phase recovery," *Acta Crystallographica Section A: Foundations of Crystallography*, vol. 61, no. 3, pp. 373–381, 2005.

[73] M. Stockmar, P. Cloetens, I. Zanette, *et al.*, "Near-field ptychography: Phase retrieval for inline holography using a structured illumination," *Scientific reports*, vol. 3, no. 1, pp. 1–6, 2013.

[74] R. J. Bean, "Domain structure imaging by bragg geometry x-ray ptychography," Ph.D. dissertation, University College London (University of London), 2013.

[75] B. Zhang, D. F. Gardner, M. D. Seaberg, *et al.*, "High contrast 3d imaging of surfaces near the wavelength limit using tabletop euv ptychography," *Ultramicroscopy*, vol. 158, pp. 98–104, 2015.

[76] P. C. Konda, L. Loetgering, K. C. Zhou, S. Xu, A. R. Harvey, and R. Horstmeyer, "Fourier ptychography: Current applications and future promises," *Optics express*, vol. 28, no. 7, pp. 9603–9630, 2020.

[77] K. Wakonig, A. Diaz, A. Bonnin, *et al.*, "X-ray fourier ptychography," *Science advances*, vol. 5, no. 2, eaav0282, 2019.

[78] P. Sidorenko, O. Lahav, and O. Cohen, "Ptychographic ultrahigh-speed imaging," *Optics Express*, vol. 25, no. 10, pp. 10 997–11 008, 2017.

[79] O. Bunk, M. Dierolf, S. Kynde, I. Johnson, O. Marti, and F. Pfeiffer, "Influence of the overlap parameter on the convergence of the ptychographical iterative engine," *Ultramicroscopy*, vol. 108, no. 5, pp. 481–487, 2008.

[80] J. C. da Silva and A. Menzel, "Elementary signals in ptychography," *Optics express*, vol. 23, no. 26, pp. 33 812–33 821, 2015.

[81] M. Guizar-Sicairos and J. R. Fienup, "Phase retrieval with transverse translation diversity: A nonlinear optimization approach," *Optics express*, vol. 16, no. 10, pp. 7264–7278, 2008.

[82] F. Zhang, I. Peterson, J. Vila-Comamala, *et al.*, "Translation position determination in ptychographic coherent diffraction imaging," *Optics express*, vol. 21, no. 11, pp. 13 592–13 606, 2013.

[83] P. Thibault and A. Menzel, "Reconstructing state mixtures from diffraction measurements," *Nature*, vol. 494, no. 7435, pp. 68–71, 2013.

[84] P. Li, D. J. Batey, T. B. Edo, A. D. Parsons, C. Rau, and J. M. Rodenburg, "Multiple mode x-ray ptychography using a lens and a fixed diffuser optic," *Journal of Optics*, vol. 18, no. 5, p. 054 008, 2016.

[85] D. J. Batey, D. Claus, and J. M. Rodenburg, "Information multiplexing in ptychography," *Ultramicroscopy*, vol. 138, pp. 13–21, 2014.

[86] P. Thibault and M. Guizar-Sicairos, "Maximum-likelihood refinement for coherent diffractive imaging," *New Journal of Physics*, vol. 14, no. 6, p. 063 004, 2012.

[87] P. Godard, M. Allain, V. Chamard, and J. Rodenburg, "Noise models for low counting rate coherent diffraction imaging," *Optics express*, vol. 20, no. 23, pp. 25 914–25 934, 2012.

[88] L. Kaup and B. Kaup, *Holomorphic functions of several variables: an introduction to the fundamental theory*. Walter de Gruyter, 2011, vol. 3.

[89] L.-H. Yeh, J. Dong, J. Zhong, *et al.*, "Experimental robustness of fourier ptychography phase retrieval algorithms," *Optics express*, vol. 23, no. 26, pp. 33 214–33 240, 2015.

[90] C. C. Margossian, "A review of automatic differentiation and its efficient implementation," *Wiley interdisciplinary reviews: data mining and knowledge discovery*, vol. 9, no. 4, e1305, 2019.

[91] H. M. Bücker, G. Corliss, P. Hovland, U. Naumann, and B. Norris, *Automatic differentiation: applications, theory, and implementations*. Springer Science & Business Media, 2006, vol. 50.

[92] A. S. Jurling and J. R. Fienup, "Applications of algorithmic differentiation to phase retrieval algorithms," *JOSA A*, vol. 31, no. 7, pp. 1348–1359, 2014.

[93] S. Kandel, S. Maddali, M. Allain, S. O. Hruszkewycz, C. Jacobsen, and Y. S. Nashed, "Using automatic differentiation as a general framework for ptychographic reconstruction," *Optics express*, vol. 27, no. 13, pp. 18 653–18 672, 2019.

[94] A. Griewank and A. Walther, *Evaluating derivatives: principles and techniques of algorithmic differentiation*. SIAM, 2008.

[95] S. Marchesini, H. Krishnan, B. J. Daurer, *et al.*, "Sharp: A distributed gpu-based ptychographic solver," *Journal of applied crystallography*, vol. 49, no. 4, pp. 1245–1252, 2016.

[96] V. Elser, "Phase retrieval by iterated projections," *JOSA A*, vol. 20, no. 1, pp. 40–55, 2003.

[97] K. Jaganathan, Y. C. Eldar, and B. Hassibi, "Phase retrieval: An overview of recent developments," *Optical Compressive Imaging*, pp. 279–312, 2016.

[98] Z. Wen, C. Yang, X. Liu, and S. Marchesini, "Alternating direction methods for classical and ptychographic phase retrieval," *Inverse Problems*, vol. 28, no. 11, p. 115 010, 2012.

[99] A. Paszke, S. Gross, F. Massa, *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *Advances in neural information processing systems*, vol. 32, 2019.

[100] NVIDIA, P. Vingelmann, and F. H. Fitzek, *Cuda, release: 10.2.89*, 2020. [Online]. Available: https://developer.nvidia.com/cuda-toolkit.

[101] G. Van Rossum and F. L. Drake, *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009, ISBN: 1441412697.

[102] J. Reppin, C. Beyer, T. Hartmann, *et al.*, "Interactive analysis notebooks on desy batch resources," *Computing and Software for Big Science*, vol. 5, no. 1, pp. 1–11, 2021.

[103] P. Dwivedi, A. Konijnenberg, S. Pereira, and H. Urbach, "Lateral position correction in ptychography using the gradient of intensity patterns," *Ultramicroscopy*, vol. 192, pp. 29–36, 2018.

[104] M. Jaderberg, K. Simonyan, A. Zisserman, *et al.*, "Spatial transformer networks," *Advances in neural information processing systems*, vol. 28, 2015.

[105] P. Bühlmann and S. Van De Geer, *Statistics for high-dimensional data: methods, theory and applications*. Springer Science & Business Media, 2011.

[106] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical imaging and vision*, vol. 20, no. 1, pp. 89–97, 2004.

[107] P. Li, T. Edo, D. Batey, J. Rodenburg, and A. Maiden, "Breaking ambiguities in mixed state ptychography," *Optics express*, vol. 24, no. 8, pp. 9038–9052, 2016.

[108] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016.

[109] M. Claesen and B. De Moor, "Hyperparameter search in machine learning," *arXiv preprint arXiv:1502.02127*, 2015.

[110] N. Qian, "On the momentum term in gradient descent learning algorithms," *Neural networks*, vol. 12, no. 1, pp. 145–151, 1999.

[111] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[112] L. N. Smith, "Cyclical learning rates for training neural networks," in *2017 IEEE winter conference on applications of computer vision (WACV)*, IEEE, 2017, pp. 464–472.

[113] M. Kahnt, J. Becher, D. Brückner, *et al.*, "Coupled ptychography and tomography algorithm improves reconstruction of experimental data," *Optica*, vol. 6, no. 10, pp. 1282–1289, 2019.

[114] M. Odstrcil, P. Baksh, S. Boden, *et al.*, "Ptychographic coherent diffractive imaging with orthogonal probe relaxation," *Optics express*, vol. 24, no. 8, pp. 8360–8369, 2016.

[115] S. Sala, B. J. Daurer, M. Odstrcil, *et al.*, "Pulse-to-pulse wavefront sensing at free-electron lasers using ptychography," *Journal of applied crystallography*, vol. 53, no. 4, pp. 949–956, 2020.

[116] H. Andrews and C. Patterson, "Singular value decompositions and digital image processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 1, pp. 26–53, 1976.

[117] M. Manfredda, C. Fava, S. Gerusina, *et al.*, "The evolution of kaos, a multipurpose active optics system for euv/soft x-rays," *Synchrotron Radiation News*, pp. 1–8, 2022.

[118] B. Keitel, E. Plönjes, S. Kreis, *et al.*, "Hartmann wavefront sensors and their application at flash," *Journal of Synchrotron Radiation*, vol. 23, no. 1, pp. 43–49, 2016.

[119] N. Banterle, K. H. Bui, E. A. Lemke, and M. Beck, "Fourier ring correlation as a resolution criterion for super-resolution microscopy," *Journal of structural biology*, vol. 183, no. 3, pp. 363–367, 2013.

[120] J. Vila-Comamala, A. Diaz, M. Guizar-Sicairos, *et al.*, "Characterization of high-resolution diffractive x-ray optics by ptychographic coherent diffractive imaging," *Optics express*, vol. 19, no. 22, pp. 21 333–21 344, 2011.

[121] S. Goh, Y. Tao, P. van der Slot, *et al.*, "Single-shot fluctuations in waveguided high-harmonic generation," *Optics express*, vol. 23, no. 19, pp. 24 888–24 902, 2015.

[122] A. Marras, C. Wunderer, J. Correa, *et al.*, "Percival: A soft x-ray imager for synchrotron rings and free electron lasers," in *AIP Conference Proceedings*, AIP Publishing LLC, vol. 2054, 2019, p. 060 060.

[123] C. Wunderer, J. Correa, A. Marras, *et al.*, "The percival 2-megapixel monolithic active pixel imager," *Journal of Instrumentation*, vol. 14, no. 01, p. C01006, 2019.

[124] K. Ozaltin, M. Lehocky, P. Humpolicek, *et al.*, "Anticoagulant polyethylene terephthalate surface by plasma-mediated fucoidan immobilization," *Polymers*, vol. 11, no. 5, p. 750, 2019.

[125] K. Gotoh, A. Yasukawa, and Y. Kobayashi, "Wettability characteristics of poly (ethylene terephthalate) films treated by atmospheric pressure plasma and ultraviolet excimer light," *Polymer journal*, vol. 43, no. 6, pp. 545–551, 2011.

[126] *Private communications with plasmawerk hamburg gmbh.*

[127] I. Levchenko, S. Xu, O. Baranov, O. Bazaka, E. P. Ivanova, and K. Bazaka, "Plasma and polymers: Recent progress and trends," *Molecules*, vol. 26, no. 13, p. 4091, 2021.

[128] M. Tuominen, J. Lahti, J. Lavonen, T. Penttinen, J. P. Räsänen, and J. Kuusipalo, "The influence of flame, corona and atmospheric plasma treatments on surface properties and digital print quality of extrusion coated paper," *Journal of adhesion science and technology*, vol. 24, no. 3, pp. 471–492, 2010.

[129] J. F. Senge, A. H. Astaraee, P. Dłotko, S. Bagherifard, and W. A. Bosbach, "Extending conventional surface roughness iso parameters using topological data analysis for shot peened surfaces," *Scientific Reports*, vol. 12, no. 1, pp. 1–12, 2022.

[130] P. Eaton and P. West, *Atomic force microscopy*. Oxford university press, 2010.

[131] X. He, C. Liu, and J. Zhu, "Single-shot fourier ptychography based on diffractive beam splitting," *Optics letters*, vol. 43, no. 2, pp. 214–217, 2018.

[132] D. Goldberger, J. Barolak, C. G. Durfee, and D. E. Adams, "Three-dimensional single-shot ptychography," *Optics Express*, vol. 28, no. 13, pp. 18 887–18 898, 2020.

[133] J. Barolak, D. Goldberger, J. Squier, Y. Bellouard, C. Durfee, and D. Adams, "Wavelength-multiplexed single-shot ptychography," *Ultramicroscopy*, vol. 233, p. 113 418, 2022.

[134] X. Pan, C. Liu, and J. Zhu, "Single shot ptychographical iterative engine based on multi-beam illumination," *Applied Physics Letters*, vol. 103, no. 17, p. 171 105, 2013.

[135] T. Edo, D. Batey, A. Maiden, *et al.*, "Sampling in x-ray ptychography," *Physical Review A*, vol. 87, no. 5, p. 053 850, 2013.

[136] U. Hahn and K. Tiedtke, "The gas attenuator of flash at desy," in *AIP Conference Proceedings*, American Institute of Physics, vol. 879, 2007, pp. 276–282.

[137] M. Manfredda, C. Fava, S. Gerusina, *et al.*, "The evolution of kaos, a multipurpose active optics system for euv/soft x-rays," *Synchrotron Radiation News*, vol. 0, no. 0, pp. 1–8, 2022.

[138] B. Enders, M. Dierolf, P. Cloetens, M. Stockmar, F. Pfeiffer, and P. Thibault, "Ptychography with broad-bandwidth radiation," *Applied Physics Letters*, vol. 104, no. 17, p. 171 104, 2014.

[139] A. Fannjiang, "Raster grid pathology and the cure," *Multiscale Modeling & Simulation*, vol. 17, no. 3, pp. 973–995, 2019.

[140] D. C. Kim, A. Hermerschmidt, P. Dyachenko, and T. Scharf, "Adjoint method and inverse design for diffractive beam splitters," in *Components and Packaging for Laser Systems VI*, International Society for Optics and Photonics, vol. 11261, 2020, p. 1 126 115.

[141] X. Huang, H. Yan, R. Harder, Y. Hwu, I. K. Robinson, and Y. S. Chu, "Optimization of overlap uniformness for ptychography," *Optics Express*, vol. 22, no. 10, pp. 12 634–12 644, 2014.

[142] C. Chang, X. Pan, H. Tao, C. Liu, S. P. Veetil, and J. Zhu, "3d single-shot ptychography with highly tilted illuminations," *Optics Express*, vol. 29, no. 19, pp. 30 878–30 891, 2021.

[143] P. Song, S. Jiang, H. Zhang, *et al.*, "Super-resolution microscopy via ptychographic structured modulation of a diffuser," *Optics letters*, vol. 44, no. 15, pp. 3645–3648, 2019.