UNIVERSITY OF SOUTHAMPTON
RWTH AACHEN UNIVERSITY

Coherent Diffractive Imaging Using Table-top Sources

Michal Odstrčil

supervised by

Dr. William S. Brocklesby, Prof. Jeremy Frey
and Prof. Larissa Juschkin

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
Faculty of Physical and Applied Sciences
Optoelectronics Research Center

May 2017

UNIVERSITY OF SOUTHAMPTON
RWTH AACHEN UNIVERSITY

Abstract

Faculty of Physical and Applied Sciences
Optoelectronics Research Center

Doctor of Philosophy

COHERENT DIFFRACTIVE IMAGING USING TABLE-TOP SOURCES

by Michal Odstrčil

Lensless microscopy, which is also called coherent diffractive imaging (CDI), is a novel and revolutionary approach to imaging. Compared to lens-based microscopy i.e. optical, fluorescence or electron microscopy, lensless microscopy does not need to rely on lenses to obtain the image of the sample. Instead of this, lensless microscopy relies on coherence of the illumination and on computational postprocessing of the measured data. This allows CDI methods huge flexibility in imaging setups compared to the lens-based microscopy. The family of CDI methods has been growing rapidly over the last years. The first simple application in the X-ray regime for imaging of aperiodic (non-crystalline) samples was done by Janwei Miao et al. [1]. In this experiment no optics were used around the sample and a single diffraction pattern was used to reconstruct the image. The current state of the art CDI applications collect thousands of images that are subsequently processed together by an appropriate method to create up to gigapixel resolution images [2–4].

CDI generally has several advantages compared to the ordinary lens-based imaging: firstly the image quality is not limited by the quality of lenses. Instead of an objective lens, a numerical iterative algorithm is used. The maximum resolution of CDI is, similarly to lens-based systems, limited by the highest spatial frequency collected by the imaging system. The advantage of CDI is the possibility of either avoiding the imaging lenses or including their limited quality into the reconstruction process [5]. However, the main advantage of the CDI methods is to recover both phase and amplitude of the exit-wave field behind the sample. This allows to obtain much higher contrast for phase objects that would have otherwise low contrast in bright field microscopy. The ability of the phase and amplitude recovery without limitation of lenses makes CDI a powerful method with a broad range of applications in nanoscience, material science and biology.

Contents

List of Figures

Declaration of Authorship

I, Michal Odstrcil declare that this thesis entitled "Coherent Diffractive Imaging Using Table-top Sources" and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated
- Where I have consulted the published work of others, this is always clearly attributed
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help; Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

Signed: _____

Date: _____

Acknowledgements

I would like to thank to Dr. Bill Brocklesby, Prof. Jeremy Frey and Prof. Larissa Juschkin for their support of my work, useful discussions and their enthusiasm. I am very grateful to Prof. Larissa Juschkin for her detailed and critical review of this thesis. Further, I would like to thank to Pete Baksh (University of Southampton) and Jan Bußmann (Forschungszentrum Juelich) for their help with building and running the EUV laboratory setups, Jo Bailey for preparation of the neurons samples.

Thanks to all my friends and family for their patience and support, it would not be possible without them.

Finally, I acknowledge Erasmus Mundus Joint Doctorate Programme EXTATIC for funding my PhD.

List of abbreviations

ASM	Angular spectrum method
AuC	Area under curve
CARS	Coherent anti-Stokes Raman spectroscopy
CDI	Coherent diffractive imaging
CG	Conjugate gradients method
DM	Difference maps algorithm
DoF	Depth of focus
DoFr	Degrees of freedom
ePIE	Extended ptychography iterative engine algorithm
ER	Error-reduction algorithm
ERFC	Error function
EUV	Extreme ultraviolet
FEL	Free electron laser
FFT	Fast Fourier transformation
FWHM	Full width at half maximum
GDP	Gas discharge plasma
HDR	High dynamic range
HHG	High harmonic generation
HIO	Hybrid input-output algorithm
HSV	Hue saturation value colour transformation
HXR	Hard X-ray
LM	Levenberg-Marquardt algorithm (dumped LSQ)
LSQ	Least squares algorithm
OPR	Orthogonal probe relaxation
PPLN	Periodically poled lithium niobate
pPIE	Parallel ptychography iterative engine algorithm
PSF	Point spread function
NA	Numerical aperture
NMBA	4-nitro-4-methylbenzilidene aniline
RAAR	Relaxed averaged alternating reflections algorithm
SHG	Second harmonic generation
SNR	Signal to noise ratio
SVD	Singular value decomposition
SXR	Soft X-ray
THG	Third harmonic generation

List of publications

Publications during PhD

2016 **M.Odstrcil**, P.Baksh, C.Gawith, R.Vrcelj, J.G. Frey, W.S. Brocklesby, *Nonlinear ptychographic coherent diffractive imaging*, Optics Express 24.18 (2016): 20245-20252.

M.Odstrcil, P.Baksh, S.A.Boden, R. Card, J.E. Chad, W.S. Brocklesby, J.G.Frey,
Ptychographic coherent diffractive imaging with orthogonal probe relaxation, Optics express 24.8 (2016): 8360-8369.

P.Baksh, **M.Odstrcil**, H.Kim, S.A.Boden, W.S. Brocklesby, J.G. Frey, *Wide-field Broadband EUV Transmission Ptychography using a High Harmonic Source*, Optics Letters (2016): 1-5.

H.S. Kim, P. Baksh, **M.Odstrcil**, M. Miszczak, J.G. Frey, L. Juschkin and W.S. Brocklesby, *Lloyds mirror interference lithography with EUV radiation from a high-harmonic source.* Applied Physics Express (2016), 9(7), p.076701

2015 **M.Odstrcil**, J.Bussman, D.Rudolf, R.Bresenitz, J.Miao, W.Brocklesby, L.Juschkin, *Ptychographic imaging with a compact gas-discharge plasma extreme ultraviolet light source*, Optics Letters 2015 Vol.40(23) pp.5574-5577

M.Odstrcil, P.Baksh, H.Kim, S.A.Boden, W.S.Brocklesby, J.G.Frey, *Ultra-broadband ptychography with self-consistent coherence estimation from a high harmonic source*, SPIE Optical Engineering+ Applications. International Society for Optics and Photonics, 2015.

D.Rudolf, J.Bussmann, **M.Odstrcil**, M.Dong, K.Bergmann, S.Danylyuk, L.Juschkin, *Interferometric broadband Fourier spectroscopy with a partially coherent gas-discharge extreme ultraviolet light source*, Optics Letters 2015 Vol.40(12) pp.2818-2821

Previous publications

2014 **M.Odstrcil**, J. Mlynar, V. Weinzettl, P. Hacek, M. Berta, T. Szabolics, A. Bencze, *Plasma Tomographic Reconstruction from Tangentially Viewing Camera with Background Subtraction*, Review of Scientific Instruments 85.1 (2014): 013509.

2013 **M.Odstrcil**, A. Murari, J. Mlynar, *Comparison of Advanced Machine Learning Tools for Disruption Prediction and Disruption Studies.* IEEE Transactions on Plasma Science. 2013 Jul;41(7):1751-9.

2012 **M.Odstrcil**, J. Mlynar, T. Odstrcil, B. Alper, A. Murari; *Modern numerical methods for plasma tomography optimisation*, Nuclear Inst. and Methods in Physics Research, A (2012), pp. 156-161

T. Odstrcil, **M.Odstrcil**, O. Grover, V. Svoboda, I. Duran, J. Mlynar, *Low cost alternative of high speed visible light camera for tokamak experiments.* Review of Scientific Instruments, 83(10)

J Mlynar, M Imrisek, V Weinzettl, **M.Odstrcil**, J Havlicek, F Janky, B Alper, A Murari, *Introducing minimum Fisher regularisation tomography to AXUV and soft x-ray diagnostic systems of the COMPASS tokamak*, The Review of scientific instrument, (2012)

Summary of thesis

The main goal of this thesis is to demonstrate that the laboratory short wavelength sources can be used for high quality coherent diffractive imaging (CDI). CDI, also known as lensless microscopy, is a technique that uses measured scatter from a coherently illuminated sample in combination with an additional knowledge to recover phase shift and attenuation introduced by a sample. Thanks to availability of high brightness Xray sources such as synchrotrons, CDI is becoming to be a relatively standard method in short wavelength microscopy in the X-ray wavelength range. However, applications of this method for laboratory short wavelength sources have been limited till now only to simple high contrast artificial samples.

In this thesis, we have demonstrated CDI and in particular its scanning based version called Ptychography, using several laboratory sources in wavelength range from infrared (IR) to the extreme ultraviolet (EUV). Ptychography uses the overlap between adjacent illuminated regions as the additional constraint that is needed to reconstruct the phase that is lost when the scattered intensity is measured.

CDI methods can provide the ultimate imaging quality because there is no optics limiting the imaging quality. However, as it is shown in this work, the imaging quality is often determined by unavoidable systematic errors. In this thesis, different limitations are discussed and appropriate corrections of the reconstruction algorithm are proposed. These corrections can improve reconstruction quality, extract additional information or make imaging possible in the case, when the data quality would not be sufficient for the standard methods.

In particular, unique methods that account for variable illumination wavefront, intensity and decoherence-like effects during the scan, sample position errors and sample anisotropy are proposed. We have also proposed multiple enhancements of the most common gradient-based solver method that improve the convergence speed and allow faster and more efficient parallel execution of the calculations on a graphical card. The proposed extensions were tested using simulated and experimental data in order to understand the limitations and issues of laboratory setups.

Along with the better numerical approach, significant improvements of our EUV experimental setups were done. It allowed us to reduce exposure time in experiments with our high harmonic generation (HHG) based EUV source at 29 nm wavelength (42 eV) by two orders of magnitude compared to the previous results published by our group. The stability of the source was significantly increased as well. All these changes made it possible to produce high quality reconstructions with field of view up to $100\times100\,\mu$m and resolution around 50 nm. Further, high resolution reconstructions of hippocampal mouse neurons samples are shown. This is the first demonstration of high resolution imaging of weakly scattering biological samples using laboratory based short wavelength CDI. We have demonstrated CDI with plasma based weakly coherent sources as well. Despite the relatively simple setup, we show the first demonstration of ptychography at 13.5 nm wavelength (92 eV) that is important for lithography industry with a laboratory source.

Finally, a novel method that we call nonlinear ptychographic CDI is introduced and tested on crystalline samples. Nonlinear CDI makes it possible to reconstruct the relative phase of electric field resulting from the nonlinear optical effects and to decouple it from the complex illumination pattern.

Chapter 1

Motivation

This work is focused on application of coherent diffractive imaging (CDI) using laboratory sources. Two optical ranges are particularly of interest for us: near-infrared illumination generated by an ultrafast laser around 790 nm wavelength for nonlinear ptychography imaging and Extreme Ultraviolet (EUV) (10-50 nm).

Microscopy in the EUV range generally makes it possible to achieve higher resolution as well as better contrast than the visible light microscopy. Simultaneously EUV has better penetration depth than the standard transmission electron microscopy and better sensitivity to the lighter elements in the specimens. EUV microscopy allows the study of thin surface layers of samples or features buried under the sample surface such as imperfections of EUV lithography masks. Further, the EUV photon energy in our setup at 42 eV is still relatively low compared to the "water-window" SXR microscopy (280-530 eV) and hence resolution for biological samples is less limited by radiation damage. In contrast to florescence microscopy, EUV microscopy provides label-free high resolution and high contrast imaging. Moreover, CDI methods provide quantitative information about the complex exit-wave at the sample position. This gives additional information that would be lost if the standard lens-based incoherent bright-field microscopy was used.

The practical advantage of using coherent EUV radiation for CDI microscopy is that tabletop sources with sufficient coherence and flux for imaging are available. In this thesis, two main laboratory sources are presented: high harmonic generation (HHG) based sources or plasma gas discharge sources. This is an advantage compared to the hard X-ray range, where reasonable coherent photon flux for

imaging can be currently produced only by large facilities such as synchrotrons or free electron lasers.

The main drawbacks of EUV microscopy are mostly shared with electron microscopy. There is a requirement of vacuum for the entire imaging setup because EUV radiation is strongly attenuated in any gas or other material. This is in particular a disadvantage for biological samples because it is complicated to observe them in their natural wet environment.

Further, laboratory sources present an additional challenge compared to the large facilities because they usually provide significantly lower coherent flux and particularly in the case of high harmonic generation EUV sources, also worse illumination stability. Therefore longer acquisition time and more robust algorithms to deal with the lower data quality are required.

The resolution of incoherent diffraction limited microscopy with a circular collecting pupil is given by the Rayleigh criterion [21]

$$\delta x \geq \frac{1.22\lambda}{\mathrm{NA}_o + \mathrm{NA}_c} \tag{1.1}$$

where λ is wavelength, NA_o, NA_c is the numerical aperture of objective and condenser. Numerical aperture is defined as $\mathrm{NA}= n\sin(\alpha)$, where n denotes refractive index of the environment and α is the maximum angle of diffraction from the optical axis passing through the optical system. The refractive index of environment is fixed to $n = 1$ in the case of EUV microscopy because of a short penetration depth of EUV in any material. The maximum available NA of EUV optics is limited by technical problems in manufacturing multilayer mirrors or Fresnel zone plates to roughly 0.3. This problem can be partly overcome if CDI is used. The imaging quality is not limited by the optics and the maximal achievable resolution is limited only by the highest scattering angle that is collected with sufficient signal to noise ratio (SNR).

The soft X-ray (SXR) and hard X-ray (HXR) energy ranges make it possible to reach even higher resolution than EUV wavelengths, however the maximum achievable resolution is often rather limited by the radiation damage of the sample. In order to avoid this problem, methods based on diffraction before destruction have been tested on free electron lasers (FELs) [22]. These methods use such short

light pulses that the pulse passes through the sample before Columbic explosion and thus resolution far beyond the damage threshold can be achieved [23, 24].

Chapter 2

Coherent Diffractive Imaging (CDI) methods

Coherent diffractive imaging has many advantages compared to the common lens-based imaging methods, however has additional requirements that need to be fulfilled otherwise the imaging quality can be limited. Generally, the optimal method always needs to be chosen for the specific task and CDI methods cannot be automatically considered superior to the lens-based methods. However, CDI has the following advantages compared to the standard lens-based methods (i.e. in focus bright field imaging).

- Resolution is not limited by quality of imaging optics
- Reconstruction of both phase shift and attenuation
- Better use of the incident photons ⇒ lower dose at sample because inefficient X-ray optics behind sample is avoided.
- Longer depth of focus [2, 25–27]
- More flexible setup [28, 29]
- More compact setup [30, 31]

The main disadvantage of CDI methods is requirement of stable coherent illumination with sufficient flux that is not easily achievable for wavelengths below 120 nm. A further disadvantage is complicated numerical processing, however

computational power will be less limiting and cheaper in future decades. In the case of table-top EUV imaging in the transmissive setup, the main advantage is reconstruction of phase shift and attenuation. Lower dose at the sample is relevant only above 100 eV photon energy [24]. If only attenuation needs to be imaged and neither dose on the sample nor the numerical aperture of the collecting optics are limiting, then the table-top lens-based methods with incoherent EUV sources can be more practical and imaging in the "water window" was already achieved [32–35].

2.1 Physics background of diffraction imaging

When light passes through a specimen, the wave front of the light will be distorted because of interaction of the specimen atoms with photons. The change in phase and attenuation will depend on optical properties of the sample such as thickness, electron density and atomic numbers of the elements in the specimen.

For most of our analysis, we will consider only sufficiently thin objects, so that variation of sample composition and propagation of the illumination function within the sample can be neglected. This means that the local refractive index $n(x, y, z)$ can be approximated by a projection into the beam propagation direction

$$\hat{n}_{\perp}(x, y) = \frac{1}{L} \int_0^L n(x, y, z) \mathrm{d}z \quad , \tag{2.1}$$

where L is local sample thickness. The maximal sample thickness depends on the illumination wavelength, resolution and illumination wave-front. In the case of the confocal microscopy, the depth of focus is given as

$$\mathrm{DoF} = \frac{\lambda}{(\mathrm{NA}_o + \mathrm{NA}_c)^2} \quad , \tag{2.2}$$

where λ denotes the illumination wavelength and NA_c, NA_o are numerical apertures of the collector and objective optics. The thin sample approximation is well describing the experiment if the sample thickness $L \ll \mathrm{DoF}$.

For a flat sample with local refractive index $n(x, y) = 1 + \Delta n(x, y) = 1 - \delta(x, y) - \mathrm{i}\beta(x, y)$, we define its object complex transmission function as

$$O(x, y) \equiv \exp\left(\mathrm{i}k(\delta + \mathrm{i}\beta)\Delta z\right) \quad ,$$

where $\delta(x, y)$ and $\beta(x, y)$ are material constants. The material is expected to be homogeneous and thin along the z-axis. This is usually fulfilled in the EUV range, where due to high absorption, the thickness of the samples is limited to a few hundreds nanometers.

In the case of thin samples, the exiting light wave contains direct information about local attenuation and phase-shift. In the case of samples thicker than the depth of focus (DoF)

$$\mathrm{DoF} = \frac{\lambda}{\mathrm{NA}^2} \quad , \tag{2.3}$$

the light propagation within a sample can be modelled using multiple thin layers as discussed in Section 2.6.2.

If an ideal detector would be placed directly behind the sample, it would detect only attenuation. The wave field needs to be propagated from the sample to a certain distance, in order to measure the diffraction pattern containing mixed information from phase and amplitude.

2.1.1 Propagation of light waves in free space

The time evolution of the electric field $\mathbf{E}(\mathbf{r},t)$ and magnetic field $\mathbf{B}(\mathbf{r},t)$ can be described in free space at any point by the Maxwell's equations

$$\begin{aligned} \nabla \cdot \mathbf{E}(\mathbf{r},t) &= 0 \\ \nabla \cdot \mathbf{B}(\mathbf{r},t) &= 0 \\ \nabla \times \mathbf{E}(\mathbf{r},t) &= -\partial_t \mathbf{B}(\mathbf{r},t) \\ \nabla \times \mathbf{B}(\mathbf{r},t) &= \varepsilon_0 \mu_0 \partial_t \mathbf{E}(\mathbf{r},t) \quad , \end{aligned} \tag{2.4}$$

where $\varepsilon_0 = 8.854 \cdot 10^{-12}\mathrm{AsV^{-1}m^{-1}}$ and $\mu_0 = 4\pi \cdot 10^{-7}\mathrm{VsA^{-1}m^{-1}}$ are vacuum permittivity and permeability constants. The equations (2.4) allow to derive two independent wave equations

$$\begin{aligned} (\varepsilon_0 \mu_0 \partial_t^2 - \nabla^2) \cdot \mathbf{E}(\mathbf{r},t) &= 0 \\ (\varepsilon_0 \mu_0 \partial_t^2 - \nabla^2) \cdot \mathbf{B}(\mathbf{r},t) &= 0 \quad . \end{aligned} \tag{2.5}$$

Because all components of the electric and magnetic field obey the same equations, the vector equations (2.5) can be replaced by an equation of scalar field $\Psi(\mathbf{r},t)$ that obeys a scalar (D'Alembert) wave equation

$$\left(\frac{1}{c^2} \partial_t^2 - \nabla^2 \right) \cdot \Psi(\mathbf{r},t) = 0 \quad , \tag{2.6}$$

where c denotes speed of light in vacuum $c = 1/\sqrt{\varepsilon_0 \mu_0}$. The solution of the Eq. 2.6 can be decomposed into monochromatic time independent components Ψ_ω multiplied by a time dependent factor $\exp(-\mathrm{i}\omega t)$

$$\Psi(\mathbf{r},t) = \int_0^\infty \Psi_\omega(\mathbf{r}) \exp(-\mathrm{i}\omega t) d\omega \quad . \tag{2.7}$$

If this decomposition is inserted into the D'Alembert equation 2.6, it results in a time independent equation for monochromatic field that is called the Helmholtz equation

$$(\nabla^2 + k^2)\Psi_\omega(\mathbf{r}) = 0 \quad , \tag{2.8}$$

where vacuum wave-number k is defined as ω/c.

2.1.2 Diffraction formulas

Propagation of waves through an aperture can be considered as a boundary solution of the Helmholtz equation (Eq. 2.8) [36]. A common solution of the diffraction problem within the first-order Born approximation [21] is given by the Fresnel-Kirchhoff diffraction formula which provides the field $\psi_\omega(\mathbf{r})$ at a point $\mathbf{r}$ placed in the upper half-space ($z > 0$) assuming all sources are in the lower half-space ($z \leq 0$). Assuming the radiation sources are placed on a flat surface S_0, the Fresnel-Kirchhoff diffraction formula at point $\tilde{\mathbf{x}}$ can be written as

$$\Psi(\mathbf{x}) = -\frac{k}{2\pi \mathrm{i}} \int_{S_0} \int_S \Psi_0(\tilde{\mathbf{x}}) \frac{\exp(\mathrm{i}k|\mathbf{r}|)}{|\mathbf{r}|} \cos(\mathbf{n}, \mathbf{r}) dS_0 dS \quad , \tag{2.9}$$

where S_0 is the source surface, S is the incidence plane, $\mathbf{r}$ is the vector from source point $\tilde{\mathbf{x}}$ to destination point $\tilde{\mathbf{x}} + \mathbf{r}$, $\mathbf{n}$ is normal to the source surface and $\Psi_0(\tilde{\mathbf{x}})$ is the potential at the source surface. This equation was derived with assumption that the distance from the source is larger than the wavelength ($r \gg \lambda$), which is valid for coherent diffraction imaging.

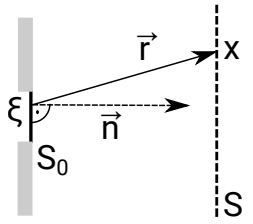

FIGURE 2.1: Coordinates used in the Fresnel-Kirchhoff diffraction formula (Eq. 2.10).

If the coordinates in the source plane S_0 are denoted ξ, η and in the incidence plane x, y, the Eq. 2.9 can be rewritten as

$$\Psi(x, y) = \frac{k}{2\pi \mathrm{i}} \int_{S_0} \Psi_0(\xi, \eta) \frac{e^{\mathrm{i}kr}}{r} \cos(\alpha) d\xi d\eta \quad , \tag{2.10}$$

where $\cos(\alpha)$ is called the directional cosinus that is effectively changing the source size (projection) dependent on the direction of $\mathbf{r}$ and r denotes distance

$$r = \sqrt{z^2 + (x-\xi)^2 + (y-\eta)^2} \quad ,$$

where z denotes the propagation distance. A common way to solve equation 2.10 is using the small angle (paraxial) approximation. Using the assumption that

$$\frac{|\xi - x|}{z} \ll 1, \quad \frac{|\eta - y|}{z} \ll 1 \quad , \tag{2.11}$$

the distance r can expressed as

$$r = \sqrt{z^2 + (x-\xi)^2 + (y-\eta)^2} \approx \left[z + \frac{(x-\xi)^2 + (y-\eta)^2}{2z} - \frac{\left((x-\xi)^2 + (y-\eta)^2\right)^2}{8z^3} \right] \quad . \tag{2.12}$$

Using only the first order approximation, the directional cosinus is approximated by 1. The field can be described by the Fresnel diffraction integral

$$\Psi(x,y,z) \approx \frac{e^{\mathrm{i}kz}}{\mathrm{i}\lambda z} \int_{S_0} \Psi_0(\xi,\eta) e^{\mathrm{i}\frac{k}{2z}\left[(x-\xi)^2 + (y-\eta)^2\right]} d\xi d\eta \quad , \tag{2.13}$$

The second order correction of phase is proportional $(x-\xi)^4/z^3$ as it can be seen from Eq. 2.12 and an additional correction factor for the intensity coming from the directional cosinus would be required.

Equation 2.13 can be rewritten as

$$\Psi(x,y,z) \approx \frac{e^{\mathrm{i}kz}}{\mathrm{i}\lambda z} e^{\mathrm{i}k\frac{x^2+y^2}{2z}} \int_{S_0} \Psi_0(\xi,\eta) e^{\mathrm{i}k\frac{\xi^2+\eta^2}{2z} - \mathrm{i}k\frac{\xi x + \eta y}{z}} d\xi d\eta \quad . \tag{2.14}$$

Using the definition of the Fourier transformation $\mathfrak{F}$

$$\tilde{f}(x) = \mathfrak{F}(f) = \int f(\xi) e^{-\mathrm{i}\pi x \xi} d\xi \quad , \tag{2.15}$$

Eq. 2.14 is equivalent to the following equation

$$\Psi(x,y,z) = \frac{e^{\mathrm{i}kz}}{\mathrm{i}\lambda z} e^{\mathrm{i}k\frac{x^2+y^2}{2z}} \mathfrak{F}\left(\Psi_0(\xi,\eta) e^{\mathrm{i}k\frac{\xi^2+\eta^2}{2z}} \right) \tag{2.16}$$

that can be very effectively calculated using the Fast Fourier Transformation

(FFT) algorithm. To summarize the Fresnel equation derivation, Eq. 2.16 describes field distribution everywhere in the upper half-space ($z > 0$) at a distance from the sample much larger than the wavelength ($z \gg \lambda$) and near to the optical axis ($|x - \xi|/z \ll 1$).

However, in some cases, an additional approximation can be beneficial. Often the illuminated region has a finite support size a and the distance between the sample and the detector z can be considered much larger then the size of the transmissive/illuminated region of the sample a. In that case, it is possible to define a unitless constant F

$$F = \frac{a^2}{\lambda z} \ll 1 \tag{2.17}$$

known as the Fresnel number. A small Fresnel number $F \ll 1$ corresponds to the far-field regime, while large $F \gg 1$ corresponds to the so-called near-field regime.

In the far-field (Fraunhofer) regime ($F \ll 1$), the curvature factor inside the Fourier transformation in Eq. 2.16 can be neglected. The resulting amplitude distribution of the field depends only on the scattering angle and it can be written as

$$\Psi(x, y, z) \approx \frac{e^{\mathrm{i}kz}}{\mathrm{i}\lambda z} \int\int_{-\infty}^{\infty} \Psi_0(\xi, \eta) e^{\frac{2\pi\mathrm{i}}{z\lambda}(x\xi + y\eta)} d\xi d\eta \sim \mathfrak{F}\left(\Psi_0(\xi, \eta)\right) \quad . \tag{2.18}$$

This equation is known as the Fraunhofer approximation.

An alternative approach to calculate the field distribution directly from the Helmholtz equation 2.8 is called the *angular spectrum method* (ASM) [36]. A monochromatic wave $\Psi(\mathbf{r}_\perp, z)$ can be decomposed laterally into spectral components $\tilde{\Psi}(\mathbf{k}_\perp, z)$

$$\Psi(\mathbf{r}_\perp, z) = \frac{1}{2\pi} \int \tilde{\Psi}(\mathbf{k}_\perp, z) \exp[\mathrm{i}(\mathbf{k}_\perp \cdot \mathbf{r}_\perp)] \mathrm{d}\mathbf{k}_\perp \quad , \tag{2.19}$$

where $\mathbf{k}_\perp$ denotes a wave vector perpendicular to the direction of propagation $\mathbf{k}$ and $\mathbf{r}_\perp$ are coordinates in the plane perpendicular to the propagation direction. As $\tilde{\Psi}(r_\perp, z)$ obeys the Helmholtz equation, if the decomposition Eq. (2.19) is inserted into the Eq. 2.8, the following equation has to be fulfilled

$$\int \left[\partial_z^2 \tilde{\Psi}(\mathbf{k}_\perp, z) + (k^2 - k_\perp^2)\tilde{\Psi}\right] \exp[\mathrm{i}(\mathbf{k}_\perp \cdot \mathbf{r}_\perp)] \mathrm{d}\mathbf{k}_\perp = 0 \quad . \tag{2.20}$$

Because only the Fourier transformation of zero function is zero, the forward propagation solution is given as

$$\tilde{\Psi}(\mathbf{k}_\perp, z) = \tilde{\Psi}_0(\mathbf{k}_\perp) \exp\left(\mathrm{i}z\sqrt{k^2 - k_\perp^2}\right) \quad . \tag{2.21}$$

If $k^2 > k_\perp^2$, then the propagation is simply a change of the relative phases of different components of the angular spectrum. However, when $k^2 < k_\perp^2$, then the angular components are rapidly attenuated during the propagation. These components are called evanescent waves and they extend only to a distance of a few wavelengths. This is the reason, why the standard microscopy methods with working distance significantly longer then the wavelength are not able to resolve smaller structures than the wavelength of the used illumination.

Equation (2.21) can be numerically solved using the real-space field Ψ_ω and the Fourier transformation

$$\Psi_\omega(\mathbf{r}_\perp, z) = \mathfrak{F}^{-1}\left[\mathfrak{F}\left(\Psi_0(\mathbf{r}_\perp)\right) \exp\left(\mathrm{i}z\sqrt{k^2 - k_\perp^2}\right)\right] \quad . \tag{2.22}$$

This equation is called angular spectrum propagation method (ASM). The only assumption needed to derive Eq. (2.22) is that the propagation distance is much longer than the wavelength. There is no paraxial approximation used, therefore Eq. (2.22) should be valid also for experiments with large numerical aperture (NA). However, if an analytical solution of the Eq.(2.22) is needed, it is possible to get a simpler expression using the Taylor expansion of the square root term

$$\Psi(\mathbf{r}_\perp, z) \approx \exp(\mathrm{i}kz)\mathfrak{F}^{-1}\left[\mathfrak{F}\left(\Psi_0(\mathbf{r}_\perp)\right) \exp\left(\frac{-\mathrm{i}k_\perp^2 z}{2k}\right)\right] \quad , \tag{2.23}$$

where $k_\perp$ denotes $\sqrt{k^2 - k_z^2}$. The integral can be also expressed in an alternative form using convolution as

$$\Psi(\mathbf{r}_\perp, z) \approx \Psi_0(\mathbf{r}_\perp) * h(\mathbf{r}_\perp, z) \quad , \tag{2.24}$$

where $*$ denotes the convolution operator and h denotes a convolution kernel

$$h(\mathbf{r}_\perp, z) = \frac{e^{\mathrm{i}kz}}{\mathrm{i}\lambda z} e^{\mathrm{i}\frac{k}{2z}\mathbf{r}_\perp^2} \quad . \tag{2.25}$$

The convolution form of the propagator in near-field can be understood as an edge enhancing filtering kernel in the phase domain which is causing fringes around

sharp objects when the light is propagated into sufficient distance behind the sample.

Limitations of the ASM method

ASM is theoretically valid even for large propagation distances, however if the propagation is calculated on a discrete pixel size using the FFT algorithm, the propagation distance is limited by the numerical precision and sufficient oversampling of the propagation matrix $\exp\left(\mathrm{i}z\sqrt{k^2-k_\perp^2}\right)$ which avoids to use ASM for distances significantly larger than the sample size. This is the reason, why ASM is sometimes called the near-field propagation.

If the discrete Fourier transformation based ASM method (Eq. (2.22)) is used, the propagating light will be calculated with a periodic boundary constraint. Therefore if the field is propagated too far, the components corresponding to the high spatial frequency will be wrapped around the propagated array. The advantage is that the ASM method is lossless (up to numerical errors) and an array that is propagated into some distance and then back-propagated will be unchanged. The drawback is that it is unphysical. If this is an issue, it is either possible to pad the array by larger empty space or apply a filter in the Fourier domain that will remove the high frequencies using the following equation.

$$\Psi_\omega(\mathbf{r}_\perp, z) = \mathfrak{F}^{-1}\left[HW\mathfrak{F}\left(\Psi_0(\mathbf{r}_\perp)\right)\right] \quad , \tag{2.26}$$

where $H = \exp\left(\mathrm{i}z\sqrt{k^2-k_\perp^2}\right)$ is a unitary matrix and the matrix $0 < W \leq 1$ serves as a positive value spectral filter that removes the high frequencies that are wrapped around due to the periodic boundary condition.

2.1.3 Effective near-field propagation: The Fresnel scaling theorem

Equation 2.16 has one very interesting property that can be seen if the sample is illuminated by a spherical wavefront. The relation between diffraction pattern from a plane-wave illuminated sample $|\Psi^{(\infty)}|^2$ and the pattern from a spherically

illuminated sample $|\Psi^{(R)}|^2$ is given by so called Fresnel scaling theorem [37]

$$|\Psi^{(R)}(\mathbf{r}_\perp, z \geq 0)| = M^{-2} \left|\Psi^{(\infty)}\left(\frac{\mathbf{r}_\perp}{M}, \frac{z}{M} \geq 0\right)\right| \quad , \tag{2.27}$$

where M denotes the geometric magnification factor. This relation is valid only for small angles from the beam propagation direction (paraxial approximation). Equation (2.27) can be obtained if $\psi^{(R)}(\mathbf{r}_\perp, 0)$ is a nearly spherical wavefront with wavefront curvature R that can be approximated as

$$\psi^{(R)}(\mathbf{r}_\perp, 0) = \psi^{(\infty)}(\mathbf{r}_\perp, 0)\frac{\exp(\mathrm{i}kR)}{R}\exp\left(\frac{\mathrm{i}kr_\perp^2}{2R}\right) \quad . \tag{2.28}$$

If this wavefront is inserted into the Fresnel equation (Eq. 2.16), then the result is

$$\Psi(\mathbf{r}'_\perp, z) = \frac{\exp(\mathrm{i}kz)}{\mathrm{i}\lambda z}\exp\frac{\mathrm{i}k{r'_\perp}^2}{2z}\mathfrak{F}\left[\Psi_0(\mathbf{r}_\perp, 0)\exp\left(\frac{\mathrm{i}kr_\perp^2}{2}\left(\frac{1}{z}+\frac{1}{R}\right)\right)\right] \quad . \tag{2.29}$$

If the geometric magnification factor M is introduced as

$$M = \frac{R+z}{R} \tag{2.30}$$

and the effective propagation distance $z_{\mathrm{eff}} = z/M$, then the amplitude of the field ψ can be written as

$$|\Psi(\mathbf{r'}_\perp, z)| = \frac{1}{\lambda z_{\mathrm{eff}} M}\left|\mathfrak{F}_{1/M}\left[\Psi_0(\mathbf{r}_\perp, 0)\exp\left(\frac{\mathrm{i}kr_\perp^2}{2z_{\mathrm{eff}}}\right)\right]\right| \quad , \tag{2.31}$$

where $\mathfrak{F}_{1/M}$ denotes the Fourier transformation in coordinates ($\mathbf{r'}_\perp = \mathbf{r}_\perp/M$). The Fresnel scaling theorem (Eq. 2.27) can be obtained by comparison of Eq. 2.31 and plane wave propagated by Eq. (2.16) to distance $z = z_{\mathrm{eff}}$. A similar idea is also valid for the ASM propagation, as it was experimentally shown in Ref. [38].

2.1.4 Extending propagation formulas beyond the paraxial approximation

Although there are no low-angle approximations in the ASM method, it is not commonly used for high resolution CDI imaging because of the relatively large effective distance between the sample and the camera (i.e. Fresnel number $F \ll 1$). On the other hand, the Fresnel method 2.13 is valid only near to the optical axis.

Resolution of the diffraction limited bright-field microscopy with circular entrance pupil is given by the Rayleigh resolution [21]

$$\delta x = \frac{1.22\lambda}{\mathrm{NA}_{\mathrm{condenser}} + \mathrm{NA}_{\mathrm{objective}}} \quad .$$

The Rayleigh resolution, which will be more discussed in Section 2.2.1, shows that in order to achieve high resolution microscopy for a given wavelength and medium, it is necessary to collect high scattering angles. This requires the study of the limitations given by the paraxial approximation and finding corrections that allow the extension of CDI beyond these limits.

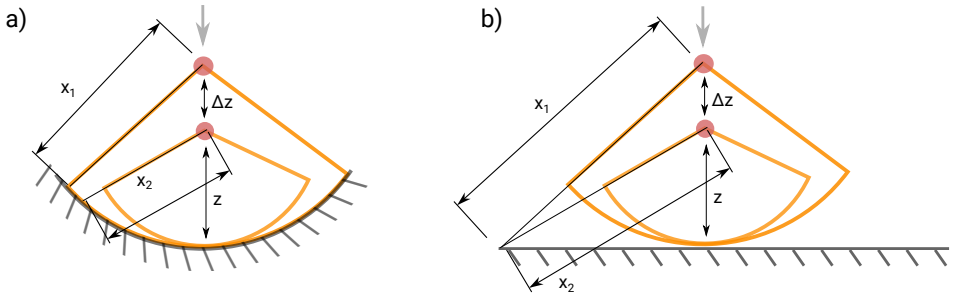

FIGURE 2.2: Schema of a simple diffraction experiment with illumination incident normal to the detector. Left image shows the case of an ideally spherical detector (e.g. optical microscope), the right image shows case of a flat detector.

In a standard high resolution CDI microscopy experiment in the EUV range, the propagated wave can be considered to be fully in the far-field regime and the illumination source size is negligible compared to the detector size. However, the collected numerical aperture (NA) can be up to 0.6 [39], therefore the paraxial approximation $\sqrt{x^2+y^2} \ll z$ is not valid any more. The two common configurations are shown in the Fig. 2.2. Figure 2.2a) shows the case of imaging with a "curved detector". This is a common configuration if a microscopy lens with flat-field correction is used to image the sample. The second case, shown in Fig. 2.2b), is a standard CDI setup, when the diffracted light is collected by a flat detector.

In the case of a "curved detector" in the far-field regime, Eq. 2.16 with a slightly modified phase curvature prefactor is valid even for high NA

$$\Psi(x,y,z) = \frac{e^{ikz}}{i\lambda z}\mathfrak{F}\left(\Psi_0(\xi,\eta)\right) \quad . \tag{2.32}$$

In the case of a flat detector, the high NA effects need to be included. The effects can be understood if higher orders of the Fresnel approximation (Eq. (2.12)) are

included. Because the detector is in the far-field regime, all the terms proportional and smaller then $\frac{|\max(\eta,\xi)|^2}{z^2}$ can be neglected. The following equation expresses the field intensity

$$|\Psi(x,y,z)|^2 \approx \frac{\cos^2\alpha}{R^2}\left|\int_{S_0}\Psi_0(\xi,\eta)\exp\left(-\mathrm{i}k\frac{\xi\left(x-\frac{xy^2+y^3}{2z^2}\right)+\eta\left(y-\frac{x^2y+x^3}{2z^2}\right)}{z}\right)d\xi d\eta\right|^2 \quad . \tag{2.33}$$

where $R^2 = x^2+y^2+z^2$ causes intensity decay for larger propagation distance. The mixed higher order terms effectively result in a slower evolution of the frequency terms ξx and ηy which means that the observed speckles for large angles will be slightly distorted (stretched) as it is shown in Fig. 2.3(b). This prevents to use the standard fast Fourier transformation to calculate the Eq. (2.33) because the FFT requires sampling on a regular grid. One option would be to calculate the integral (2.33) using the nonuniform Fourier transformation [40], which is however computationally significantly slower than the standard FFT algorithm. The second option is to re-interpolate the measured intensity to a uniform grid.

If the detector is perpendicular to the incidence beam, the corrected normalized radial coordinates $\tilde{\mathbf{r}}_{\perp\mathrm{corr}}$ (i.e. $|\tilde{\mathbf{r}}_\perp| = 1$ for edge of the detector) can be expressed using the original normalized coordinates $\tilde{\mathbf{r}}_\perp$ as

$$\tilde{\mathbf{r}}_{\perp\mathrm{corr}} = \frac{\tilde{\mathbf{r}}_\perp}{\sqrt{1+(\tilde{r}_\perp \mathrm{NA})^2}} \tag{2.34}$$

with intensity correction I_{corr}

$$I_{\mathrm{corr}} = I\left(1+\frac{r_\perp^2}{z^2}\right) \quad . \tag{2.35}$$

Examples of the distorted Cartesian grids are shown in Fig. 2.3.

These corrections are only valid for thin samples. In the case of thick samples, multiple scattering inside the sample cannot be neglected. If the sample is shifted by the distance Δz allong the beam axis as it is illustrated in the Fig. 2.2, a phase shift is introduced for a given scattering angle that can be calculated as path difference between $|x_1|$ and $|\Delta z + x_2|$. For the ideal "curved" detector, the phase shift difference between the layers $\Delta\varphi$ can be approximated as

$$\Delta\varphi = k\frac{\Delta z}{2}\mathrm{NA}^2 \quad , \tag{2.36}$$

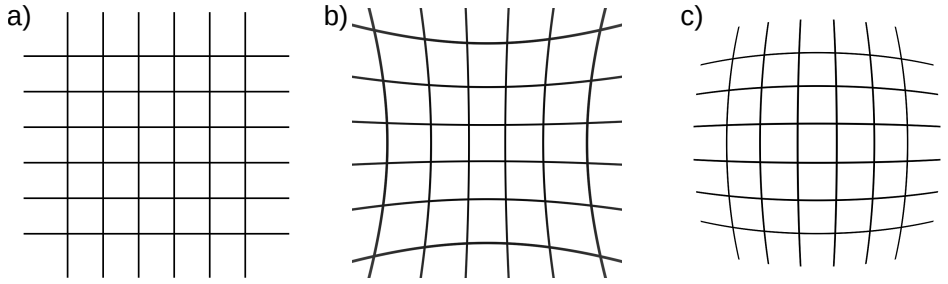

FIGURE 2.3: An example of a Cartesian grid (a) and a non-Cartesian grid coordinates (b) of the flat detector resulting from high NA diffraction. Image (c) shows the transformation that needs to be applied in order to correct (b) to linear coordinates (a).

which for $\Delta\varphi = \pi$ corresponds to the expected depth of focus (DoF) [41]

$$\mathrm{DoF} = \frac{\lambda}{\mathrm{NA}^2} \quad . \tag{2.37}$$

Therefore, a shift of the sample by distance Δz can be understood as propagation to the detector by the Fraunhofer equation 2.18, phase correction using the Eq. 2.36 to a different Ewald sphere and back propagation. Note that this correction is equivalent to the paraxial approximation of ASM propagation (Eq. 2.23). In case of a flat detector (Fig. 2.2b), the idea is very similar. The phase difference $\Delta\varphi$ can be approximated as

$$\Delta\varphi = k\left(\sqrt{1+\mathrm{NA}^2}\Delta z - \Delta z\right) \tag{2.38}$$

Figure 2.4 shows phase correction for the numerical aperture NA=0.4 and $\Delta z = \lambda/(\mathrm{NA})^2$. In the case of the small angle approximation (NA<0.1), the phase correction for a flat and a curved detector are identical as expected. However, for higher NA, the phase on the flat detector is changing slower, therefore the DoF given by the Eq. 2.37 is not valid any more and the flat detector experiment will be less sensitive to sample thickness than an experiment with an ideal curved detector. This slightly relaxes the constraints on sample flatness and thickness for high NA CDI microscopy.

The second effect of the sample shift along the z-axis is scaling of the diffraction pattern as it can be seen from the Eq. 2.18. When the sample is shifted by distance Δz, then for the same pixel grid of the detector, the sample will be sampled slightly finer by a factor of $1 + \Delta z/z$. The consequence is that the parts of the sample closer to the detector will appear slightly larger. This effect is usually negligible

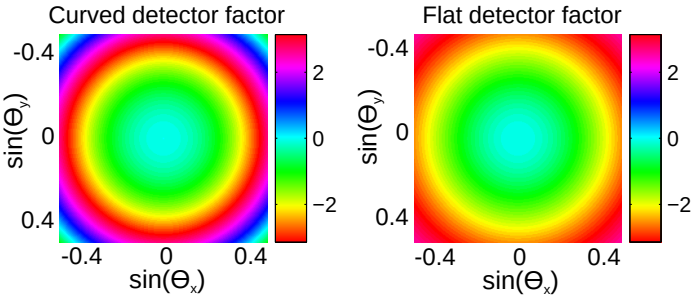

FIGURE 2.4: Phase correction factor for angular spectrum propagation on a curved detector (left) and a flat detector (right) simulated for NA=0.4. Propagation distance was equal to depth of focus ($\Delta z = \lambda/\mathrm{NA}^2$).

in the EUV due to low penetration if the sample is sufficiently flat, i.e. the sample to illumination forming optics distance changes less than the DoF.

2.1.5 Off-axis illumination in far-field diffraction

A more general case of diffraction is obtained if the angle of the sample with respect to the illumination is arbitrary. In order to simplify the calculations, the sample is expected to be sufficiently thin. In both cases of transmission and reflection, it is expected that multiple scattering can be neglected.

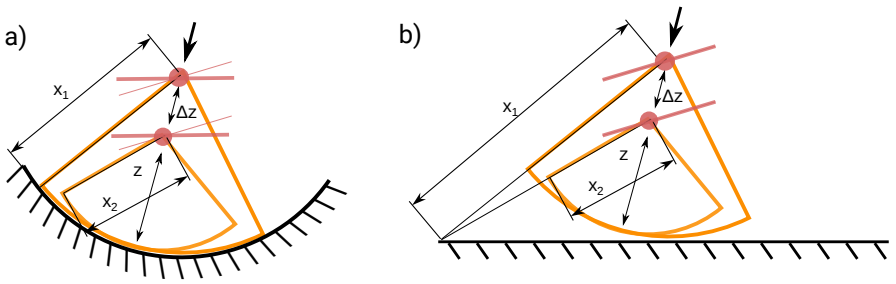

FIGURE 2.5: A schema of a diffraction experiment when the illumination incidences on the sample non-normally. Image (a) shows the case of an ideally spherical detector (e.g. optical microscope with flat plane correction) with different sample orientation and image (b) shows the same but for a flat detector.

The simplest case is off-axis, close to flat illumination of a sample imaged with an ideal curved detector (Fig. 2.5(a)). The only effect will be a shift of the far-field diffraction pattern dependent on the direction of the incident illumination wave. If the size of the curved detector is finite, then only a limited fraction of the far-field diffraction will be collected. This principle of oblique illumination is used by the Fourier ptychography [2, 26, 27] to extend the NA of the imaging optics.

In the more general case with arbitrary angle of the sample with respect to a flat detector (Fig. 2.5(b)), there is an additional effect of a different propagating distance in dependence on the sample orientation. The tilted plane correction (Fig. 2.6) can be separated into two steps:

1. Interpolation of the measured data onto the Ewald sphere in the case of high NA geometry using Eq. 2.34
2. Tilted sample correction described in Ref. [42–44]

If the sample is rotated about the y-axis by angle α, then the transformation from detector coordinate system $\vec{q}(x, y, z)$ to sample coordinate system $\vec{q}(x', y', z')$ is given by [42]

$$\vec{q}(x', y', z') = (q_x \cos\alpha - q_z \sin\alpha, q_y, q_z \cos\alpha) \quad . \tag{2.39}$$

The effects of off-axis illumination of an thin sample are shown in Fig. 2.7. The propagation effects were simulated by ASM propagation on a tilted plane and then far-field propagation on the detector.

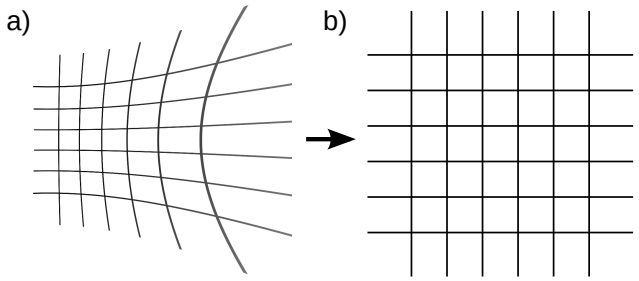

FIGURE 2.6: Distorted coordinates by off-axis diffraction (a) need to be corrected to the linear coordinate system (b).

The off-axis correction is generally valid for both transmission and reflection if the sample is sufficiently flat. However, the relation between the reconstructed exit wave field and the material constants is more complex for off-axis reflection geometry [28, 29]. More rigorous derivations but with the same conclusions can be found in [28, 42–44].

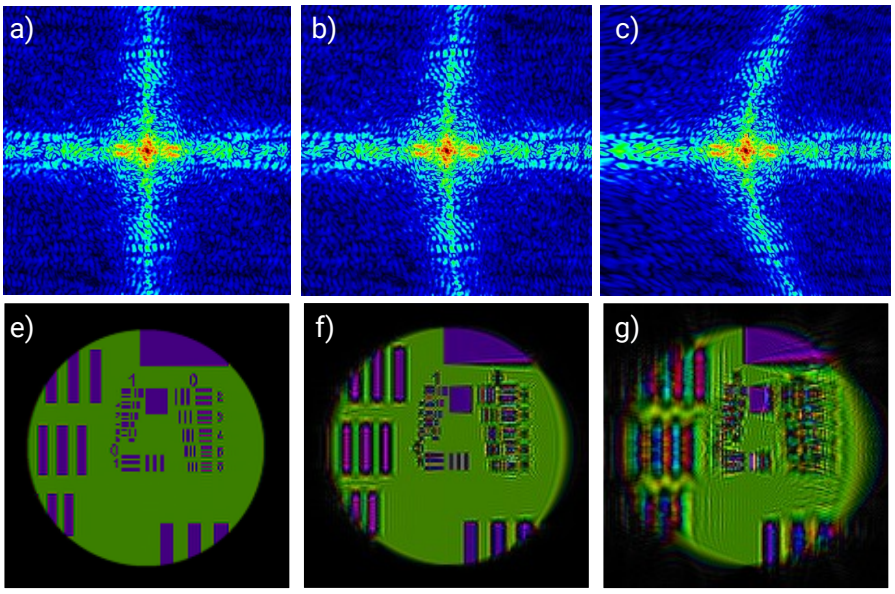

FIGURE 2.7: (a-c) Simulation of the angular spectrum intensity distributions in the logarithmic colour-scale for different rotations of a sample (0°, 22.5°, 45°) and corresponding complex electric field in the original (unrotated) plane (e-g). Simulation parameters: $\lambda = 30\,\mathrm{nm}$, NA=0.5, illumination diameter is $10\,\mu\mathrm{m}$.

2.2 Background of coherence diffractive imaging

2.2.1 Diffraction limit

Resolution of CDI methods is limited by the maximal diffraction angle passing the imaging system with sufficient signal to noise ratio, similarly to most of the other microscopy methods. The Rayleigh resolution of a lens-based incoherent microscopy system with circular aperture is given by [21]

$$\Delta p = \frac{1.22\lambda}{\mathrm{NA}_O + \mathrm{NA}_C} \quad , \tag{2.40}$$

where Δp is the smallest resolvable period (i.e. 2 pixels), λ is the imaging wavelength and $\mathrm{NA}_O, \mathrm{NA}_C$ are effective numerical apertures of the objective or collector respectively. In the case of coherent illumination, the resolution for circular aperture system is more limited and it is given by the Abbe criterion [21]

$$\Delta p = \frac{1.64\lambda}{\mathrm{NA}_O + \mathrm{NA}_C} \quad , \tag{2.41}$$

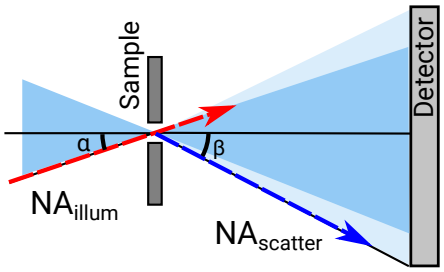

FIGURE 2.8: An illustration showing effect of large illumination NA on the achievable resolution. The largest angle collected by the detector is limited by sum of scattering angles $\alpha + \beta$ if the signal intensity is sufficient.

CDI methods provide the reconstruction on a regular grid with the pixel size Δx given for small scattering angles by the formula

$$\Delta x = \frac{\lambda}{2\mathrm{NA}_O} = \frac{\lambda z}{Np} \quad , \tag{2.42}$$

where N is number of pixels, p is size of the camera pixel and z is the sample to camera distance. In the case of larger diffraction angles, it is necessary to use the precise formula

$$\Delta x = \frac{\lambda z}{Np}\sqrt{1 + \left(\frac{Np}{2z}\right)^2} \quad . \tag{2.43}$$

This reconstruction grid samples sufficiently the imaged object with the resolution given by the Abbe criterion 2.41 for the case of flat illumination wavefront. This means that NA_C is negligible and only NA_O determines the resolution. This is a common case for the standard single diffraction CDI. In the case of scanning CDI (ptychography), NA of the illumination cannot be neglected [2, 26, 27, 45]. Contrary to the traditional lens-based microscopy, the NA_C is not defined by the illumination cone angle, but rather by the maximum illumination angle that is given by any illumination forming optics such as a pinhole, an iris or some diffusion element as it is shown in Fig. 2.8. If the illumination NA is not negligible, then there will exist features in the diffraction pattern that cannot be described by a reconstruction and a probe sampled on a grid with the step size given by Eq. 2.42.

If the signal or any other sources of errors are not limiting, then the maximum resolution that can be recovered is given by Eq. 2.41 and it cannot be sufficiently described by a solution sampled on grid with pixel size given by Eq. 2.42. Therefore a reconstruction on an up-sampled grid can provide higher resolution and thus better describe the measured data. This idea is used by the "super-resolution" ptychography method [45] that uses an up-sampled grid to improve the resolution

up to the diffraction limit if the signal is sufficient and other systematic errors are negligible.

However, SNR is usually the major resolution limitation in the short wavelength microscopy. According to the Rose criterion, the SNR should be larger than 5 in order to clearly distinguish the measured features. This is not entirely true for the scanning CDI methods when several overlapping exposures of adjacent scanning positions improve the achievable resolution as it is shown in Ref. [46] and the resolution is given rather by the total number of photons passing though the sample than by information content in a single diffraction pattern. This idea is similar to the dose-fractionation theorem in tomography [47]. However, the maximum scattering angle, where SNR is above one is usually a good estimate of the theoretically achievable resolution in the case of negligible readout noise.

It is also important to consider that the resolution significantly depends on the sample and it can vary for different parts of the sample in dependence on transmission and contrast. Absorption contrast C_a in bright-field microscopy is defined as

$$C_a = \frac{|N_{\text{feature}} - N_{\text{bg}}|}{N_{\text{bg}}} \quad , \tag{2.44}$$

where N_{feature} is number of photons passing through the smallest resolvable feature and N_{bg} is number of photons passing through a pixel around the feature.

As we will show in Chapter 5, it is relatively easy to get a high resolution reconstruction of strongly scattering binary samples, while achieving the same resolution with weakly scattering biological samples is significantly more challenging.

Further effects limiting the resolution can be temporal coherence that is limiting reconstruction complexity. That means the diameter of illumination divided by resolution (pixel size) of the reconstruction as it is mentioned in Section 2.2.3. Another common resolution limitations are:

1. dynamic range of the collected diffraction pattern
2. precision of the sample shifting actuators
3. precise knowledge of the experimental geometry
4. aberrations of optics between sample and camera

These limitations will be further discussed in the experimental chapters.

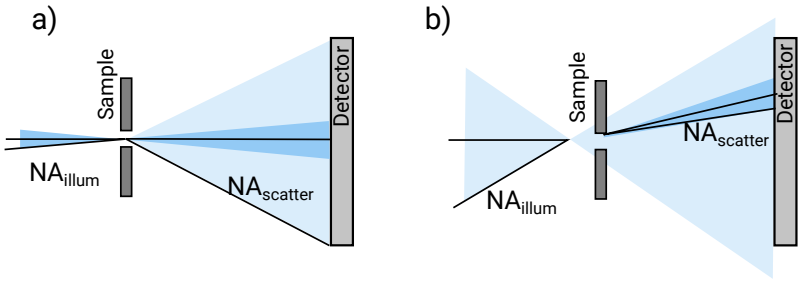

FIGURE 2.9: Schematics of a far-field CDI setup (a) and an effectively near-field CDI setup with high NA illumination (b).

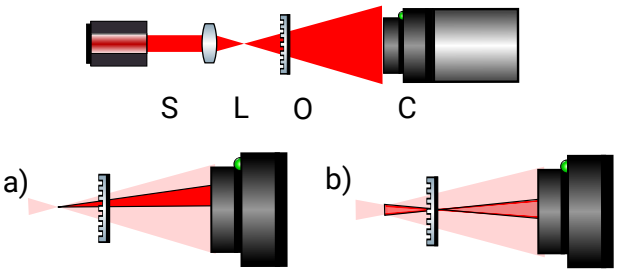

FIGURE 2.10: A simplified schema of a near-field imaging setup. Light source (S) is focused by a lens or a mirror (L) into a smaller virtual source. The expanding light passes through the sample (O) and it is captured by camera (C). Figure (a) shows resolution limited by the pixel size on the camera while (b) is the virtual source size (coherence) limited resolution.

2.2.2 Coherence

Coherence expresses the degree of correlation between two points in space and time. Fully incoherent radiation eliminates interference effects that are required in CDI. Therefore CDI needs sufficient level of coherence in order to find a unique solution. Correlation in the direction of propagation is called longitudinal or temporal coherence and it depends on the bandwidth of the illumination spectrum. Spatial coherence measures correlation in the orthogonal plane to the propagation direction and it is also sometimes called transverse coherence. Mathematically, the coherence can be described by a mutual coherence function Γ defined as the time averaged correlation of the electric field at two points as a function of their separation [21, 41]

$$\Gamma_{12}(\tau) = \langle E_1(t+\tau)E_2^*(t)\rangle \tag{2.45}$$

where E_1, E_2 denotes electric fields at points 1, 2 and τ is their time delay. The mutual coherence function is usually normalized by the average intensity of the

electric field and the result is called complex degree of coherence γ_{12}

$$\gamma_{12}(\tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{\langle |E_1|^2 \rangle \langle |E_2|^2 \rangle}} \quad . \tag{2.46}$$

Light is fully coherent if $|\gamma_{12}| = 1$ everywhere and incoherent if $|\gamma_{12}|$ approaches zero. Real sources are never fully coherent, although continuous wave lasers can often be close to this limit. The temporal coherence is often expressed using the coherence length l_{coh} that is defined for sources with the spectral width $\Delta\lambda$ as

$$l_{\mathrm{coh}} = \frac{\lambda^2}{2\Delta\lambda} \quad . \tag{2.47}$$

This temporal coherence corresponds to the distance, where two waves separated by $\Delta\lambda$ and with initially identical phase will have the opposite phase.

The spatial coherence expresses, how small should be the radiation source to produce a sufficiently coherent wavefront between two points [41]. If the source size is so small that it cannot be distinguished from a point source, then the light can be considered to be coherent. For given observation half-angle Θ (or numerical aperture), the maximum size d of a partially coherent source can be

$$d = \frac{\lambda}{2\pi\theta} \quad . \tag{2.48}$$

The transverse coherence length is defined for convenience and similarity with the temporal coherence. If one defines $l_{\mathrm{transverse}} = z\theta$, where z is the propagation distance, then the transverse coherence length is expressed as

$$l_{\mathrm{transverse}} = \frac{z\lambda}{2\pi d} \quad . \tag{2.49}$$

Using the van Citter-Zernike theorem [21], the transverse coherence length $l_{\mathrm{transverse}}$ can be related to the complex degree of coherence γ_{12} for a Gaussian shaped source as [41]

$$|\gamma(l_{\mathrm{transverse}})| = 0.88 \quad . \tag{2.50}$$

The absolute value of the degree of coherence $|\gamma|$ (Eq. 2.46) determines the contrast of the fringes in the measured far-field diffraction pattern. The fringe contrast (visibility) is defined as

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad , \tag{2.51}$$

where $I_{\max}$ is maximum and $I_{\min}$ is the minimal intensity of the measured fringes. In real measurements, it is necessary to measure the visibility on a sufficiently fine grid in order to reduce the effect of the finite pixel size.

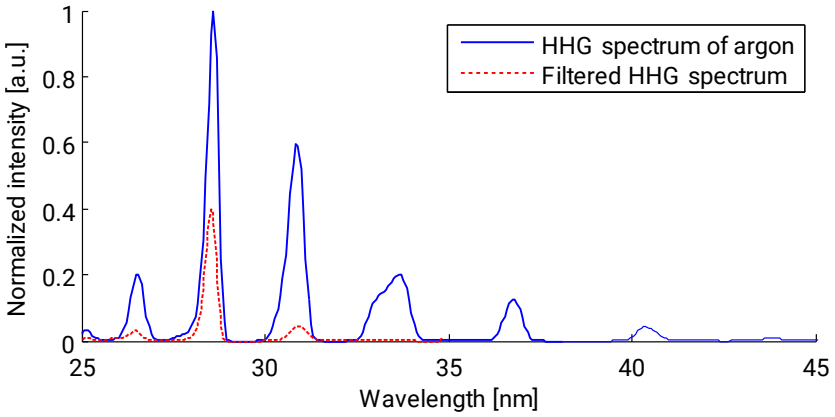

FIGURE 2.11: Examples of the HHG spectra reconstructed from the Young's double slits diffraction pattern. The blue line is HHG spectrum without filtering by a single multilayer mirror and the red line is after filtering. The maximum spectral resolution of our this method was in our setup around 0.5 nm (see Chapter 5.5.1).

Transverse coherence of our illumination HHG beam in focus was estimated to 7 μm [11] with the beam size $\approx 15\,\mu$m. The temporal coherence is determined by number of harmonic orders that are generated (Fig. 2.11). If the HHG radiation is used for imaging and separation of the harmonics is sufficient, single harmonic can be selected and focused by a multilayer mirror or a grating monochromator. The measured width of the peaks in Fig. 2.11 is diffraction limited by resolution of our Young's double slits spectrometer to 0.5 nm (see Chapter 5.5.1), however we have observed speckles even for large diffraction angles NA$\approx$ 0.4.

2.2.3 Coherence requirements for CDI

The requirements for CDI are relatively high compared to other microscopy techniques. The recommended lateral coherence length for far-field CDI is twice the object size D [48] which is however quite conservative and usually the coherence is lower in laboratory experiments. The relation for temporal coherence in far-field CDI is slightly different because the blurring effect of temporal coherence depends on the path difference and it differs for each part of the diffraction patten. The temporal coherence is giving a limit on complexity of the reconstruction that given

as

$$\frac{D}{\Delta D} \leq \frac{2\lambda}{\Delta\lambda} \quad , \tag{2.52}$$

where ΔD is the pixel size and $\frac{D}{\Delta D}$ is number of reconstructed pixels.

These limits are only recommendations and multiple relaxation methods were proposed to reduce these strong assumptions (see Section 2.6). These relaxation methods can help with the reconstruction, however they usually lead to slower convergence and higher SNR requirements. Furthermore, if the coherence limits are significantly exceeded, the CDI reconstruction will fail to converge.

Examples of different incoherent effects in the far-field regime are shown in Fig. 2.12.

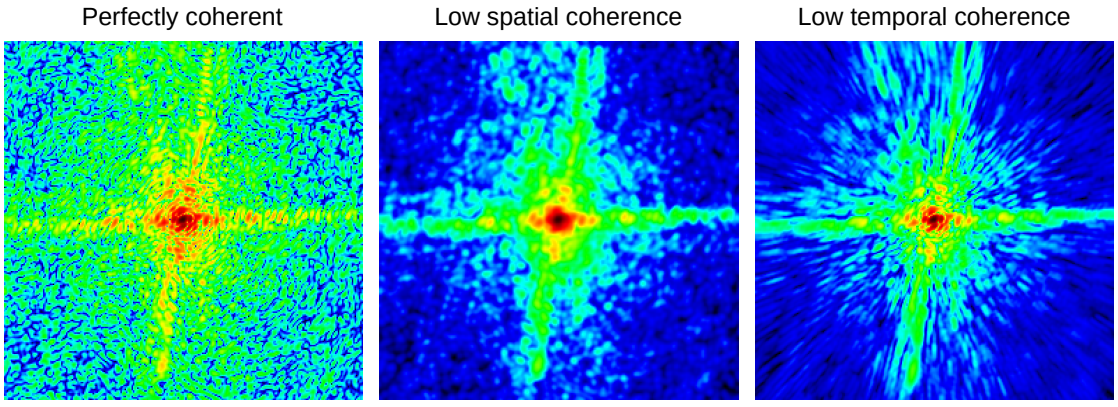

FIGURE 2.12: Simulations of influence of temporal and spatial incoherence on the collected far-field diffraction pattern. Effects similar to the spatial incoherence can be also observed in case of the illumination beam or the sample movements faster then the exposure time.

In the far-field mode, the spatial incoherence can be modelled as a convolution with a positive kernel [49] or equivalently described as an incoherent sum of multiple illumination waves [16, 50]. The effect of convolution (blurring) in the Fourier space can be understood as suppressing values far from the centre in the real-space domain and therefore limiting the maximum extent of the reconstructed object.

Temporal incoherence can be modelled as an incoherent sum of diffraction patterns for each wavelength in the far-field regime. If the transmitting sample is achromatic, then the diffraction pattern at wavelength λ is equal to the monochromatic pattern at the central wavelength scaled by the factor λ/λ_0, where λ_0 is the central wavelength [51]. If the sample is not sufficiently achromatic or the diffraction

pattern cannot be considered to be in far-field regime, then full-multicolour ptychography should be used [52].

Both these incoherence effects are in the far-field regime effectively simple linear operations, therefore it can be removed by direct deconvolution or unfolding methods shown in Section 2.6, however only if the signal to noise ratio is sufficiently high and convolution kernel and spectrum are known.

Finally, the resolution of a near-field experiment is limited by the longitudinal coherence as well. For given full-period resolution Δx the required bandwidth $\Delta\lambda/\lambda$ can be expressed [53]

$$\frac{\Delta\lambda}{\lambda} \leq \frac{(2\Delta x)^2 M}{\lambda z} \quad . \tag{2.53}$$

If we consider an imaging setup with magnification $M = 200$, wavelength $\lambda = 13\,\text{nm}$, sample to camera distance $z = 2\,\text{cm}$ and bandwidth 2%, the resolution limit would be 80 nm.

Ideally, the coherence should be improved in the experiment. The spatial coherence can be increased by a pinhole in sufficient distance that will effectively reduce the size of the source. The spectral bandwidth can be improved by multilayer mirrors, spectral gratings or just spectral filters. Multilayer mirrors are mostly used in the EUV range for spectral filtering, however it is challenging to produce them for wavelengths below 5 nm.

In practice, there are also other systematic errors that can be modelled as low coherence. The most common example is the finite size of the pixels on a camera or point spread function of the scintillator coupled CCD [50]. The pixels cannot be considered to be points but they collect signal from a finite region. If oversampling of the diffraction pattern is not sufficient, the lost of visibility effect will be similar to the spatial incoherence. Another example are variations of the sample or source faster then the exposure time [16, 54] which also lead to blurring and similar methods can be applied.

2.2.4 Noise and other sources of errors

Every microscopy method collects data for a limited amount of time and reads out the signal from a sensor by electronics with limited sensitivity and dynamic range.

Therefore, noise will be always present in the measured data, however particularly in visible light bright field microscopy it may not be the main source of errors. The errors can be divided into two groups: random and systematic.

Random errors

There are three main sources of random errors caused by limitations of the detectors: Poisson noise of the collected photons (photon noise), readout noise and Poisson noise of the dark current after subtraction. In the case of short wavelength microscopy, the photon noise is dominating at least for pixels with more than a few photons. The photon noise can be approximated for higher number of collected photons N by

$$\sigma_{\text{photon}} = \sqrt{N} \quad , \tag{2.54}$$

therefore signal to noise ratio (SNR) is proportional to $1/\sqrt{N}$.

The next source of noise is the dark signal that is produced by the flow of dark current during the exposure and readout time. Dark current exponentially decrease with lower temperature. The dark signal can be easily subtracted from the measured data, however thanks to the random nature of the dark current noise, will add random noise to each measured pixel proportional to $\sqrt{N_{\text{dark}}}$.

Finally, readout noise is mainly caused by limited sensitivity of the readout sensor. It can be reduced if a lower readout speed is used, however at the cost of long total acquisition time.

The expected SNR in one collected pixel is

$$\text{SNR} = \frac{F_p c_e Q_e t}{\sqrt{F_p t Q_e c_e^2 + \sigma_{\text{readout}}^2 + F_{\text{dark}} t}} \quad , \tag{2.55}$$

where F_p is photon flux on a pixel, c_e is number of electrons generated by each photon, Q_e is quantum efficiency of the detector for given wavelength, t is integration time and σ_{readout} is readout noise and F_{dark} is dark current "flux". In contrast to the X-ray range, photon counting cameras are not available in the EUV range. The theoretical ultimate limit is when the photo-electron flux is equal to the dark current flux, however usually the limit is given by stability of the source, maximum acceptable exposure time or dynamic range of the camera.

Systematic errors

In contrast to the random errors, systematic errors are often difficult to detect and even if the source of error is recognized, it is very difficult to remove. Several methods that allow us to deal with some will be shown in Section 2.6. However, especially in the visible light range when the achievable SNR is not limiting, the systematic errors will usually dominate.

In the case when systematic errors cannot be removed in experiment, they have to be correctly included into the diffraction model otherwise they will cause artefacts in the reconstructions.

Common sources of systematic errors are listed here:

- **reflections** - one of the common errors, in the case of visible light, are back-reflections from different pieces of optics or from the CCD chip itself to the sample and back. Although the anti-reflection coating removes most of the reflected light, still if a diffraction pattern with dynamic range of 60 dB is collected, reflections will be present. Reflections are not issue in EUV and X-ray range.
- **background** - background can be separated into several groups:
  - Readout offset and fixed pattern noise - easy to measure and subtract from data;
  - Background from external sources - it either needs to be suppressed mechanically in the experimental setup or it can be measured and subtracted similarly to the dark current;
  - Incoherent background from the source. This can be caused by incoherent scattered X-ray photons [50] or scatter from dust and scratches on the optics behind the sample.
- Insufficient coherence, beam movements, sample vibrations
- Errors in the sample positions, when the ptychography method is used (Sec. 2.4)
- Wrong geometry estimation e.g. sample to camera distance

- Sample is not sufficiently thinner than the depth of focus for given experiment
- The collected diffraction pattern does not have sufficient linearity for example due to slow shutter used to collect multiple exposures for a single high dynamic range acquisition

Missing data

In some cases, it is not possible to measure certain part of the diffraction pattern. Often it is the centre, because the intensity is too high and the collected diffraction pattern is over-saturated even for the shortest achievable exposure time. Another common case is that the sensor consists of several individual areas with a blind space between them or a static beam block is placed in front of the detector in order to avoid damage of the detector or leaking of the central over-saturated pixels into the neighbouring pixels.

Because the missing regions always have negative effects on the reconstruction reliability and convergence, it is preferred to limit the size of these regions. If it is not technically possible, the CDI algorithms can include the missing regions as data with infinite errors using the relaxed modulus constraint (Section 2.4.1). If the missing regions are caused by over-saturation, the additional knowledge that can be used is that the reconstructed missing data has to have higher amplitude than the maximum of the other measured pixels.

The missing data results in unconstrained modes as shown in Ref. [55]. If the ptychography method (Sec. 2.4) is used, the additional information provided by the overlap can significantly help to constrain the missing data especially if the support of the missing region is small or different in every collected pattern.

2.2.5 Method for estimation of the reconstruction quality

The reconstruction quality is easier to be estimated is a simple test sample is used because then many options to estimate the reconstruction quality are available. If an SEM image is available, it could be possible to compare it with the reconstruction and calculate the mutual transfer function as it was shown in Ref. [29].

However, this is not suitable for EUV/visible light samples due to the additional phase effects.

Therefore, the self-consistent methods are preferred. The most common methods are:

1. **Knife-edge method** – a sharp edge in the image is fitted by a sigmoid function and distance between 10%-90% intensity is used to estimate the resolution [29, 56, 57]. The main advantage is robustness, but the disadvantage is requirement of sharp features. Moreover, the knife-edge method has only limited correspondence with the overall reconstruction quality. This criterion provides only local information and therefore it can be easily abused to claim a high resolution based on a single sharp feature.

2. **Phase retrieval transfer function - PRTF** – A relatively standard method that was widely used in single shot CDI [12, 13], but it was also used in ptychography [29]. However, for ptychography it can be used only if the object diffraction is dominating the measured diffraction pattern otherwise the resolution for weakly scattering objects is overestimated [58]. The PRTF for ptychography can be defined as

$$\mathrm{PRTF}(\mathbf{k}) = \left\langle \frac{|\left\langle \Psi_{ij} \right\rangle_i|}{\left\langle |\Psi_{ij}| \right\rangle_i} \right\rangle_j \quad , \tag{2.56}$$

   where Ψ_{ij} denotes i-th reconstruction of exit-wave field (EWF) in the Fourier plane of the j-th position. PRTF shows how reliably each spatial frequency of the EWF was reconstructed, however it does not prove reliability of the reconstructed object and probe. PRTF can be used for resolution estimation, although there is no agreement on the proper threshold. A conservative criterion is when the PRTF drops below 0.5. Another options are $1/e$ or $1/\sqrt{N}$, where N is the number of the averaged reconstructions.

3. **Fourier ring correlation - FRC** – A general method to compare similarity of two complex images at each spatial frequency [59]. The FRC is defined as

$$\mathrm{FRC}(r_i) = \frac{\sum_{\mathbf{k}\in r_i} \left(\mathfrak{F}O_1\right)^* \mathfrak{F}O_2}{\sum_{\mathbf{k}\in r_i} |\mathfrak{F}O_1|^2 \sum_{\mathbf{k}\in r_i} |\mathfrak{F}O_2|^2} \quad , \tag{2.57}$$

   where $\mathfrak{F}O_i$ denotes the Fourier transformation of the i-th image. The FRC values are averaged over all spatial frequencies $\mathbf{k}$ within the ring r_i. The FRC

can be used to estimate the resolution using the 1-bit or 1/2-bit criterion [60]. For comparison of two independent datasets, we have used the 1-bit criterion that is defined as

$$T_{\text{1-bit},i} = \frac{0.5 + 2.41/\sqrt{N_i}}{1.5 + 1.41/\sqrt{N_i}} \quad , \tag{2.58}$$

where N_i is number of pixels in the i-th ring. The drawback of the FRC method is requirement for two independent datasets.

4. χ^2 **- goodness of fit** – Measure of goodness of fit normalized by degrees of freedom (DoFr)

$$\chi^2(k)/\text{DoFr} \approx \frac{1}{N_{\text{pos}} N_{\text{pix}}(k)} \sum_k \sum_i \left(\frac{|\Psi_i(k)|^2 - I_i(k)}{\sigma_i(k)} \right)^2$$

at different spatial frequencies k and $N_{\text{pix}}(k)$ is number of pixels in the k frequency ring. This criterion assumes Gaussian noise in the data. This criterion cannot be used to estimate resolution and also estimation of the real number of degress of freedom for ptychography is nontrivial.

2.3 Classical CDI methods

If it would be possible to directly measure both amplitude and phase of the scattered signal on the detector, the exit wave field at the sample position could be simply recovered using the inverse Fourier transformation. However, because only intensities of the diffraction pattern can be collected, the phase is lost and needs to be recovered numerically. This is known as the "2D phase retrieval problem". This is different from the traditional lens-based microscopy when the lenses are using full information to create a real image on the detector, where the phase is lost again. The idea of CDI was first demonstrated at visible light wavelengths [61] and then at X-ray wavelengths [1]. Since then, many different variations of the CDI were proposed such as Fresnel CDI [62], Keyhole CDI [63], ptychography method [64] or Fourier ptychography [2]. Particularly, ptychography is a very powerful method that allows us to relax coherence requirements and obtain a robust reconstruction.

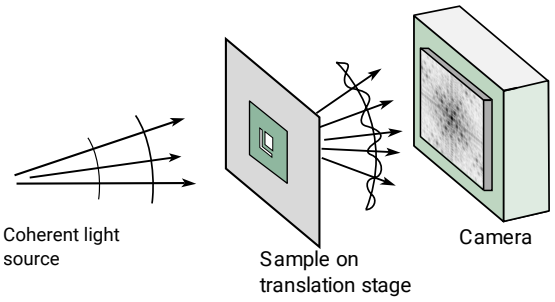

FIGURE 2.13: A typical setup for the CDI with support constraint. The whole sample is illuminated with coherent light and the diffraction pattern is collected by a pixel detector. The image is reconstructed computationally.

Coherent diffractive imaging (CDI) methods are generally based on searching of a solution that satisfies two sets of constraints [65, 66]. One set is applied in the real space (object plane) and the second in the Fourier space (detector plane) as shown in Fig. 2.14. Usually an iterative algorithm is used to find the lost phase information on the detector that satisfies all the constraints. Other phase retrieval methods (e.g. Fourier holography) [67] may allow the phase recovery directly without need of multiple iterations, however these methods have severe limitations on the sample or reference wave preparation, required coherence and limited field of view, in order to fulfil oversampling requirements on the detector plane.

The first successfully used CDI method in short wavelength microscopy was CDI with support and positivity constraints [1]. It is also sometimes called classical CDI

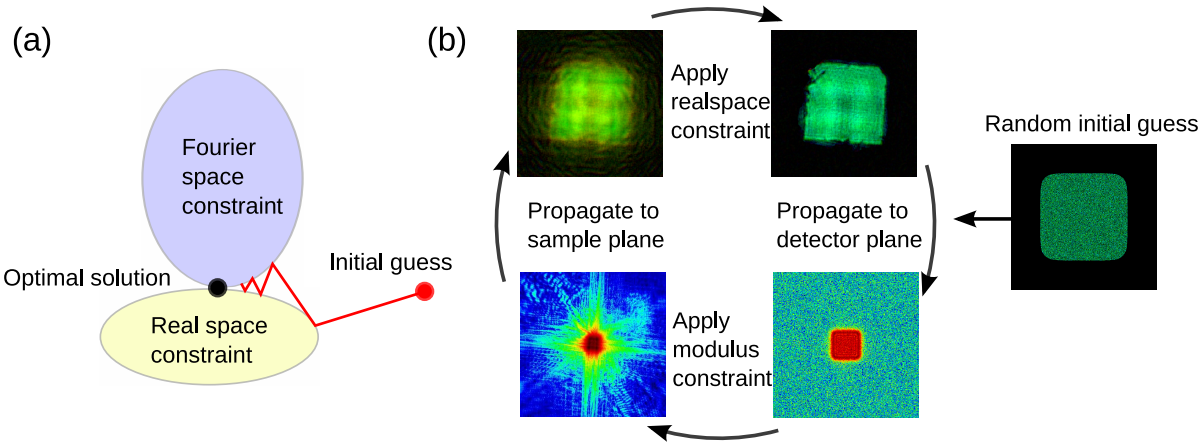

FIGURE 2.14: An illustration of the phase retrieval process (a) A schema of a simple projection based algorithm. The solution is obtained by iterative projections into a given set of constraints. This schema shows the optimal case, when there is only one unique solution, however in real applications the two sets can overlap in multiple solutions or even more commonly, there is no overlap and thus no solution satisfies exactly all constraints. (b) Reconstruction process for the classical CDI with limited support. The initial guess is propagated to the detector, where the modulus of the model is forced to correspond to square root of the measured data. Then the model is back-propagated to the sample plane, where the real-space (e.g. support) constraint is enforced. This procedure is repeated until convergence.

or single diffraction pattern CDI because only one diffraction pattern is needed to achieve reconstruction. The methods are still used in specific applications, where it is not possible to get multiple exposures of the sample. However, the main focus of this thesis is on the more robust scanning methods, therefore only a brief introduction is provided.

2.3.1 Oversampling based CDI

The CDI with limited support is based on three main assumptions:

1. illumination wave-front is precisely known
2. the object is surrounded by empty space
3. measured diffraction pattern is sufficiently oversampled

It is also expected that precise parameters for the light propagation from the sample to detector are well known. The support constraint is also called the

oversampling constraint and it guarantees that according to the Nyquist-Shannon theorem no information is lost if the signal is sampled at twice the frequency of the highest frequency in the original signal. The support constraint in the real space in combination with the *modulus constraint* that ensures that modulus of the diffraction pattern corresponds to the measured data are iteratively applied as shown in the Fig. 2.14(b).

If the illumination wavefront is sufficiently flat and sample to detector distance is large, then this configuration is called far-field CDI and propagation is given directly by the Fraunhofer equation (Eq. 2.18). If the wavefront is curved or the sample size is not negligible compared to the detector size, then it is called Fresnel CDI [62], Keyhole coherent diffractive imaging [63]. The main advantage of the far-field CDI is lower sensitivity to illumination beam position instabilities. On the other hand, the reconstruction is not unique. This means that there may exist a set of different solutions that satisfy both constraints. In case of far-field CDI it is position shift of the reconstruction and so called twinning when an object and a twin object created by flipping over the centre have the same diffraction pattern. Fresnel CDI provides an unique solution, however it needs precise knowledge of the illumination wavefront and the propagation distance. Also the curved illumination wavefront makes the imaging more sensitive to beam position instabilities. Another advantage of the Fresnel CDI is the reduced dynamic range of the collected diffraction pattern compared to the far-field CDI.

The optimal support needs to be either *a priori* known or in the case of sample with high contrast boundary, it is possible to use so called Shrink-wrap algorithm [68]. The Shrink-wrap algorithm tries to find the most compact support around the sample during the iterative convergence and therefore only a coarse initial guess is needed.

The most common algorithms to solve the single diffraction CDI are hybrid input-output (HIO) [65], difference maps (DM) [12], or relaxed averaged alternating reflections (RAAR) [69].

Another option is to use a method based on double blind Fourier holography (DBFH) that uses two different but unknown objects that are sufficiently separated [70, 71]. DBFH avoids the twinning problem and allows us to search for the solution with a convex optimization algorithm that has convergence to the global optimum guaranteed.

The drawback of the single exposure CDI methods it that the maximum imaged region is limited by the coherence length of the illumination. The sample must be roughly as small as the lateral coherence length [48] for CDI and this requirement is even more strict for holography.

2.4 Scanning CDI – ptychography

The standard CDI method with the Shrink-wrap algorithm can be applied only on compact, isolated objects with sufficient oversampling. Another method called ptychography [64] uses information from multiple independent but overlapping measurements. Therefore, ptychography is not restricted only to compact objects and the maximum imaged region is theoretically unlimited. Moreover, the overlap constraint in real space is a stronger and more versatile constraint than oversampling, therefore the reconstruction is more robust and additional parameters can be recovered [16, 49, 72]. With a sufficient overlap, it is possible to achieve multicolour reconstruction [52], reconstruct multiple layers of the sample [25–27] or reconstruct higher spatial frequencies than the maximal frequency given by the collected numerical aperture (i.e. super-resolution) [45].

The basic idea of standard ptychography is to scan a specimen with a finite extent structured illumination probe and collect a diffraction pattern for each scanning position (see Fig. 2.15). In order to benefit from the redundancy, the scan step must be sufficiently smaller than the illumination probe size on the sample and the probe size should be sufficiently smaller than the scanned region. This idea was first proposed by W. Hoppe in 1969 [73]. The first attempts to solve the 2D ptychography problem were based on the Wigner distribution deconvolution [74]. In X-ray range, the Wigner distribution deconvolution method was first demonstrated by H. Chapman [75]. The first iterative solver based ptychography simulations were published by J. Rodenburg and H. Faulkner in 2004 [76, 77]. Later, the iterative ptychography was demonstrated in the visible [78] and the X-ray range [79].

However, ptychography started to be a really useful technique when more efficient numerical solvers were developed [80–82]. Mainly, the possibility to recover the illumination probe along with the object significantly improved the imaging quality.

Standard ptychography is based on the projection approximation, when it is assumed that the exit wave field ψ_j at j-th position can be factorized into a part that is related to the illumination probe and part that is related to the object shifted to position j

$$\psi_j(\mathbf{x}) = O(\mathbf{x})P(\mathbf{x} - \mathbf{x}_j) \quad , \tag{2.59}$$

where P and O denotes probe and object functions and $\mathbf{x}_j$ is a relative shift to the j-th scanning position.

This relatively general definition of ptychography provides many ways to arrange a ptychography setup. The most common is an illumination forming aperture or lens in front of the sample as shown in Fig. 2.15 (e.g. [79–82]). Another option is placing the aperture downstream of the sample (e.g. [77, 83]) and reconstruct near-field propagated object illuminated by a plane wave. This configuration needs to back-propagate the reconstructed object back and the sample plane and it assumes close to flat illumination wavefront. Another very powerful method called Fourier ptychography [2] crops the spatial frequency spectrum by an aperture in the Fourier plane and the detector that is placed in the imaging plane observes a magnified image of the Fourier filtered image.

The standard ptychography method needs to fulfil the following requirements:

a) Stable illumination, i.e. no intensity or position distribution fluctuations over the scan time and also the distance from the sample to the illumination forming optics needs to be unchanged

b) Precise knowledge of propagation distance between the sample and the camera

c) Precise the knowledge of sample scanning positions

d) Monochromatic light source, temporal coherence length $l_t > X_a$.NA, where X_a is aperture size and NA numerical aperture

e) Lateral coherence length twice as large as the aperture size

f) Sufficient oversampling of the speckle pattern ideally at higher than the Nyquist frequency.

g) Sample is thin, i.e. it can be approximated by a single diffraction interface

h) Sample is not changing during the scan

Most of these requirements are usually easily fulfilled with visible light coherent sources or synchrotron experiments in the hard X-ray range. However, laboratory short wavelength sources provide significantly lower illumination quality and therefore the stability and sufficient coherent flux are severe issues. Multiple methods that allow relaxation of some of the ptychography constraints are described

in this chapter. However, a drawback of the relaxation methods is usually slower convergence, higher computational demands or requirement of larger overlap in the scan.

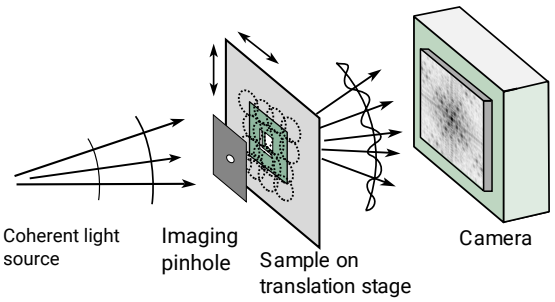

FIGURE 2.15: A simplified schematic of a typical setup for the ptychographic measurement. In contrast to the single exposure CDI (Fig. 2.13), only a fraction of the sample is illuminated and the sample is moved so that the illumination probe is scanning over the entire sample. Each diffraction pattern is collected by a pixel detector and the full image is reconstructed computationally.

2.4.1 Projection based iterative algorithms

The majority of the algorithms used for ptychography are projection based [81, 82, 84] and, similarly to the previous CDI method, the solution is searched as an intersection between two constraints sets [85].

The sets of constraints can vary dependent on the available prior knowledge; however the most common are modulus constraint, overlap constraint, support constraint, positivity constraint or smoothness. The constraints are expressed in the form of projectors, where a projector is defined as a map that for each point x returns point y that satisfies the given constraint so that the distance $\|x - y\|$ is minimized. Often the standard projection cannot be used because the problem to be solved does not obey the projection precisely. In that case, it is advantageous to use the relaxed projector $\mathcal{P}_\alpha$ defined as [85]

$$\mathcal{P}_\alpha(x) = x + \alpha(\mathcal{P}(x) - x) \quad , \tag{2.60}$$

where α is a scalar number that is changing the force of the relaxed constraint. If $\alpha \in (0, 1)$ then the projector is called under-relaxed and over-relaxed for $\alpha > 1$. Some of the projection based algorithms such as difference maps [82] or relaxed averaged alternating reflections [84, 86] are based on reflectors. Reflector is called a relaxed projector in the case of $\alpha = 2$. In the similar way, we can define relaxed

reflector for $\alpha \in (1, 2)$. An example of the projection and reflection is shown in Fig. 2.16.

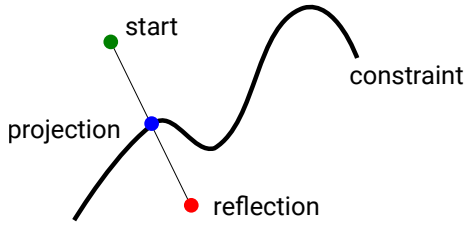

FIGURE 2.16: The green initial point is projected on the constraint (black line) resulting in the blue point. Reflection of the green point over the constraint is denoted by the red point.

Modulus constraint

The modulus constraint P_m is the fundamental common constraint and usually the only one that is applied in the Fourier space (detector plane). If Ψ denotes the exit-wave ψ propagated to the detector plane using the optimal propagator, then the modulus projector can be expressed as

$$P_m\Psi = \sqrt{I}\exp(\mathrm{i}\arg(\Psi)) = \frac{\sqrt{I}}{|\Psi|}\Psi \quad , \tag{2.61}$$

where $\sqrt{I}$ is modulus of the measured data. The modulus constraint simply replaces the amplitude of the field Ψ by the known measured values $\sqrt{I}$ while the phase is conserved. The factor $\frac{\sqrt{I}}{|\Psi|}$ can be also considered as a correction factor, where only the spatial frequencies with ratio $\sqrt{I}/|\Psi|$ unequal to one are changed and the rest is untouched. This idea is used for different Fourier space corrections as shown in Section 2.5.

All real measurements are noisy and thus the model can never be closer to the data then the noise level σ. An intuitive approach to include this into the modulus constraint is to correct only the values with distance $\|I - |\Psi|^2\| > \sigma$. The corrected value needs to be replaced either by $\sqrt{I+\sigma}$ or $\sqrt{(I-\sigma)_+}$, where $()_+$ denotes positive part, dependent on which is closer to the value $|\Psi|$ as it was shown in Refs. [13, 87]. This idea can be written as

$$P_m\Psi = \frac{\sqrt{I} + \tilde{\sigma}\mathrm{sgm}\left(\frac{|\Psi|-\sqrt{I}}{\tilde{\sigma}}\right)}{|\Psi|}\Psi \quad , \tag{2.62}$$

where $\tilde{\sigma}$ denotes the noise level σ transformed according to distribution of $\sqrt{I}$ and "sgm" denotes the sigmoid function $\mathrm{sgm}(x) = x/(1+|x|)$. The previous equation can be understood as a form of a relaxed modulus constraint with a weighting matrix W

$$P_m\Psi = \left[(1-W) + W\frac{\sqrt{I}}{|\Psi|}\right]\Psi \quad , \tag{2.63}$$

where the weighting factor W is defined by the following formula

$$W = 1 - \left[1 + \frac{|\sqrt{I} - |\Psi||}{\tilde{\sigma}}\right]^{-1} \quad , \tag{2.64}$$

The relaxed modulus constraint (Eq. 2.63) leaves almost unchanged values of Ψ, where the exit-wavefront model $|\Psi|$ is closer to the measured data $\sqrt{I}$ than expected error σ. Moreover, equation (2.62) shows that for large errors, the modulus values are replaces by $\sqrt{I+\sigma}$ or $\sqrt{(I-\sigma)}$. Lower weight W results in slower convergence speed and more relaxation in the noisy regions. The weighting matrix W can be also easily used to include missing data by setting their relaxation weight to 0. Note that if only the Poisson noise is considered and the number of captured photons is sufficiently high, $\tilde{\sigma}$ is equal to 0.5 independent of the value $\sqrt{I}$. This is the reason why the simple modulus constraint (Eq. 2.61) works usually very well for short wavelength CDI.

If the measured data are limited only by the Poisson noise and other sources of errors can be neglected, it is possible to used more optimal solvers based on the maximal likelihood method [46].

Another ways of statistical regularization by effectively rescaling the low intensity values of the measured data are described in Ref. [88], however the final effect is very similar to background removal procedure (Section 2.6.2).

Overlap constraint

The overlap constraint is a crucial part of the ptychography algorithm. The idea that illumination passing through adjacent regions of a sample must carry shared information is very general. The overlap can be generally expressed as

$$\psi_j(\mathbf{x}) = P(\mathbf{x})O(\mathbf{x}+\mathbf{x}_j) \quad , \tag{2.65}$$

where $\psi_j(\mathbf{x})$ is the exit wave from j-th scanning position $\mathbf{x}_j$ with illumination probe P and sample O. The overlap constraint has the following assumptions:

1. The illumination is expected to be constant for all scanning positions. This is usually fulfilled for visible light experiments and synchrotron experiments, however high harmonic generation sources suffer from limited pointing stability, which can deteriorate reconstruction, as shown in Section 3.2.

2. Sample must be thin and flat. Thick samples cannot be decomposed into the form of Eq. 2.65. Thick and sufficiently transparent samples can be reconstructed using the multi-slice ptychography extension [25] that uses a modified overlap constraint. The effect of a sample that is tilted with respect to the optical axis is discussed in Section 2.1.5.

3. Sample must be isotropic. If the sample is anisotropic and illumination is not plane wave, then different parts of the illumination probe sees slightly different material properties of the sample. Anisotropy can be also included into ptychography as shown in Section 2.6.2.

4. Scanning positions need to be known precisely. If there is any imprecision of the probe positions, then the overlap constraint cannot be valid. The positioning system should have ideally precision higher then the reconstructed pixel size, because the total position error may add up over longer scans and it may cause non-negligible drifts of the scanning positions. Finally, it is necessary to know precisely an estimate of the sample to camera distance because it defines the reconstructed pixel size (i.e. scaling of all the positions). Wrong distance estimation is equivalent to wrong shift of all probe positions.

   Scaling and probe positions can be corrected using multiple correction methods during reconstruction [89–91] or using our gradient descent method shown in Section 2.6.2.

The optimal amount of overlap depends on the data quality and sample structure, however, for our experimental data, the optimal linear overlap is between 60-85%.

The optimal overlap projector P_π can be derived from the least-squares method as a solution that best describes all diffraction patterns in each iteration

$$\hat{O}(\mathbf{x}) = \frac{\sum_j P^*(\mathbf{x}-\mathbf{x}_j)\hat{\psi}_j(\mathbf{x}-\mathbf{x}_j)}{\sum_j |P(\mathbf{x}-\mathbf{x}_j)|^2 + \varepsilon} \quad \hat{P}(\mathbf{x}) = \frac{\sum_j O^*(\mathbf{x}+\mathbf{x}_j)\hat{\psi}_j(\mathbf{x})}{\sum_j |O(\mathbf{x}+\mathbf{x}_j)|^2 + \varepsilon} \quad , \tag{2.66}$$

where $\hat{O}$ and $\hat{P}$ are the updated estimations of the object and the probe, * denotes complex conjugate, ε is a small number to avoid dividing by zero and $\hat{\psi}_j$ are updated estimates of the exit wave field $\hat{\psi}_j = P_m(\psi_j)$. The advantage of expression 2.66 is that all updates $\hat{\psi}_j$ can be solved in parallel and thus the calculation speed can be significantly improved. This parallel update scheme is used for example difference maps (DM) [80, 82] or relaxed averaged alternating reflections (RAAR) ptychography algorithms [84, 86].

If the parallel scheme is used in a simple error (ER) reduction scheme (Fig. 2.14)

$$\psi_j^{(i+1)} = P_\pi \left[P_m \left(\psi_j^{(i)} \right) \right] \tag{2.67}$$

this "optimal solution" often converges into the closest local minimum. Another possible solution, proposed by Maiden *et al.* [81] for extended ptychographic iterative machine (ePIE), applies the overlap constraint serially

$$\begin{aligned} \hat{O}(\mathbf{x}) &= O(\mathbf{x}) + \alpha_O \frac{P^*(\mathbf{x}-\mathbf{x}_j)}{\|P\|^2_{\max}} \Delta\psi_j(\mathbf{x}) \\ \hat{P}(\mathbf{x}) &= P(\mathbf{x}) + \alpha_P \frac{O^*(\mathbf{x}+\mathbf{x}_j)}{\|O\|^2_{\max}} \Delta\psi_j(\mathbf{x}) \quad , \end{aligned} \tag{2.68}$$

where $\alpha_P, \alpha_O \in (0, 1]$ are relaxation parameters, $|\cdot|_{\max}$ denotes the maximum norm and $\Delta\psi_j(\mathbf{x}) = \hat{\psi}_j(\mathbf{x}) - \psi_j(\mathbf{x})$.

In order to avoid trapping in a local minimum, the ePIE method applies the relaxed ER scheme on each scanning position in a random order

$$\psi_{\pi(j)}^{(i+1)} = P_\pi \left[P_m \left(\psi_{\pi(j)}^{(i)} \right) \right] \tag{2.69}$$

where $\pi(j)$ denotes a random permutation over the indices j. Finally, in Sec. 2.5.2 we will show a better way how to find a more optimal relaxation parameter α that can be generally much larger than 1.

Support constraint

The support constraint is crucial for single diffraction CDI, however in specific cases it can be also used in the ptychography method if the illumination probe has a limited support. The projection P_s into the support $\mathcal{S}$ is

$$P_s\psi(\mathbf{x}) = \begin{cases} \psi(\mathbf{x}) & \text{if } x \in \mathcal{S} \\ 0 & \text{if } x \notin \mathcal{S} \end{cases} \tag{2.70}$$

Often it is advantageous to use a relaxed version of the projector with $\alpha \ll 1$. In the case of transmission ptychography setup with illumination defined by a pinhole, the support constraint has to be first back-propagated on the pinhole plane (see Section 5.5.1). The support $\mathcal{S}$ can be very loose, therefore the dynamic support estimation is generally a simple thresholding using a similar approach to the Shrink-wrap algorithm [68]:

$$\mathcal{S} = |P| * k > T \quad , \tag{2.71}$$

where $*$ denotes convolution with a positive kernel k thresholded by a value T.

Positivity constraint

The positivity is very strong constraint because it significantly limits the space of possible solutions as it was shown for example in Refs. [65, 66]:

$$P_+\psi(\mathbf{x}) = |\psi(\mathbf{x})| \quad . \tag{2.72}$$

This strict version of the positivity constraint is only rarely valid. As we will show in Sec. 5.6, it is not even valid for binary samples if they are not sufficiently thin. However, the relaxed positivity constraint with $\alpha \ll 1$ can be used for majority of the EUV samples because the phase shift is usually relatively small compared to attenuation. Moreover, relaxed positivity constrain can be used as a regularization if additional degrees of freedom are added (see e.g. Section 2.7)

Smoothness constraint

Using a smoothness constraint in high resolution microscopy experiments seems to be illogical, but despite the overlap constraint, ptychography can be still an ill-posed method. A simple smoothness constraint can be generally defined as

$$\hat{O}(\mathbf{x}) = O(\mathbf{x}) * \gamma \quad , \tag{2.73}$$

where $*$ denotes convolution operator and γ is a small positive kernel. Note that relaxation of this "constraint" is not needed because it can be simply adjusted by the kernel size.

The smoothness constraint applied in the object real space is advantageous during initial convergence of reconstruction with highly diffusive probe (see Fig. 3.2). The smoothness constraint then forces first the sharp features into the probe. This speeds up the convergence and helps to the position correction methods.

Scanning positions

In order to use the standard ptychographic reconstruction, the scanning positions need to be chosen so that the illuminated regions are overlapping. The ideal linear overlap is around 60-90% [92]. The minimal empirically recommended number of scanning positions for a simple high contrast sample is 4×4 with ≈70% linear overlap. However, larger scans than 10×10 should be used in the case of a lower quality data or weakly scattering samples.

Further, it is advantageous to keep the distance between the scans aperiodic since a periodic grid can result in a periodic artefacts in the reconstruction [82]. This is mainly an issue if the measured diffraction pattern has a positive offset (imperfectly removed background) that is not correctly included in the reconstruction method as shown in Ref. [20]. It is possible to use either some irregular patterns such as a spirals [93] or simply a regular rectangular grid with random offset. The advantage of the spiral shape is better overlap between temporally distant regions that allows us to see more clearly long term intensity fluctuations and illumination probe drifts.

For practical reasons, the maximum size of the scans is limited by operation memory of the GPU card used (Nvidia GeForce GTX Titan X). In our case, the memory

limitation is 12 GB. This corresponds to 22000 diffraction patterns stored as unsigned integer 16-bit at resolution 512×512 pixels using the ePIE method. Due to other hardware limitations, the maximal CUDA matrix size is $512 \times 512 \times 8192$ elements, however this is only a technical problem that can be avoided. Larger scan can be reconstructed using parallelization over multiple cards, computer clusters or using CPU and RAM. A disadvantage of CPU at the similar price range is $\approx 20\times$ lower performance if the BLAS/FFTW libraries are used compared to CUDA (see Fig. 2.18).

2.4.2 Reconstruction process

The reconstruction process usually consists of the following steps:

1. Get initial guess of the illumination probe support and sample to camera distance (i.e. the numerical aperture)
2. Object is initialized as a random pattern and probe as a shape of the expected support
3. Apply the modulus constraint P_m on the current estimation of the exit wave ψ_j for the current probe position
4. Apply the overlap constraint
5. Apply additional constraints if possible
6. Repeat steps 3-5 until convergence

Stability of the ptychography algorithm depends on the amount of independent information provided by each probe shift. If the probe support size is large compared to the maximum scanning range or the probe is very smooth, the algorithm will have similar behaviour as in the case of CDI and further *a priori* knowledge may be needed for a good reconstruction.

Furthermore, the object and probe patterns should not be sparse in the Fourier domain. The ptychographic projection propagated into the Fourier plane

$$\Psi = \mathfrak{F}\left[P(\mathbf{x})O(\mathbf{x}+\mathbf{x}_j)\right]$$

can be equivalently rewritten as a convolution of the probe and the object diffraction

$$\Psi = \mathfrak{F}\left[P(\mathbf{x})\right] * \mathfrak{F}\left[O(\mathbf{x}+\mathbf{x}_j)\right] \quad . \tag{2.74}$$

Figure 2.17a) shows the diffraction intensity resulting from convolution of a circular illumination probe and a periodic grid. If the whole sample is sparse in the Fourier

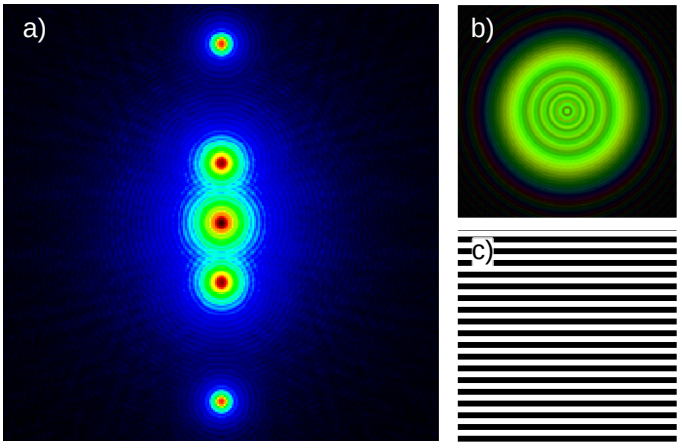

FIGURE 2.17: (a) An example of a sparse diffraction pattern intensity in logarithmic colour-scale. The pattern was created by multiplicative projection of near-field propagated illumination through a circular aperture (b) and a periodic grid sample (c)

domain, then most of the Fourier space cannot be explored by the ptychography scan and it is not possible to reliably reconstruct the probe 2.17(b). A similar issue is raised if the sample contrast is very low compared to the probe pattern, then the dominating diffraction pattern belongs to the probe and the object contribution may not be sufficient to over-constrain the ptychographic reconstruction [58].

In these degenerate cases, it is recommended to use a high contrast aperiodic sample to obtain an initial guess of the ptychographic probe and other ptychography parameters.

2.4.3 Standard reconstruction algorithms

Despite many small modifications, the most common ptychography reconstruction algorithms can be divided into two groups: projection based and reflector based. The most common representative of the projection based group is the extended ptychographic iterative machine (ePIE) [94] that is based on application of the modulus and real-space constraint serially on each scanning position in random

order. As it was already mentioned, this allows to more easily avoid trapping in local minima. The real-space constraint is given by the Eq. 2.68 that is averaging the corrections given by the modulus constraint into a single illumination probe and single object. Other common algorithms in this group are the PIE method [76], the maximum likelihood method [46] or the translation diversity algorithm [95].

The reflectors based group of algorithms is most commonly represented by the difference maps (DM) ptychography algorithm [82]. The DM method was proposed by V. Elser [96, 97]. DM search the solution in parallel for all positions, which makes this methods suitable for parallelization and solving large scale reconstructions. The real space constraint is given by the Eq. 2.66. The $n+1$ iteration is searched by the difference maps using the following formula

$$\boldsymbol{\psi}^{(n+1)} = \boldsymbol{\psi}^{(n)} + P_m[2P_\pi(\boldsymbol{\psi}^{(n)}) - \boldsymbol{\psi}^{(n)}] - P_\pi(\boldsymbol{\psi}^{(n)}) \quad , \tag{2.75}$$

where P_m denotes the modulus constraint and P_π is the overlap constraint given by Eq. 2.65

$$P_\pi(\psi_i) \rightarrow \hat{P}(\mathbf{x})\hat{O}(\mathbf{x}+\mathbf{x}_i) \quad , \tag{2.76}$$

where $\hat{P}, \hat{O}$ are given by Eq. 2.66. DM scan the space around the vector ψ_n. Instead of going in the direction of the projection, DM rather "overshoot" the constraint and go in direction of the reflector. This helps to avoid trapping in local minima, but also it can prevent convergence if any systematic errors (e.g. noise, incoherence, etc) are present. Therefore, it is preferable to use relaxed projectors. The final solution is given by the values $\hat{P}, \hat{O}$. One drawback of the DM is larger memory requirement compared to ePIE, because the vector ψ_n needs to be stored in memory. It is twice the size of the dataset because ψ_n is complex, therefore the memory requirements are at least three times higher than for ePIE.

The second method in this group, which is however less common, is the relaxed averaged alternating reflections (RAAR) method [69, 86].

2.4.4 Optimization of the reconstruction speed

One of the common issues of the CDI methods is the reconstruction time. On contrary to the lens-based methods, the imaging only starts when the data are collected. Fast reconstruction speed is crucial for visible light and X-ray synchrotron experiments, where the flux is not limiting the data acquisition time and thus it is

possible to generate datasets with size of several GB in several minutes minutes. Even larger amounts of data can be produced by the modern X-ray detectors used on synchrotrons that have negligible readout time and the main limitation in data collection is the setting time of the piezo actuators [98–100]. Large GPU clusters are used in order to deal with the amounts of data produced [101]. However, this is infeasible for small laboratory scale experiments.

Our laboratory scale CDI experiments include a large variety of problems with significantly varying data quality, dataset sizes and convergence issues. Sometimes, fast convergence is preferred and in other experiments high robustness is needed. But in any case, fast iteration speed is crucial. In order to process the large amount of data, we have used single GPU card: Nvidia GTX Titan X 12 GB. GPU CUDA implementation can provide significant speed-up compared to a CPU implementation based on the BLAS and FFTW libraries, however only for large matrices. Figure 2.18 shows the dependence of a single FFT calculation time on the dataset size performed by Matlab parallel processing toolbox. The maximum speed-up is up to 50× for large matrices. However, it also shows that GPU performance is less efficient than CPU for smaller sizes of matrices. That is not a problem for the difference maps, parallel PIE or conjugate gradient methods, where all positions are solved in parallel. However, it is a significant drawback for the ePIE method that is in general the most common method for small scale experiments thanks to its robustness and flexibility.

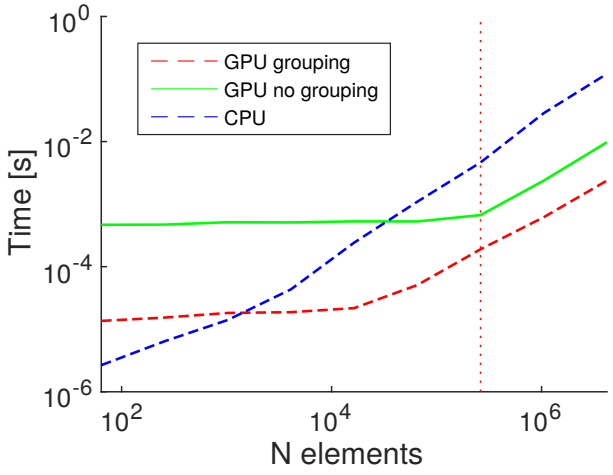

FIGURE 2.18: Tests of the calculation speed of a CPU (Intel Core i7 4970, 4 cores) compared to GPU (Nvidia Titan X). The tests were based on calculating the near-field propagation ($\mathfrak{F}^{-1}[H\mathfrak{F}(x)]$). The CPU code was using FFTW algorithm and the GPU code was based on CudaFFT. The red line denotes the case when the calculation of 20 positions was performed in parallel. The dotted vertical line denotes resolution 512×512 pixels.

2.5 Improved iterative ptychography solvers

2.5.1 Sequential gradient descent solver based on Levenberg-Marquardt method

Many methods have been already proposed to improve the convergence speed and stability of ptychography [102–104]. Here, we will summarize the state of the art methods and based on them, we will introduce our improved ptychography solver. Most of the ptychography iterative solvers are based on adjustments of the overlap constraint Eq. (2.68). A simple way to understand the ePIE-like update (i.e. the overlap constrain Eq. 2.68) is as a relaxed projector Eq. 2.60 in the following form

$$O_{n+1}(\mathbf{x}+\mathbf{x}_j) = (1-\alpha(\mathbf{x}))O_n(\mathbf{x}+\mathbf{x}_j) + \alpha(\mathbf{x})\frac{P_n^*(\mathbf{x})\psi_j(\mathbf{x})}{|P_n(\mathbf{x})|^2} \quad . \tag{2.77}$$

The vector of the relaxation parameters $\alpha(\mathbf{x})$ is defined for the ePIE method as $\alpha(\mathbf{x}) = \hat{\alpha}\frac{|P(\mathbf{x})|^2}{\|P\|^2_{\max}}$ and it reduces the convergence speed for the regions, where intensity of the probe P is smaller then $\|P\|_{\max}$. The main implication of the relaxed constraint is that if the positive relaxation parameter α is less than 1, the error should be monotonically decreasing [65]. This is not true if the constraint is applied sequentially as in the ePIE method because the global error may be increased, however our empirical tests show that selection of $\alpha < 1$ prevents divergence.

Sequential solvers have the advantage of fast convergence and they are not easily trapped in a local minimum. On the other hand, parallel solvers are generally more robust against systematic errors [6, 103] and they can be faster because of easier parallelization between many clusters. If a single most optimal update is selected for whole dataset, Eq. 2.68 can be also understood as a variation of the standard gradient descent method

$$O_{j+1} = O_j - \gamma \nabla f(O) \tag{2.78}$$

with the step length $\gamma = 1/\|2P\|^2_{\max}$. The gradient of the error function $f = (|\Psi| - \sqrt{I})^2$ is equal to [46]

$$\nabla f(O) = -2P_j^*(\mathbf{x})\Delta\psi_j(\mathbf{x}) \quad . \tag{2.79}$$

Using the framework developed for the gradient descent methods, it is possible to optimize convergence. One example, proposed in Ref. [103], is to use the preconditioned gradient descent method

$$O_{j+1} = O_j - \gamma K^{-1} \nabla f(O) \tag{2.80}$$

with the following preconditioner

$$K(\mathbf{x}) = \sum_{i=1}^{N} |P(\mathbf{x} - \mathbf{x}_i)|^2 + \delta \quad , \tag{2.81}$$

where δ is a small number preventing dividing by zero. If the selected step is equal to $\gamma = 1/2$, then this preconditioner leads to the least squares solution shown in Eq. 2.66.

$$O_{j+1} = O_j - \frac{\sum_{i=1}^{N} P_i^* (P_i \cdot O_j - \tilde{\psi})}{\sum_{i=1}^{N} |P_i|^2 + \delta} = \frac{\sum_{i=1}^{N} P_i^* \tilde{\psi}}{\sum_{i=1}^{N} |P_i|^2 + \delta} \quad , \tag{2.82}$$

where P_i denotes $P(\mathbf{x} - \mathbf{x}_i)$ and $\tilde{\psi}$ is the exit wave after application of the modulus constraint.

Instead of the gradient descent method, it is possible to use more advanced methods such as the conjugate gradients (CG) [46, 102].

$$\begin{aligned} O_{j+1} &= O_j + \gamma \Delta_j \\ \Delta_j &= \nabla f(O) + \beta_j \Delta_{j-1} \end{aligned} \quad , \tag{2.83}$$

where Δ_j denotes direction of the j-th step and β is the conjugate projection into the previous direction that is calculated using the Polak-Ribiere formula [105]

$$\beta_j = \frac{\Re[O_j(O_j - O_{j-1})]}{\|O_{j-1}\|^2} \quad . \tag{2.84}$$

The conjugate gradient should improve convergence, however in our tests the value of β_j was for iteration > 10 practically negligible of order of $10^{-2} - 10^{-3}$.

Gradient descent and the conjugate gradients method are both methods of the first order. Second order methods should be able to provide faster convergence. Second order methods such as the Newton method uses both first and also the second derivative of the object function (Hessian matrix) [106], which provides a better approximate of the object function and makes possible to reach faster convergence.

The main drawback is that the Hessian matrix size scales quadratically with the size of the object and calculation of its inverse is computationally expensive. This is the reason why an approximate method such as the Gauss-Newton method is preferable. The Gauss-Newton method replaces the Hessian matrix H by the outer product of two Jacobian matrices J

$$H \approx \left(\frac{\partial f}{\partial O_k}\right)^* \left(\frac{\partial f}{\partial O_l}\right) = J^T J \quad . \tag{2.85}$$

The updated iteration can be expressed as

$$O_{j+1} = O_j - (J^T J)^{-1} J^T \nabla f(O) \quad . \tag{2.86}$$

In order to avoid issues with singularity of the Jacobian, it is preferable to use a damping parameter δ. This results in the Levenberg-Marquardt (LM) algorithm also called damped least-squares method [106]

$$O_{j+1} = O_j - (J^T J + \delta I)^{-1} J^T \nabla f(O) \quad . \tag{2.87}$$

The inversion of the Jacobian matrix is still computationally very costly operation. The inversion can be avoided if the matrix is approximated only by its diagonal form [103]

$$J = Q_l^* \mathrm{diag}(|P|) Q_l \quad ,$$

where Q denotes the displacement matrix shifting the probe P. The resulting LM ptychography update formula will be

$$O_{j+1} = O_j - \frac{\alpha}{\|P\|_{\max}} \frac{\sum_i |P_i|}{\sum_i |P_i|^2 + \delta} \nabla f(O) \quad . \tag{2.88}$$

A method for estimation of the optimal step α is shown in Section 2.5.2. Note that if the probe is a top hat function, the LM update is identical to the ePIE method update. The gradient $\nabla f(O)$ can be estimated to be proportional to $|P|^2$, then the LM update is proportional to $|P|^3/(|P|^2 + \delta)$ while the ePIE update is proportional to $|P|^2$ and the least squares (LSQ) solution (Eq. 2.82) is proportional to $|P|^2/(|P|^2 + \delta)$. The LM update relaxes the low intensity regions of the probe/object more than the LSQ method, and it provides faster convergence than ePIE in the regions close to the maximum probe / object intensity.

It is also possible to solve the Hessian matrix without approximations, as it is

shown for example in Ref. [107], however it requires additional nested solver loop to find the inversion of the Hessian matrix.

There are still other ways to make the sequential solvers such as ePIE faster:

1. Non-overlapping regions can be solved in parallel
2. The optimal relaxation constant α can be larger than 1

As it is shown in Fig. 2.18, it is beneficial to process multiple scanning positions together in small bunches, especially when the calculations are done on GPU. We are using a simple heuristic algorithm to select the least overlapping subsets of the randomly sorted scanning positions. We use a sequential LM update defined

$$O_{j+1,k}(\mathbf{x}) = O_{j,k}(\mathbf{x}) + \frac{\alpha_O}{\|P\|_{\max}} \frac{\sum_j^{\pi(k)} |P_j| P_j^* \Delta\psi_j}{\sum_j^{\pi(k)} |P_j|^2 + \delta} \quad , \tag{2.89}$$

where P_j denotes $P(\mathbf{x} + \mathbf{x}_j)$ and $\pi(k)$ denotes k-th set of N partially overlapping positions. This constraint behaves as the LM constraint Eq. 2.88 in the case of full parallelization and it benefits from random serialization similarly to the ePIE method in the case of lower parallelization.

2.5.2 Gradient descent solver with variable step

The LM constraint works for any bunch size between 1 and N_{pos}, however for larger grouping the convergence speed is reduced. This can be easily seen from Eq. (2.89), where the optimal relaxation constant can be generally anything between $\frac{\alpha}{N_{\text{pos}}}$ to α. We propose to use the variable step size in a similar way as it was introduced in Ref. [46] for the conjugate gradient method. The optimal step size is selected in order to minimize the Fourier space error summed over the scanning positions and detector pixels

$$\chi^2 = \sum_i \sum_{\mathbf{k}} \left[W_i (|\mathfrak{F}\left[(O_i + \alpha\Delta O_i)(P + \alpha\Delta P_i)\right]|^2 - I_i) \right]^2 \quad , \tag{2.90}$$

where ΔP, ΔO denote probe and local object patches updates respectively defined by Eq. 2.68. Matrix W is a weighting factor defined as $1/\sigma$ in the measured regions and 0 elsewhere. $\sigma(\mathbf{k})$ is the noise estimation at position $\mathbf{k}$. Direct minimization of

Eq. (2.90) would be highly inefficient. Fortunately, it is possible to find a quadratic approximation of the Eq. (2.90) that has a fast analytic solution. We will use the following notation:

$$\begin{aligned} A_0 &= \mathfrak{F}(O\Delta P + P\Delta O) \\ A_1 &= |\Psi|^2 - I \\ A_2 &= 2\Re\left(\Psi A_0\right) \\ A_3 &= |A_0|^2 \\ W &= 1/\sigma^2 \quad , \end{aligned} \tag{2.91}$$

where A_0-A_3 denote stack up 2D frames for N scanning positions. Note that contrary to the method proposed in Ref. [46], we use only a single FFT instead of three time number of the incoherent modes, which makes the calculation significantly faster. Moreover, $|\Psi|^2$ denotes the current estimation of the Fourier space model in the Fourier plane and therefore it can be directly replaced by our generalized linear model $\mathbb{A}\sum_i |\Psi_i|^2 + \mathbf{b}$ that is introduced in the Section 2.6.2. The Eq. (2.90) can be approximated as

$$\chi^2 \approx W\left(A_1 + \boldsymbol{\alpha}A_2 + \boldsymbol{\alpha}^2 A_3\right) \approx W\left[A_1^2 + \boldsymbol{\alpha}2A_1A_2 + \boldsymbol{\alpha}^2(2A_1A_3 + A_2^2)\right] \quad . \tag{2.92}$$

This is a quadratic equation in $\boldsymbol{\alpha}$ and the minimum can be found as the root of the first derivative of χ^2

$$\alpha_i = \frac{\sum W A_1 A_2}{\sum W\left(2A_1A_3 + A_2^2\right)} \quad . \tag{2.93}$$

Unlike the maximum likelihood (ML) based conjugate gradient (CG) method presented in [46], we estimate a different optimal α_i for every scanning position, which allows us to automatically select the optimal convergence speed for each region without additional calculation costs. On the other hand, our method cannot garantee improvement of global likelihood in every step as the ML-CG method [46].

The variable step idea can be used for position correction as shown in Section 2.6.2. However, it is still necessary to keep in mind that step size α_i given by equation Eq. 2.90 is the best value to optimize the Fourier space errors regardless of the real space (overlap) constraint. Hence, the optimal α_i may lead to increase of the Fourier error in the next iteration. It can be easily seen that choice of $\alpha_i = 0$ fully satisfies the overlap constraint in the real space and therefore the optimal

value that optimize both constraints lays between these two extremes. We have selected

$$\hat{\alpha}_i = \alpha_i/2 \quad . \tag{2.94}$$

This criterion is relatively conservative, however we have tested that it provides optimal and consistent results for both simulated and measured datasets.

2.5.3 Numerical tests of the iterative ptychography solvers

The gain of the proposed improvements was first tested on simple numerical examples, in order to compare behaviour without systematic errors in data. However because numerical tests can never include all issues of the real experiments, we will show many reconstruction examples from the real data ptychography in the next chapters. The simulation parameters were set to common values in EUV microscopy. The illumination wavelength was 30 nm and real space pixel size 50 nm and 256×256 pixels size detector. Attenuation and phase of the simulated sample were created from an image of a tissue sample (Fig. 2.19). The modulus was normalized between 0.7 and 1, and phase was from -0.2 to 0.2 in order to simulate rather a weak contrast biological sample. The illumination probe was simulated by a near-field propagation of weakly curved illumination cropped by a circular aperture. The resulting linear oversampling in the Fourier plane was approximately 2.5 and the scanning step length was selected to provide 70% linear overlap between the adjacent illuminated regions. The total scan consisted of 454 positions. The first simulation shows performance of different methods with ideal data limited only by the Poisson noise. The maximum number of photons per scanning position was 10^8. The simulated data were fully coherent and the probe positions were assumed to be precisely known. The initial guess of the probe was a top hat function with roughly the same diameter as the original probe and the object was initialized as a random array. The first test included ePIE with and without variable step size (Sec. 2.5.2) and ePIE with and without noise relaxation modulus constraint (Eq. 2.63). Methods were run for 1000 iterations in order to reach convergence and the probe started to be updated in the second iteration. The relaxation constant α in the ePIE method was 1. The Fourier error of the reconstruction is defined as

$$\chi^2/\mathrm{DoFr} \approx \frac{1}{N_{\mathrm{pos}} N_{\mathrm{pix}}^2} \sum_i \left(\frac{|\Psi_i|^2 - I_i}{\sigma_i} \right)^2$$

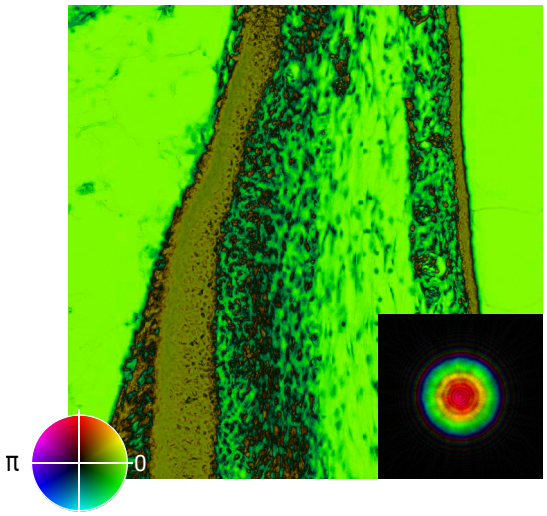

FIGURE 2.19: Image shows the simulated sample with the probe in the inset. The color intensity corresponds to the amplitude and hue to the phase.

where $|\Psi|^2$ is the Fourier plane model estimate and σ is known variance of the data I. The ideal model $|\Psi|$ should have the χ^2 metric close to 1. Note that we have approximated the degrees of freedom (DoFr) by the number of the measured pixels.

The convergence properties are shown in Fig. 2.20. The initial convergence speed is significantly improved if the variable step method (full lines) is used. However, the convergence speed is comparable in the final convergence region, because the automatically selected optimal step is close to 1. It can be also seen that larger parallelization (blue line) leads to slower convergence, because it is not possible to find non-overlapping regions anymore. On the other hand, it is surprising that the fastest convergence was reached for grouping 5 instead of the standard ePIE without grouping. Our explanation is that the small grouping provided additional relaxation that helped with convergence in the final stage. The optimal selection of the grouping depends on the convergence but also on the calculation time per one iteration. This is because similar calculation time can be achieved by faster convergence or shorter time per iteration. Figure 2.21 shows the calculation time for different bunching for our simulation (probe size 256×256 px) and it can be seen that the calculation speed gain is minimal roughly from bunching 20.

The second test shown in Fig. 2.22, was testing difference between convergence speed for noise relaxed modulus constraint (dashed lines) and hard modulus constraint (full lines). Both methods were run with the variable step length method. Again, additional relaxation helped in convergence and similarly to the previous

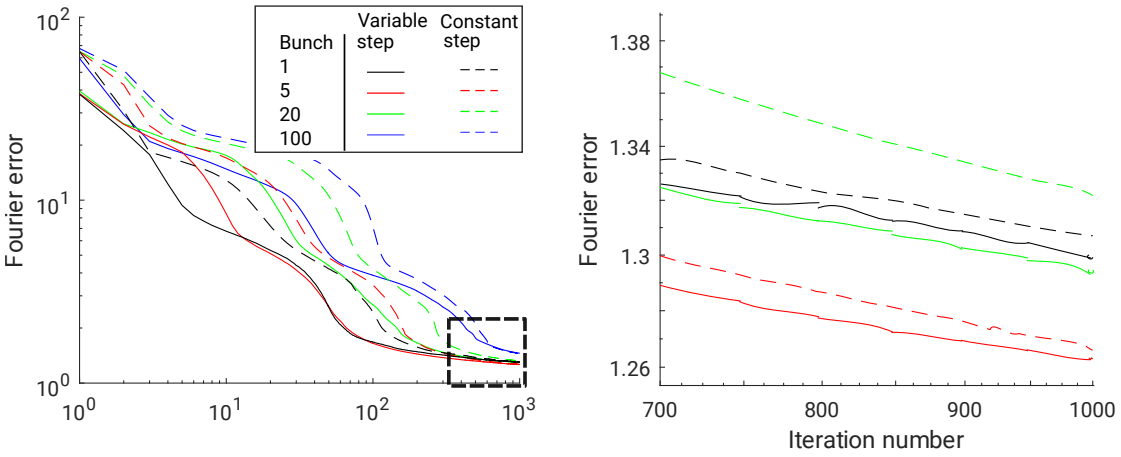

FIGURE 2.20: Dependence of the Fourier space error on the iteration number is showing significant improvement in the convergence speed if the variable step method (full lines) was used compared to the standard ePIE method (dashed lines). The performance was tested for different grouping parameter.

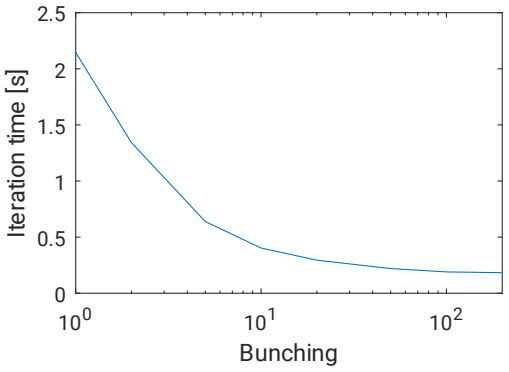

FIGURE 2.21: Dependence of the iteration time on the grouping parameter for our GPU ePIE implementation.

test, small grouping even further improved the convergence. Note that the non-relaxed standard ePIE method (full black line) finished with the highest Fourier error.

Further, we have tested the damped LSQ (Levenberg-Marquardt) solver (Eq. 2.88) and compared the convergence to the standard ePIE solver with noise relaxation. The damped LSQ method depends on the damping parameter δ that effects the convergence. In order to make the selection of δ scaling independent, we have multiplied δ by $\|P\|^2_{\max}$

$$O_{j+1} = O_j - \alpha \frac{|P|}{|P|^2 + \delta \|P\|^2_{\max}} \nabla f(O) \quad , \tag{2.95}$$

where α was selected using the variable step method (Sec. 2.5.2). Figure 2.23(a)

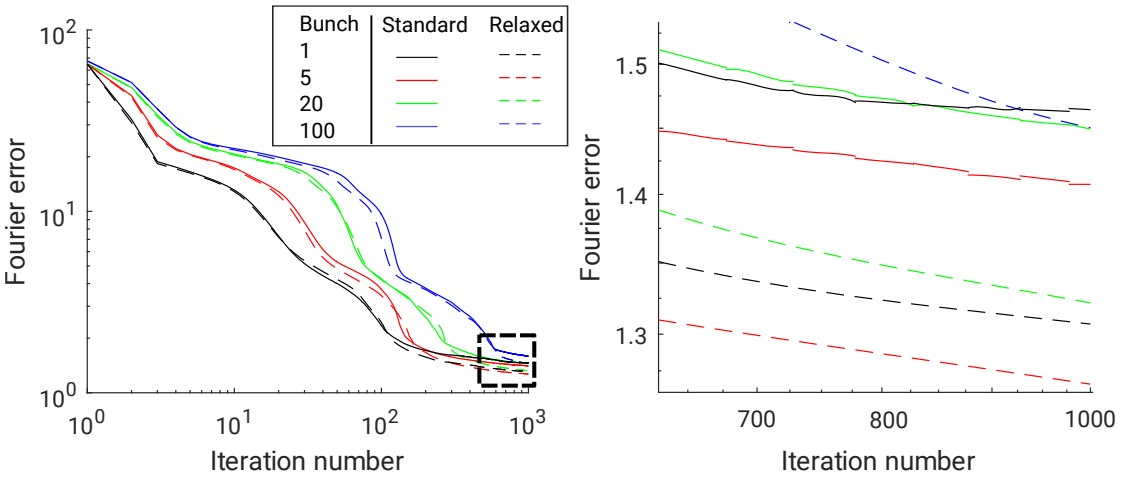

FIGURE 2.22: Dependence of the Fourier space error on the iteration number is showing improved convergence during the final iterations (right image) for noise relaxed modulus constraint (dashed) and different grouping.

shows the residual error in the Fourier space for different damping constants compared to the ePIE method. The plots show that selection of the right damping in combination with variable step length makes possible to reach the same Fourier space error in 100 iterations as ePIE method reaches in 1000 iterations. Furthermore, the damped LSQ reached lower final residuum practically for any choice of the damping parameter. The optimal damping δ that leads to the fastest convergence was around 10^{-2}. Figure 2.23(b) shows that the lower χ^2 metric corresponds to better reconstruction of the original object.

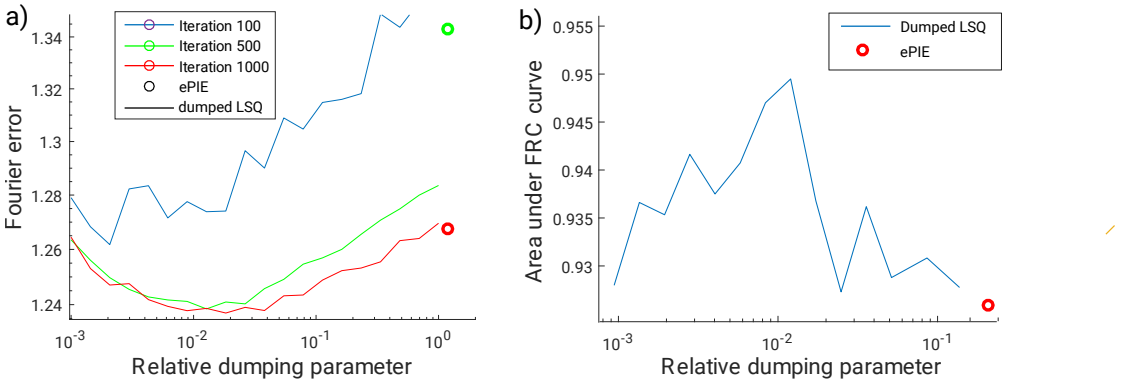

FIGURE 2.23: The plot (a) shows convergence of the damped LSQ ptychography solver. The lines show dependence of the χ^2 metric on the damping parameter δ using the LM based algorithm. The circles correspond to the error of the ePIE algorithm. Different colours show progress at 100, 500, 1000 iterations. Iterations 100 for ePIE is not shown because the Fourier error was 3.3. Plot (b) shows area under the Fourier ring correlation curve (AuC) (see Section 2.2.5) between the reconstructed and the original image. Higher AuC value means better result.

2.6 Correction of systematic errors

The basic idea of the ptychography method is very simple and numerical simulations with Poisson noise only show high robustness againts trapping in local minima. However, in the real experiments, the reconstruction has to often deal with systematic errors described in Sec. 2.2.4 such as lateral and temporal incoherence or power and beam fluctuations. Other common errors are for example imprecise knowledge of the setup geometry e.g. distance from sample to the camera or probe positions. These errors are the reasons why the constraints (Sec. 2.4.1) may not be valid and the optimal solution that satisfies all the constraints cannot be found.

In this section, we will show that some of these errors can be included into the ptychography model and thus relax the strict constraints and produce high quality reconstruction with lower quality data. Note that the achievable improvement will always depend on many factors such as oversampling, overlap or contrast of the sample.

The data corrections can be divided to two groups: offline and online corrections. Offline corrections are the corrections that can be performed before the reconstruction itself while online corrections run during the ptychography reconstruction. An example of reconstruction with corrections (incoherence/fast beam vibrations) and without is shown in Fig 2.24 using real data from the nonlinear ptychography experiment (Chapter 4).

2.6.1 Offline corrections

The advantage of the offline correction is that the ptychography solver is not slowed down during the reconstruction and that they do not depend on the convergence. The disadvantage is that additional *a priori* knowledge is needed in order to perform the correction. Moreover, statistical properties such as noise distribution of the corrected dataset maybe different from the original data due to smoothing caused by interpolation and therefore the expected noise distribution data may not be valid any more.

Background removal The background introduced by the readout bias and dark current of the detector needs to be first removed from the collected data before

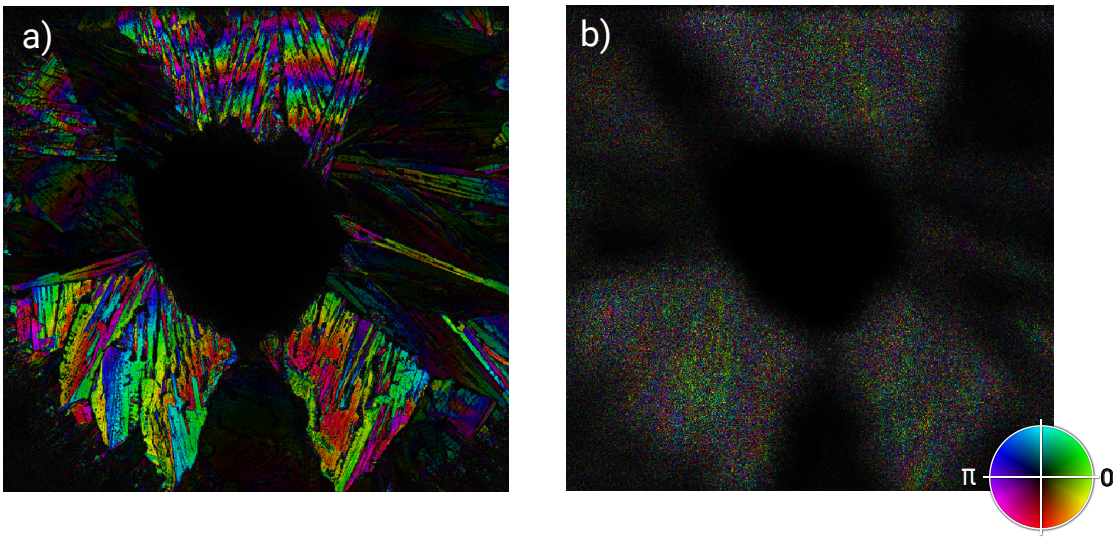

FIGURE 2.24: An example showing how inclusion of systematic errors into the model (a) can significantly improve the reconstruction compared to uncorrected modulus constraint (b) in our nonlinear (SHG) ptychography experiment with a diffusion probe (see Chapter 4). The main challenge in this dataset was rather low oversampling (≈ 1.5) in combination with limited pointing stability of the driving IR laser. It would be possible to use larger oversampling, however it was preferable to keep the acquisition time as short as possible in order to avoid other laser instabilities.

further processing can be applied. If the background offset is not negligible and it is constant over the image frame, the bias can be roughly estimated by a low quantile of the data. This rough method should be used only as a first guess. A more precise estimation can be obtained using the online methods (Sec. 2.6.2)

Geometrical corrections The first group are geometrical corrections, when the collected Fourier pattern does not correspond to regular grid coordinates. The most common examples are:

1. High NA correction in a lensless setup (Sec. 2.1.4)
2. Geometrical correction for the off-axis reflection mode [28, 42–44] (Sec. 2.1.5)

Lateral incoherence and other blurring sources

Blurring of the diffraction pattern can be caused by insufficient pointing stability of the EUV source, low coherence of the light or just the pixel size limitation. Lateral incoherence of any source can be expressed using the mutual optical intensity $J_0(r_1, r_2)$ [41, 108, 109]:

$$J_0(r_1, r_2) = \Psi(r_1)\Psi^*(r_2)g(r_1 - r_2) \quad ,$$

where $g(r_1 - r_2)$ is the coherence factor. In the case of a fully incoherent Gaussian source, the coherence factor can be estimated using the van Cittern Zernike theorem

$$g(r_1 - r_2) = \exp\left(-\frac{|r_1 - r_2|^2}{2l_c^2}\right) \quad . \tag{2.96}$$

In the case of the far-field approximation, the measured diffraction pattern can be simulated using convolution with a positive kernel g

$$|\Psi|^2 = |\Psi_0|^2 * g \quad . \tag{2.97}$$

Because convolution is a linear operation, the equation 2.97 can be also written is a matrix form

$$|\Psi|^2 = \mathbb{A}_{\text{conv}}|\Psi_0|^2 \quad , \tag{2.98}$$

where $\mathbb{A}_{\text{conv}}$ is a sparse multidiagonal matrix and Ψ, Ψ_0 are matrices reshaped to vectors.

The deconvolution can be either solved using iterative deconvolution methods for image processing as Richardson-Lucy deconvolution [72] or direct methods such as the Tikhonov regularization [110] but only if an estimation of the kernel is available.

Note that also other sources of blurring can be described as a convolution with a positive kernel [54, 98, 99] and solved if the appropriate kernel is known.

Broadband spectra corrections – temporal incoherence

A far-field diffraction pattern from a polychromatic source illuminating an achromatic sample [51] can be written as

$$|\Psi|^2 = \sum_{\lambda} W_\lambda |\Psi_\lambda|^2 \quad , \tag{2.99}$$

where W_λ is relative intensity and Ψ_λ is the Fourier transformation of the exit wave for the given spectral line λ. Furthermore, a change in the wavelength is only scaling the size of the diffraction pattern as λ/λ_0, where λ_0 is the central wavelength. Because scaling by the bilinear interpolation is only a linear operation, it can be also written as a matrix multiplication

$$|\Psi|^2 = \mathbb{A}_{\text{scaling}}|\Psi_0|^2 \quad . \tag{2.100}$$

If the spectrum is known and the specimen can be approximated as achromatic for given wavelength range, the system can be solved by the Tikhonov regularization method [110]. This method is however only an approximation for the ptychography method, when the diffraction pattern is created by a nearfield propagated illumination pattern interacting with the object. This approximation would be valid only for a smooth illumination probe (with negligible NA compared to the collected NA) or if is the spectral range is sufficiently narrow.

2.6.2 Online corrections

A disadvantage of the predeconvolution or unfolding methods is that they are all based on some prior knowledge such as smoothness or positivity that is not fully valid for diffraction datasets. Therefore, it is preferable to use priors that are consistent with the CDI methods. In the case of ptychography it is the overlap constraint. If the overlap constraint is fully valid, the "online" methods can reach significantly better results.

A potential drawback of the online corrections is that poor data quality or bad initial guess may result in convergence failure. Further, the online corrections may significantly increase the time per iteration and increase the number of iterations needed to achieve convergence.

The online corrections methods can be divided into two main group: Fourier space and real space. The Fourier space methods apply corrections on the model of the exit wave propagated onto the detector plane $\Psi(\mathbf{k})$ (Eq. (2.61)). The aim is to reach as precise description of the measured data I as possible by applying correction to the intensity distribution model $|\Psi(\mathbf{k})|^2$ in order to make the modulus constraint (Sec. 2.4.1) valid. Similarly, the real-space corrections are used to correct the overlap constraint (Sec. 2.4.1) if some of the assumptions are not valid.

Fourier space corrections

As it was shown in the previous section, if the diffraction pattern is collected in the far-field mode, most of the common Fourier space corrections can be written in a simple matrix form

$$\mathbb{A}|\Psi(\mathbf{k})|^2 + \mathbf{b} \tag{2.101}$$

where $\mathbb{A}$ is generally a sparse matrix with a dominating diagonal applied on the intensity distribution model in a vector shape and $\mathbf{b}$ contains an estimation of the background. The correction is applied using a generalized modulus constraint [20]

$$\hat{\Psi}(\mathbf{k}) = \sqrt{\frac{I(\mathbf{k})}{\mathbb{A}|\Psi(\mathbf{k})|^2 + \mathbf{b}}}\Psi(\mathbf{k}) \quad . \tag{2.102}$$

condition number κ of the matrix $\mathbb{A}$

$$\kappa(\mathbb{A}) = \|\mathbb{A}\| \cdot \|\mathbb{A}^{-1}\|$$

significantly affects the convergence speed and reconstruction quality. CDI methods effectively search for a solution Ψ that satisfy

$$I(\mathbf{k}) = \mathbb{A}|\Psi(\mathbf{k})|^2 + \mathbf{b} \tag{2.103}$$

with an additional constraint in the real space. This can be understood as a deconvolution or more generally an unfolding problem. Bad conditionality (i.e. $\kappa(\mathbb{A}) \gg 1$) result in amplification of the noise proportional to $\kappa(\mathbb{A})$. Therefore, in order to achieve the same reconstruction quality, either much higher signal to noise ratio is needed or a stronger constraint in the real-space (i.e. more overlap) that provides sufficient regularization.

Estimation of the background offset

In the ideal case, the background in CDI data should be removed already during data acquisition, however often it is not possible for example when the background is created by incoherent scatter from dust and scratches on the microscope objective or when the background offset is caused by truncated readout noise or by small changes in the CCD camera readout offset long scans. Several different methods have been published for automatic estimation of the background [84, 111]. However, none of them is really robust, fast and universal, therefore we have implemented a different method based on maximum likelihood estimation.

If the background for each diffraction pattern can be sufficiently approximated by a single offset value b_i, then using the following Fourier space error estimation

$$\chi^2 = \sum_i \sum_{(\mathbf{k})} \left[W(\mathbf{k})(|\Psi_i(\mathbf{k})|^2 + b_i - I_i(\mathbf{k}))\right]^2 \quad . \tag{2.104}$$

The optimal nonnegative background correction that minimizes the χ^2 criterion can be estimated as

$$b_i^{(j+1)} = \left[b_i^{(j)} + \alpha \frac{\sum_{\mathbf{k}} W(\mathbf{k}) \left(I_i(\mathbf{k}) - |\Psi_i(\mathbf{k})|^2 - b_i^{(j)} \right)}{\sum_{\mathbf{k}} W(\mathbf{k})} \right]_+ \quad , \tag{2.105}$$

where $b_i^{(j+1)}$ is an updated value of the background and $0 < \alpha < 1$ is a relaxation constant that prevents oscillations. The background is forced to be nonnegative. Finally, W denotes distribution of reliability of each measured pixel $\mathbf{k}$. W is estimated using the following empirical formula

$$W(\mathbf{k}) = \begin{cases} \frac{1}{\sigma^2} & \text{if pixel } k \text{ is measured} \\ 0 & \text{otherwise} \end{cases} \tag{2.106}$$

This formula eliminates the effects of the non-measured and the high intensity regions which may contain large misfit between data and model but absolutely no information about the background.

Estimation of the incoherence blurring kernel

The estimation of an unknown blurring kernel shape is generally a blind deconvolution task, however in ptychography the overlap constraint provides additional constraints which make it a simple optimization task.

The most straightforward method is based on the maximum likelihood solution. If the blurring kernel is described by a low number of parameters, the optimal parameters are iteratively selected to minimize misfit (maximize likelihood) between data and model [49]. The simplest model is a Gaussian kernel parametrized by the kernel width σ that is inverse proportional to the lateral coherence length

$$g(x_1 - x_2) = \exp\left(-\frac{|x_1 - x_2|^2}{\sigma^2} \right)$$

In some cases the full matrix of the PSF $\mathbf{g}$ needs to be recovered. The convolution equation

$$|\Psi|^2 * \mathbf{g} = I \quad ,$$

can be rewritten to a matrix form

$$\mathbb{A}\mathbf{g} = I \quad , \tag{2.107}$$

where $\mathbf{g}$ is the searched point spread function (PSF) of size $(N_{\mathrm{ker}}; N_{\mathrm{ker}})$, I are all measured diffraction patterns in a long vector and $\mathbb{A}$ is an appropriate matrix composed from the current model estimation $|\Psi|^2$ shifted by integer distances $[-N_{\mathrm{ker}}/2, \ldots, N_{\mathrm{ker}}/2]$ and the size of the matrix is $(N_x \cdot N_y \cdot N_{\mathrm{pos}}; N_{\mathrm{ker}}^2)$. Note that Eq. (2.107) is an example of strongly ill-posed problem, so although the number of equations is several orders higher than the number of variables, the task is still difficult to solve. Therefore, we have used the nonlinear Tikhonov regularization with positivity constraint [112]:

$$\mathbf{g} = (\mathbb{A}^T\mathbb{A} + \lambda\mathbb{I}/|\mathbf{g}|)^{-1}\mathbb{A}^T I \quad ,$$

where $\mathbb{I}$ is the identity matrix and the optimal regularization constant λ is chosen using the L-curve method [113]. The optimal PSF estimated by a simple maximum likelihood and by the Tikhonov method is shown in Fig. 2.26 and Fig. 2.25 shows measured diffraction pattern and pattern after "deconvolution".

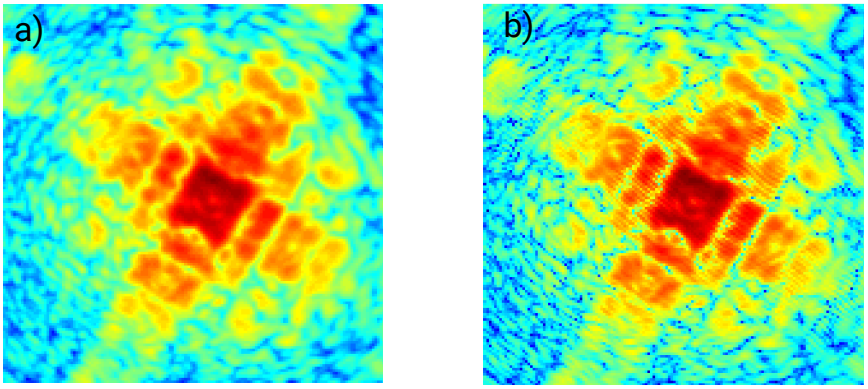

FIGURE 2.25: Comparison of measured diffraction pattern (a) and "deconvolved" model (b) using convolution kernel estimated by the ML method. The diffraction pattern is produced by scatter of HeNe beam through a metal mesh. The small "incoherence" was caused by fast air fluctuations and the finite pixel size.

Thibault *et al.* [16] proposed a different way for correction of the modulus constraint. A decomposition of the exit wave into a set of orthogonal modes Ψ_i is searched resulting in a modulus constraint is in the following form

$$\hat{\Psi}_i(\mathbf{k}) = \sqrt{\frac{I(\mathbf{k})}{\sum_i |\Psi_i(\mathbf{k})|^2}}\Psi_i(\mathbf{k}) \quad . \tag{2.108}$$

The main advantage but also drawback of the Thibault's "mixture of states"

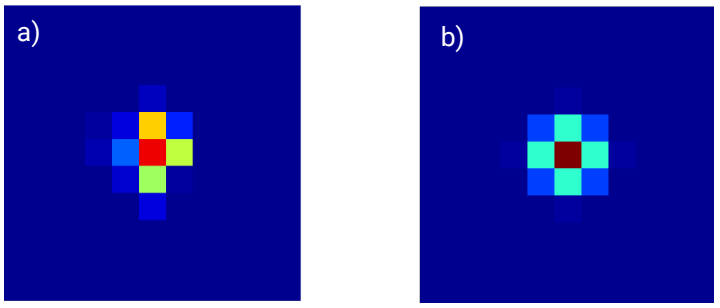

FIGURE 2.26: Difference between PSF (point spread function) estimated by the Tikhonov method and maximum likelihood estimation using the diffraction pattern in the Fig. 2.25. The left image shows that the additional freedom was used to perform a sub-pixel shift of the model in order to fit better the data.

method is that it has no prior knowledge about the searched imperfection. Therefore, it is a versatile method that can deal with a large variety of systematic errors as shown in Ref. [50]. On the other hand, linear corrections methods can provide more robust results with lower computational complexity. A good option is to combine both methods to obtain the following modulus constraint

$$\hat{\Psi}_i(\mathbf{k}) = \sqrt{\frac{I(\mathbf{k})}{\mathbb{A}\sum_i |\Psi_i(\mathbf{k})|^2 + \mathbf{b}}}\Psi_i(\mathbf{k}) \quad . \tag{2.109}$$

which makes possible to use less modes and thus speed up the calculation.

Broadband illumination

As it was already mentioned in the Section 2.6.1, non-monochromatic illumination can be modelled as a matrix product

$$I(\mathbf{k})_{\text{broad}} = \mathbb{A}_{\text{broad}}|\Psi(\mathbf{k})|^2 \quad , \tag{2.110}$$

where $\mathbb{A}_{\text{broad}}$ is a precalculated sparse matrix. If this approximation is not valid, it is necessary to propagate each wavelength independently using the ptychography information multiplexing (PIM) algorithm [52].

Multicolor ptychography (ptychography information multiplexing)

The idea of reconstructing multiple orthogonal incoherent modes proposed by Thibault *et al.* [16] was extended by Batey *et al.* [52] into recovery of multiple colour modes called PIM. The PIM method uses following modification of the modulus constraint

$$\hat{\Psi}_i(\mathbf{k}) = \sqrt{\frac{I(\mathbf{k})}{\sum_i |\Psi_i(\mathbf{k})|^2}}\Psi_i(\mathbf{k}) \tag{2.111}$$

where $\Psi_i(\mathbf{k}) = \mathfrak{F}(P_i \cdot O_i)$ and P_i, O_i are i-th colour probe and object. In the PIM method, the modes are significantly less constrained compared to the "mixture of states" method, where the orthogonality is enforced. Therefore, the convergence is slower and it is very difficult to separate the colours without artefacts, in particular at low spatial frequencies [52]. Therefore, if the object can be considered sufficiently achromatic, more robust results can be achieved when the object reconstructions are linked together by proper upscaling as it is shown in Section 5.7.

Real-space corrections

Similarly to the Fourier space constraint, there exist many different effects that violate the overlap constraint. The most common in laboratory scale experiments are:

- Sample is not thin
- Sample is not isotropic (i.e. properties of the sample change in dependence on the illumination)
- Positions of the illumination probe are wrong or the estimation of the reconstructed pixel size is imprecise
- Intensity and position fluctuations of the illumination probe

If these effects are not eliminated in the experiment or included into the CDI algorithm, it will result in artefacts and deterioration of the reconstruction quality.

Estimation of the probe position correction and the pixel scale

In small laboratory experiments, it is practically impossible to measure the effective numerical aperture (NA) with sufficient precision required by the ptychography method. Therefore, the optimal NA that determines the pixel size of the reconstruction needs to be recovered from the measured dataset.

This problem is closely related to probe position estimation, because incorrect estimation of the pixel size scales the distance between the probe positions. It is possible to estimate the position of each probe separately, however this works

only for well structured objects with good contrast and signal for all regions and sufficient data quality. Moreover, the precision of the piezoelectric positioners is not limiting reconstruction quality of our visible light experiments.

The basic idea of different position corrections algorithms [25, 90, 91, 95, 114] is similar: minimize distance between data and model by adjusting the (x,y) coordinates of the illumination probe for each measured position. If the position errors are large, the minimization task gets strongly nonlinear and non-convex, therefore robust minimization methods such as simulated annealing [25] need to be applied. If the errors are reasonably small, i.e. less than 10% of the probe diameter, it is possible to use a simpler method similar to [90] that compares each object patch before and after applying the modulus constraint (Eq. 2.61). If the probe position is wrong, the updated object O_{new} gets slightly shifted by a distance that is much smaller than one pixel. The optimal shift can be estimated after every iteration as the centre of the cross-correlation between the patches

$$X_{\mathrm{corr}} = \mathfrak{F}^{-1}\left[\mathfrak{F}\left(O_{\mathrm{new}}\right)^{*}\mathfrak{F}\left(O_{\mathrm{orig}}\right)\right] \quad , \tag{2.112}$$

where O_{new} and O_{orig} are the updated object and original object patch. The optimal shift is determined as position of maximum of X_{corr}. Note that if the probe has a limited support, the patch size does not need to contain more than the probe support area. Because the centre needs to be estimated with high precision, it is necessary to use a sub-pixel precision image registration method e.g. [115]. We propose an alternative and much faster method which gives comparable or better results. In Appendix B, we have demonstrated on our measurements that it can provide more than 1000× better precision than the image resolution. The method is based on calculation of the centre of mass of the cropped cross-correlation matrix X_{corr}

$$\mathrm{d}x = \mathrm{CoM}\left\{\left[X_{\mathrm{corr}} - (1-\alpha)\max\left(X_{\mathrm{corr}}\right)\right]_{+}^{2}\right\} \quad , \tag{2.113}$$

where CoM denotes centre of mass, the threshold constant α can be selected as roughly 0.1 and $[]_{+}$ denotes the positive part.

The cross-correlation method is very fast, usually the additional computational cost is 5-10% of the iteration time. However, from our experience, it can often fail to provide solution that minimize the Fourier error and it can lead to inferior results. In these cases it is necessary to use a different method. Therefore, we have developed two methods that provide better stability. The orthogonal probe

relaxation method (OPRP) presented in Sec. 2.7 that is able to recover the probe positions even in the case of unstable illumination. Another option that can be also easily combined with the OPRP is to use a gradient descent algorithm with variable step size similar to the method introduced in Section 2.5.2. The only difference is that instead of Eq. 2.90

$$\chi^2 = \sum \left[W(|\mathfrak{F}\left[(O + \alpha\Delta O)(P + \alpha\Delta P)\right]|^2 - I) \right]^2 \quad ,$$

in this case, we minimize two equations

$$\begin{aligned} \chi_x^2 &= \sum \left[W \left\{ \left| \mathfrak{F} \left[O \left(P + \alpha_x \frac{\mathrm{d}P}{\mathrm{d}x} \right) \right] \right|^2 - I \right\} \right]^2 \\ \chi_y^2 &= \sum \left[W \left\{ \left| \mathfrak{F} \left[O \left(P + \alpha_y \frac{\mathrm{d}P}{\mathrm{d}y} \right) \right] \right|^2 - I \right\} \right]^2 \quad , \end{aligned} \tag{2.114}$$

where α_x, α_y are the optimal steps in the horizontal and vertical direction respectively. The probe gradient $\frac{\mathrm{d}P}{\mathrm{d}x}$ is equal to

$$\frac{\mathrm{d}P}{\mathrm{d}x} = \mathfrak{F}^{-1}[2\mathrm{i}\pi \mathbf{q}_x \mathfrak{F}(P)] \quad . \tag{2.115}$$

It would be also possible to use derivatives of the object $\frac{\mathrm{d}O}{\mathrm{d}x}$, $\frac{\mathrm{d}O}{\mathrm{d}y}$, however it would require more computational power. Equation 2.114 can be approximated by a second order polynomial that is solved in the same way as it was shown in Section 2.5.2, where the optimal step is given by Eq. 2.93:

$$\alpha_i^{(x,y)} = \frac{\sum W A_1 A_2}{\sum W \left(2A_1 A_3 + A_2^2\right)} \quad .$$

Because the $\alpha_{x,y}$ is solved separately for each position i, it is already equal to the optimal probe position correction in the given direction. Note that we solve vertical and horizontal direction separately, which may not lead to the most optimal solution. However, it is a sufficient approximation that is valid for small correction steps. This method is computationally more expensive than the cross-correlation based method. The computational cost is between 10-30% of the iteration time.

Another simple yet robust method is to use the idea of the simulated annealing [25].

In order to make the calculation fast and still compatible with the multiple incoherence corrections methods, the following expression was optimized:

$$\begin{aligned}\chi_0^2 &= \sum \left[W\left(|\Psi_{\text{incoh}}|^2 - I\right)\right]^2 \\ \chi_{\text{new}}^2 &= \sum \left[W\left(|\Psi_{\text{incoh}}|^2 - |\Psi|^2 + \left|\mathfrak{F}\left[\hat{O}\cdot P\right]\right|^2 - I\right)\right]^2 \quad ,\end{aligned} \tag{2.116}$$

where $\hat{O}$ denotes object array shifted by a random vector $\mathbf{x}$, $|\Psi_{\text{incoh}}|^2$ is the incoherent model of the illumination at the detector and $|\Psi|^2$ denotes only the first (i.e. coherent) mode. Only one random step is investigated in every iteration and if it reduces the χ^2, then the probe position is corrected. The advantage of this method is simplicity and robustness against trapping in local minima.

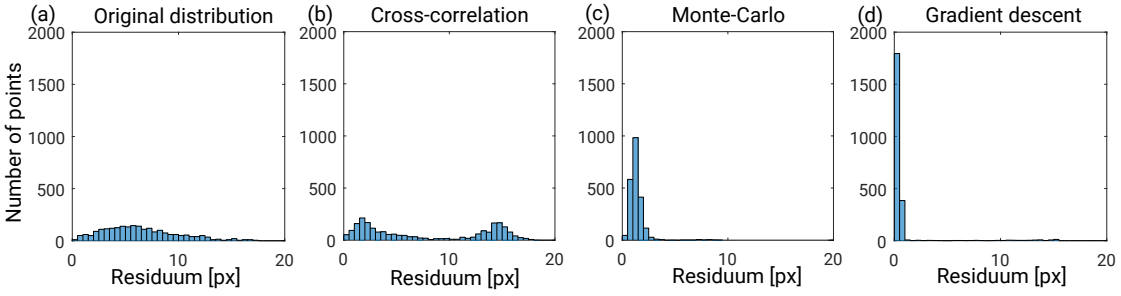

FIGURE 2.27: Comparison of different position correction methods applied on the simulated dataset presented in Section 2.5.3. Each test was repeated 5× with a random position shift. (a) shows the original residual distribution with no position correction, (b) is the cross-correlation method, (c) is a simple Monte-Carlo method, (d) is our new gradient descent solver. Bin width is 0.5 px.

Figure 2.27 shows residua of the tested position correction methods. The testing dataset was described in Section 2.5.3. Note that the test object (Fig. 2.19) has a large region in top left corner with very low contrast. The original residua without correction are shown in Fig. 2.27(a). The cross-correlation method 2.27(b) improved the positions that were close to the optimum but the distant positions were not able to converge and they drifted away. A simple Monte-Carlo (MC) method that tested several random shifts in every iteration and selected the one with lower error was tested in Fig. 2.27(c). MC was not able to get the average position error below ≈1.5 px, however it had the lowest number of outliers. A simulated annealing with additional parameters as shown Ref. [89] would need to be implemented in order to reach lower errors. Finally, the gradient descent method in Fig. 2.27(d) was able to recover the positions with better than 0.5 px accuracy, despite a few outliers. Our recommended method is therefore the gradient descent combined with the Monte-Carlo method run every n-th iteration to avoid outliers.

In order to guarantee robust probe position recovery even for low SNR or low contrast regions and speed up the convergence, it is preferable to use a simple geometrical model that is sufficiently describing errors in the experimental geometry using a low number of variables. Furthermore, if the accuracy of the actuators is measured (see Appendix B), then the probe positions can be allowed to freely move only inside a circle defined by the known accuracy. The geometry model is then only loosely following the average trend and thus provides much faster information propagation between the distant regions of the reconstruction. In our experiments, we have used a model that includes scaling and rotation of the coordinates. We have tested also asymmetry that would be caused by horizontal rotation of the camera with respect to the sample, however it was always negligible.

Generally, the drawback of position correction methods is that the probe shifts estimation is not very good for weakly scattering samples and in the case of variable illumination.

Thick sample

If the sample thickness is not much smaller than the depth of focus (Eq. 5.9), the projection approximation is not fully valid [80, 83, 116]

$$\psi_i(\mathbf{x}) = P(\mathbf{x})O(\mathbf{x}+\mathbf{x}_i) \quad . \tag{2.117}$$

Optical thickness is a common issue particularly for visible light microscopy, but it can be limiting also for X-ray nanotomography. In the EUV range, the thickness of the imaged specimen is usually limited by the attenuation length although thin but not flat specimens result in similar issues (see Sec. 5.6).

In order to simulate the effect of the sample thickness, Maiden *et al.* [25] proposed to model the thick sample using multiple thin layers and the light is propagated between the layers by ASM propagation. The separation of the layers should be roughly corresponding to the depth of focus for the given experiment (Eq. 2.37)

$$\mathrm{DoF} = \frac{\lambda}{(\mathrm{NA}_c + \mathrm{NA}_o)^2} \quad , \tag{2.118}$$

where NA_c and NA_o are numerical apertures of the collector and objective. DoF roughly gives the resolution limit in the z-direction.

An example of the reconstruction procedure for a two layer sample reconstructed by the 3PIE or maximum likelihood method [25, 83]:

1. $\psi_1 = P_1 O_1$ standard projection approximation for the first layer
2. $P_2 = \mathcal{P}(\psi_1, \Delta z)$ propagate to the next layer by distance Δz
3. $\psi_2 = P_2 O_2$ apply standard projection approximation for second layer
4. $\Psi = \mathcal{P}(\psi_2, z)$ propagate to the detector plane
5. $\tilde{\Psi} = \frac{\sqrt{I}}{|\Psi|}(\Psi)$ apply standard modulus constraint
6. $\tilde{\psi}_2 = \mathcal{P}(\tilde{\Psi}, -z)$ backpropagate to the second layer and update O_2 and P_2
7. $\tilde{\psi}_1 = \mathcal{P}(P_2, -\Delta z)$ backpropagate to the updated probe P_2 to the first layer and update O_1 and shared probe P_1

Simple repeating of the algorithm allows us to reconstruct several layers of the sample. One of the drawbacks of this method results from the quadratic dependence of DoF on NA. Most of the depth information is kept in the high spatial frequencies (high NA), therefore convergence of the low spatial frequencies is slower and the zero spatial frequency (e.g. constant offset) cannot be assigned at all.

The idea of 3PIE is very similar to ideal used in the incoherent X-ray tomography method called laminography, when a cone beam is used for scanning a thin sheet object and multiple layers are reconstructed [117]. And similarly to laminography, the reconstruction quality along the beam propagation axis is limited by so called missing cone in angular space. This results in very poor reconstruction of the low spatial frequencies within the layers and limited resolution along beam propagation.

There are also other methods that claim the possibility to extract 3D information from a single complex valued projection of the object such as Ankylography [56, 118]. However, these methods only use the phase and amplitude information to recover a pseudo-3D reconstruction that satisfies additional constrains such as positivity, smoothness and compactness.

Multilayer probe

The multilayer ptychography method (3PIE) can be easily extended also for a reconstruction of a single thin sample with a multilayer probe. This is a relatively common case, diffuser optics or a semi-transparent beam-block is used behind a sample as it will be shown in Section 3.1. Furthermore, multilayer probe recovery can be used in the near-field ptychography, when one probe is given by the sample illumination and the second is usually assumed to be known and it is the ASM propagation factor (Eq. 2.22). If the propagation factor is not precisely known or the assumption of spherical illumination wavefront is not valid for example due to astigmatism, then the ASM factor can be recovered as an additional phase-only probe.

The implementation of the multilayer probe is almost identical to the multilayer object reconstruction presented in the previous section. The only difference is in two steps:

2. $\mathbf{O_2} = \mathcal{P}(\psi_1, \Delta z)$ propagate to the next layer by distance Δz
7. $\Psi = \mathcal{P}(\mathbf{O_2}, -\Delta z)$ backpropagate to the updated object P_2 to the first layer and update O_1 and shared probe P_1

The advantage of this method is that the probe information is more constrained than the object information, therefore the convergence and stability is better then if a multilayer object is reconstructed.

Anisotropic sample

Anisotropy of the sample is an effect that is mostly negligible in X-ray microscopy. However, this effect should be included if CDI methods are used for microscopy of crystalline materials. Anisotropy of the sample is not negligible if the illumination probe cannot be approximated as a nearly flat wavefront (see Section 4.3). In that case, the projection approximation

$$\psi_j(\mathbf{x}) = P(\mathbf{x}) \cdot O(\mathbf{x} - \mathbf{x}_j) \tag{2.119}$$

is not valid because each part of the probe will see different phase-shift and possibly also attenuation in the same point of the object $O(\mathbf{x}, \boldsymbol{\alpha})$ depends on the local wavefront gradient of the illumination probe (i.e. relative illumination angle).

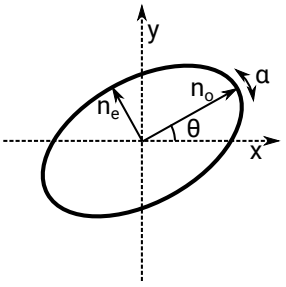

FIGURE 2.28: A schematic of the coordinates system in the anisotropy calculation. The crystal axes are rotated with respect to the beam by angle $\Theta \ll 1$.

Here, we will show, how to use the discrepancy in the real space constraint to estimate the local derivative of the complex refractive index with respect to the illumination angle. There are several effects that change the object properties if the illumination NA is very high

1. Refractive index change due to anisotropy of the sample. The refractive index can be approximated for small angle α when the crystal optical axes are rotated by angle Θ with respect to the imaging coordinates as shown in Fig. 2.28. The refractive index of the anisotropic crystal can be calculated for given ordinary and extraordinary refractive indices (n_o, n_e) and rotation of the crystal Θ

$$n = \sqrt{(n_o \cos\Theta)^2 + (n_e \sin\Theta)^2} \tag{2.120}$$

A small rotation α of the crystal or coordinate system around the angle Θ can be expressed as:

$$\begin{aligned} n &= n_0 + \frac{\mathrm{d}n}{\mathrm{d}\alpha}\alpha \\ n &= n_0 + \frac{(n_o^2 - n_e^2)\sin 2\Theta}{2\sqrt{(n_o \cos\Theta)^2 + (n_e \sin\Theta)^2}}\alpha \quad . \end{aligned} \tag{2.121}$$

Note that for $0 < \Theta < \pi/2$ this equation depends linearly on the value of α. If the crystal axes are close to the optical axes (i.e. $\Theta = 0$), it is possible to use the following approximation of the Eq. 2.120

$$n = \sqrt{(n_o \cos\alpha)^2 + (n_e \sin\alpha)^2} \approx n_o + \frac{2n_e - n_o}{4}\alpha^2 \quad , \tag{2.122}$$

which shows that the change of the refractive index depends on the illumination angle α only as a second order correction.

2. Path through the sample is elongated

$$z = \frac{z_0}{\cos\alpha} \approx z_0(1+\alpha^2) \quad . \tag{2.123}$$

3. Reflection from the sample air interface - using the Fresnel coefficients, the reflected electric field is equal to

$$r_{TE} = \frac{n\cos\alpha - n'\cos\alpha'}{n\cos\alpha + n'\cos\alpha'} \quad , \tag{2.124}$$

where n, n' are refractive indices on the interface and $\cos\alpha' = \sqrt{1-\left(\frac{n}{n'}\sin^2\alpha\right)}$. This can be approximated for small α

$$r_{TE} = \frac{n-n'}{n+n'}\frac{n\alpha^2}{n+n'} \quad . \tag{2.125}$$

4. Thick sample - in our approximation, we have to assume thin sample with sufficiently flat surface

The previous list shows that the only effect that depends on the angle α in the first order approximation is the local derivative of the refractive index due to anisotropy. Moreover, all even-order corrections will be naturally eliminated as we will show later.

The effects of anisotropy on the exit wave field can be expressed from Eqs. (2.123, 2.121) using only the first order expansion

$$\begin{aligned} \hat{\psi} &= \exp\left(\mathrm{i}k_0 n(\alpha) z(\alpha)\right) \\ &\approx \exp\left[\mathrm{i}k_0\left(n_0 + \frac{\mathrm{d}n}{\mathrm{d}\alpha}\alpha\right) z_0\right] \\ &= \psi_0 \psi^{\alpha}_{\mathrm{corr}} \quad , \end{aligned} \tag{2.126}$$

where k_0 is the wave number in vacuum. This equation shows that it is possible to express the anisotropy effect as a product of the original exit wave function ψ_0 and a correction factor that depends on α. Note that till now, all calculations were done in one dimensional case. In the real samples, we need to assume different angle in each axis (α_x,α_y) and different correction factor for the anisotropy of the

crystal (O_x,O_y). Because the angles $\alpha_{x,y}$ depend on the probe, it is possible to use the following notation

$$\hat{\psi}_j(\mathbf{x}) = P(\mathbf{x})O(\mathbf{x}+\mathbf{x}_j)O_x(\mathbf{x}+\mathbf{x}_j)^{\alpha_x(\mathbf{x})}O_y(\mathbf{x}+\mathbf{x}_j)^{\alpha_y(\mathbf{x})} \quad , \tag{2.127}$$

where $O_x(\mathbf{x}), O_y(\mathbf{x})$ contain a correction factor for vertical and horizontal direction and $\alpha_x(\mathbf{x}), \alpha_y(\mathbf{x})$ are matrices of vertical and horizontal angles. The equation 2.127 cannot be easily solved by ptychography because it is not possible to separate terms related to the object and probe. A further approximation that uses the assumption of small α will help to separate it

$$\tilde{\psi}_j(\mathbf{x}) = P(\mathbf{x})O(\mathbf{x}+\mathbf{x}_j)+\sum_{xy}[P(\mathbf{x})\alpha_{xy}(\mathbf{x})][O(\mathbf{x}+\mathbf{x}_j)(\ln(|O_{xy}(\mathbf{x}+\mathbf{x}_j)|)+\mathrm{i}\ \mathrm{angle}(O_{xy}(\mathbf{x}+\mathbf{x}_j))] \quad . \tag{2.128}$$

Now, the first square bracket contains only terms related to probe position and the second to the object position. The previous equation 2.128 can be written in a general form

$$\tilde{\psi}_j(\mathbf{x}) = \sum_k P_k(\mathbf{x})O_k(\mathbf{x}+\mathbf{x}_j) \quad , \tag{2.129}$$

where $P_{2,3}(\mathbf{x}) = P(\mathbf{x})\alpha_{x,y}(\mathbf{x})$ and $O_{2,3}(\mathbf{x}) = O(\mathbf{x})[\ln(|O_{xy}(\mathbf{x})|)+\mathrm{i}\ \mathrm{angle}(O_{xy}(\mathbf{x})]$. If the objects O_k are reconstructed, the anisotropic correction ψ_{corr} can be expressed as

$$\psi_{\mathrm{corr},(2,3)} = \exp\left(\frac{O_{2,3}O^*}{|O|^2}\right) \quad . \tag{2.130}$$

Despite Eq. 2.129 seems very similar to the result from the multicolour ptychography (Sec. 2.6.2) that has the far field diffraction intensity distribution $I(\mathbf{k})$ in the following form

$$I \sim \sum_k |\mathfrak{F}\left[P_k(\mathbf{x})O_k(\mathbf{x}+\mathbf{x}_j)\right]|^2 \quad , \tag{2.131}$$

while the anisotropy results in the form

$$I \sim \left|\mathfrak{F}\left[\sum_k P_k(\mathbf{x})O_k(\mathbf{x}+\mathbf{x}_j)\right]\right|^2 \quad , \tag{2.132}$$

the method solving of the anisotropy needs to be fundamentally different. This is because the waves are added coherently, therefore the modulus constraint is fully valid and only the real space constraint is violated. Therefore, it is necessary to find another method that uses the available knowledge to find the solution.

We propose the following procedure:

1. We assume a linear angle gradient $A_{x,y}$ over the illuminated r that is proportional to

   $$A_{x,y} = c_{x,y}\left(-N_{\mathrm{pix}}/2, \ldots, N_{\mathrm{pix}}/2\right)$$

   The probes P_k are defined as $P_1 = P$ and $P_k = P_1 A_k$ for $k \in [2, 3]$ and thus they are mutually orthogonal.

2. The object is searched in a similar way as in the multicolour ptychography [52], because we use a single shared modulus constraint and then we reconstruct three objects using three different orthogonal probes. In order to allow separation of the O_k functions, we use the linearity of the searched correction. The signs of the correction objects and probes are flipped in every iteration

   $$\begin{aligned} \hat{O}_k &= -O_k \\ \hat{P}_k &= -P_k \quad . \end{aligned} \tag{2.133}$$

   If we assume the sample to have the same transmission along each axis (i.e. isotropic attenuation), it is possible to use an additional relaxed constraint

   $$O_k = -\left[(1-\alpha)O_k + \alpha \mathrm{i} O_1 \Im\left(\frac{O_1^* O_i}{|O_1|^2}\right)\right] \quad . \tag{2.134}$$

   This alternating direction scheme will remove from the reconstruction all effects that depend on the angle of the 0th and 2nd order and only the 1st order corrections will be recovered. We assume that the higher order corrections can be neglected.

The results from our anisotropic ptychography method are shown in the Section 3.3. We have applied the method on a thin crystal, however we expect similar anisotropy effects to be also observed in the EUV wavelength range for multilayer mirrors with some buried damage beneath the reflexive layers.

2.7 Orthogonal Probe Relaxation (OPR)

One of the critical issues of high harmonic EUV sources is the illumination stability. The position of the beam on the illumination forming aperture is changing and the intensity is fluctuating during the ptychography scan. This results in slight variations of the illumination probe for each diffraction pattern and therefore the standard overlap constraint is not valid any more because each scanning position has effectively a different probe. This effect can be also observed with visible light experiments, when the optical table bends during the scan due to small temperature variations or in synchrotron experiments due to sample and optics drifts. The issue is that the standard ptychography method does not allow reconstruction of an independent probe for every scanning position because the probes would not be separable from the object.

We propose a method called Orthogonal Probe Relaxation (OPR) [6] that provides additional relaxation of the real-space overlap constraint without need of additional constraints or prior information except the standard constraints already used in ptychography. In order to allow sufficient variability of the illumination probe but still guarantee unique separation of the object and probe, it is necessary to restrain the probe reconstructions and do not allow them to be entirely independent. If the beam movements can be considered sufficiently small and the probe can be considered sufficiently smooth then the probe movement can be approximated by a decomposition into a low number of coherent orthogonal modes. Therefore, the projection into a lower dimensional space defined by several orthogonal modes was selected as the optimal constraint. As it was already shown in the previous sections, ptychography allows recovery of large amounts of additional information. Therefore, if the number of the orthogonal modes is significantly smaller than number of scanning positions, reconstruction of a lower number of modes is possible and also the separability of the object and probe can generally be done for an aperiodic sample. However, additional regularizations such as the relaxed probe support constraint (Sec. 2.4.1) or the relaxed positivity constraint (Sec. 2.4.1) can be beneficial in the case of imperfect datasets and they can speed up the convergence.

There has been already published methods for reconstruction of several incoherent modes that can be used to include the coherence properties into the ptychography

reconstruction [16, 52]. These methods are based on a principle that incoherence-like effects in the Fourier plane are captured into the higher order orthogonal modes. Their reconstruction methods use the fact that each of the modes can be propagated into the Fourier plane independently and they are incoherently summed ($I \sim \sum |\Psi_i|^2$) at the detector. The modes are then reconstructed in order to obey the modulus constraint that is not satisfied by a single mode. On the contrary, coherently added modes ($I \sim |\sum \Psi_i|^2$) satisfy the same modulus constraint as a single mode $\Psi = \sum \Psi_i$ and therefore the modulus constraint cannot be used to separate them. This is the reason why a different method that is sensitive to mismatch in the real space domain needs to be used.

The standard ptychography reconstruction algorithms are designed to calculate a single average probe for all positions. However, it is possible to modify them to calculate a separate probe for each scanning position. These separate reconstructions then need to be linked together by a projection into a lower dimensional space at the end of each iteration cycle. We have used the fast randomized singular value decomposition (fSVD) [119, 120] in order to keep the decomposition time low even for large datasets. fSVD is an approximation of the standard SVD method that is used to recover only several most important singular vectors. The basic idea of fSVD is that if we need a decomposition of matrix A, we need to find an orthonormal matrix Q with n columns so that $A \approx QQ^TA$. Then the SVD decomposition of a matrix B defined as $B = Q^TA = USV^T$ can be easily calculated because B can be very small. Finally, a decomposition of the original matrix A is recovered as

$$A \approx QQ^TA = Q(USV^T) = (QU)SV^T \quad ,$$

where QU are the singular vectors (topos), S is a diagonal matrix of the singular values and V is an evolution matrix (chronos) of the singular vectors. The orthonormal matrix Q is formed by the action of the matrix A on a collection of random vectors X. In the simplest case, the orthonormal matrix Q can be calculated as the QR decomposition of the matrix product AX

$$QR = AX \quad ,$$

however there are also other more advanced methods available [119].

We have used fSVD to find n components (eigenprobes) of the most optimal projection basis U for the matrix form of the probes P

$$[U, S, V] = \text{fSVD}(P, n) \quad , \tag{2.135}$$

where the orthonormal projection basis U we call eigenprobes, S are the singular values and V contains complex time evolution of the eigenprobes U. Note that the principal component analysis (PCA) and SVD are practically identical methods because columns of US are the principal components and columns of V are the principal axis. The lower dimensional projection $\hat{P}$ is then defined as

$$\hat{P} = USV^* \quad , \tag{2.136}$$

where * denotes complex conjugate. The low dimensional projection $\hat{P}$ still provides a different probe for every position and the information propagation is provided by the selection of the optimal SVD modes.

If the orthogonal probe modes are weak, the approximative SVD methods can fail to recover the initial guess. Therefore, the ability to recover the OPR modes can be improved either by using full (without approximation) SVD for initial iterations, which may be however extremely slow. The second option it is to use a relaxed version of projector 2.136

$$P = \alpha P + (1 - \alpha)USV^* \tag{2.137}$$

where $0 \leq \alpha \leq 1$ is a relaxation constant.

The orthogonality and low number of modes is mostly a sufficient constraint to provide a stable solution, however in some cases additional constraints mentioned in Section 2.4.1 can be beneficial at least for the initial convergence. Also, if the data quality is sufficient it can be beneficial to use the standard single probe ptychography for the initial convergence and OPRP only for final refinement.

Implementation of the OPR overlap constraint is straightforward for both the most commonly used methods:

1. ePIE:

$$\hat{P}(\mathbf{x})_j^{(n+1)} = P(\mathbf{x})_j^{(n)} + \alpha_P \frac{O^*(\mathbf{x} + \mathbf{x_j})}{\|O\|^2_{\max}} \Delta\psi_j(\mathbf{x}) \tag{2.138}$$

2. Difference maps:

$$\hat{P}(\mathbf{x})_j^{(n+1)} = (1-\alpha_P)P(\mathbf{x})_j^{(n)} + \alpha_P \frac{O^*(\mathbf{x}+\mathbf{x}_j)\hat{\psi}_j(\mathbf{x})}{|O(\mathbf{x}+\mathbf{x}_j)|^2 + \varepsilon} \tag{2.139}$$

Note that the relaxed version of the real space projection must be used for the Difference maps methods otherwise the method may diverge.

The orthogonal probe modes (eigen-probes) are applied in the following simple procedure that is also shown in Fig. 2.29:

1. A separate probe is applied for each scanning position using the (Eq. 2.138.
2. The independent probes reconstructions are linked together by the truncated SVD (Eq. 2.137)
3. The low dimensional projected probes are used in the following iteration to calculate the exit wave needed for the modulus constraint
4. Iterations 1-3 are repeated till convergence

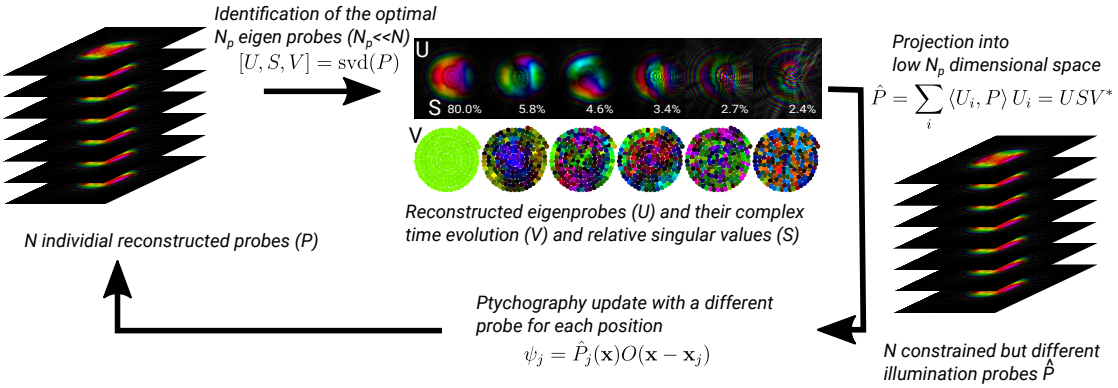

FIGURE 2.29: A simple schema showing the working principle of the OPR method. The singular value decomposition and low dimensional filtering is applied only once per iteration on all the independent probes together.

A potential drawback of the OPR method is usually slower convergence that is caused by slower information propagation compared to the ePIE-like solvers. The convergence can be improved if the variable step method presented in Section 2.5.2 is used. Experimental tests of the OPR method are shown in Section 3.2.

The OPR method is able to include wavefront shape and intensity variations caused by beam or optics movements into the ptychography. Additionally, it

is able to correct small sample positions errors and small variations of the sample to probe distance as well. Note that thanks to the symmetry of the probe and the sample, the position errors of the sample are equivalent to probe shift. If the probe shift is sufficiently small or the probe is smooth, it can be described by a low number of orthogonal modes and OPR method will automatically include it. However, larger shifts cannot be easily decomposed and can lead to unnecessarily high number of the eigenprobes as it is shown in Fig. 2.30.

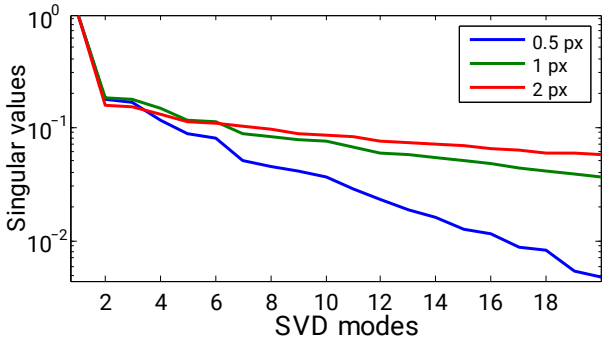

FIGURE 2.30: A numerical simulation shows the decrease of the singular values with the higher mode numbers for a randomly shifted illumination probe shown in Fig. 2.19. The plot shows that only small (sub-pixel) shifts can be decomposed into a low number of eigen probes.

2.7.1 Sparse OPR method

We propose a simple but yet robust method for the sample position correction based on OPRP. Our method is using the fact that if a certain probe has a wrong position then the decomposition

$$\hat{P}_i = \sum_j \langle P_i, U_j \rangle U_j \tag{2.140}$$

will be inefficient. If there exists an integer shift of each probe P_i that results in a more efficient description of P_i, then the probe should be shifted. We measure the "efficiency" by a simple $-l^p$ sparsity measure [121]

$$l_i^p = \left(\sum_j^N |c_{ij}|^p \right)^{1/p} \quad , \quad 0 < p \leq 1 \quad , \tag{2.141}$$

where l_i^p denotes sparsity of i-th probe and $c_{ij} = \langle P_i, U_j \rangle$. Parameter $p = 1$ and N is the number of the eigen probes. This sparsity definition penalizes use of

the weaker modes that need to be suppressed. We have tested that this method gives consistent results in simulations and with real datasets. The requirement of sparsity is simply expressing the fact that only a single eigen probe is sufficient to describe an ideal errorless ptychography system. Also, position errors lead to much lower sparsity of the modes compared to only errors caused by movements of a relatively smooth beam.

The procedure to find the optimal shift is following:

1. Shift the eigen probes U_j to distances of -1,0,1 pixels in vertical and horizontal direction. It results in 4 $U_j^{(n)}$ shifted eigen probes and one original $U_j^{(0)}$
2. Calculate projection value

$$c_{ij}^{(n)} = \left\langle P_i, U_j^{(n)} \right\rangle = \sum_k \left\langle U_k, U_j^{(n)} \right\rangle S_k V_i^* \tag{2.142}$$

   for each probe with each of the eigen probes $U_j^{(n)}$.
3. Calculate sparsity l_i^p score for each position i and shift n
4. Perform shift of every probe P_i in order to maximize the sparsity l_i^p

Note that the steps 1-3 can be written as simple matrix operations without need to operate directly with the full expression of the probes P, which allows very efficient GPU implementation with linear computational complexity. Also it is not necessary to deal with the sub-pixel shifts, because they are recovered by the OPRP method automatically.

On contrary to other position correction methods [25, 90, 91], our method does not need additional tuning parameters and it provides robust results even for significantly limited illumination stability. On the other hand, we have observed that it may be beneficial to combine the proposed position correction method with the gradient descent based method introduced in Section 2.6.2 in order to speed up the convergence.

2.7.2 Incoherent modes relaxation

Until now, only relaxation of the main "coherent" probe mode was discussed and relaxation of the higher incoherent modes [16] was neglected. The first mode carries most of the information, therefore usually it is a valid approximation to assume that the higher order incoherent modes are constant. However, this may be incorrect if the illumination intensity is fluctuating significantly or if the decoherence effects are changing during the scan. This can be for example if the sample changing speed during the on-fly ptychography [100] or if the beam movements over the optics are faster than the exposure time. Using the presented OPR method for each of the incoherent modes would be highly inefficient because the number of the incoherent modes can be high and it would be very costly to calculate SVD of each mode.

Here, we present a simplified version of the OPR idea that can be easily used to include variable coherence properties of the measured diffraction pattern or it can just serve as an easy intensity correction method if applied to the first "coherent" mode. Instead of projection into a lower dimensional space defined by the SVD method, it is possible to use only a projection of the i-th incoherent mode estimation in position j (P_{ij}) into the space defined by the average probe mode $P_{\mathrm{i,\ avg}}$ over the scanning positions

$$\begin{aligned} P_{\mathrm{i,\ avg}} &= \sum_j P_{ij}/N_{\mathrm{pos}} \\ \hat{P}_{ij} &= \langle P_{ij}, P_{\mathrm{i,\ avg}} \rangle \, P_{\mathrm{i,\ avg}} \quad , \end{aligned} \tag{2.143}$$

where $\hat{P}_{ij}$ denotes projected mode estimation P_{ij} . The main advantage is that once the projections $\langle P_i, P_{\mathrm{avg}} \rangle$ are known, the rest of the ptychography code can be identical and there is no need to keep in memory reconstructions of all the probe estimations P_{ij}. This is crucial if the ptychography method is memory limited for example in the case of GPU implementation. The modified reconstruction procedure is performed as follows:

1. Relative weights of the probes w_{ij} are initialized as 1, $\hat{P}_i$ are initialized as the single probe ptychography estimates P_i

2. Each incoherent illumination mode $\hat{P}_i$ is multiplied by its relative weight that is different for every position j and mode order i

$$P_{ij} = w_{ij}\hat{P}_i$$

3. Each of the modes P_{ij} is updated by a ptychography algorithm

4. Update the weights that are calculated as

$$P_{\text{i, avg}} = \sum_j P_{ij}/N_{\text{pos}}$$
$$w_{ij} = \frac{\langle P_{ij}, P_{\text{i, avg}} \rangle}{\|P_{\text{i, avg}}\|^2}$$

5. Update the shared incoherent illumination mode $\hat{P}_i$

$$\hat{P}_i = \frac{\sum_j P_{ij}}{\sum_j w_{ij}}$$

6. Repeat steps 2-5

Each update is performed over a small group of N positions in parallel. This allows serialization from the ePIE method and the advantage of faster processing speed of GPU as well (see Fig. 2.21) without additional memory requirements.

2.8 Conclusion and outlook

The most fundamental *a priori* knowledge used in ptychography is the overlap constraint and the measured data (modulus constraint), however as we have mentioned in this chapter and we will show in the following experimental chapters as well, both these constraints are commonly invalid due to systematic errors. In Section 5, we will show that despite the low flux in the EUV laboratory setups, the reconstruction quality of the weakly scattering samples is often rather limited by systematic errors than by low photon flux.

In this chapter, some of the most important published methods allowing relaxation of the ptychography constraints were summarized. Additionally, we have introduced several novel methods for the real space relaxation: anisotropy correction

(Sec. 2.6.2), orthogonal probe relaxation with sparsity based position correction Sec. 2.7 and a new gradient based position correction solver (Sec. 2.6.2). As we will show in the next chapters, the OPR method can correct for variable illumination wavefront and variable coherence properties. We have also shown a simple method to include known noise distribution into the standard ePIE based algorithms (Sec. 2.4.1), optimal background offset estimation using a gradient based method (Sec. 2.6.2) or how to include imperfect optics behind the sample into the reconstruction (Sec. 2.6.2). In order to improve the convergence speed, we have presented a sequential gradient descent solver with variable step length (Sec. 2.5.2) and a better update scheme for the ePIE-like methods that is based on the Levenberg-Marquardt (LM) algorithm known also as the damped least-squares method (Sec. 2.5). This method allows us to calculate larger number of positions in parallel similarly to the pPIE method and improve the convergence speed as well.

The aim of this chapter was mainly to show that the reconstruction quality can be significantly increased if the hard ptychography constraints are relaxed and additional degrees of freedom that allow to describe more precisely the real imaging system are added. Furthermore, if additional valid *a priori* knowledge is provided, it can further improve the convergence and the reconstruction quality.

Chapter 3

Experimental verification of ptychographic corrections using visible light

Ptychography method is most commonly used in the short wavelength regions, where the quality of the images is limited by optics or optics is expensive and inefficient. Lensless ptychography can be also used with the visible light for high quality phase imaging or 3D microscopy [25, 45, 122], however due to availability of high quality optics, it is loosing its main advantage. Moreover, there are more appropriate methods for visible light CDI such as the Fourier ptychography [2, 3, 5, 26] or spatial light interference microscopy (SLIM) [123] that are using the advantage of the availability of good quality optics to get fast reconstructions of both phase and attenuation.

Despite this, visible light ptychography has many advantages as a simple testing setup and many novel ideas were first proposed and tested with visible light [45, 52, 81, 82] and then later implemented in the short wavelength regions. In our case, the intention to build a visible light ptychography microscope was however mainly motivated by applications of the nonlinear ptychography that is explained in the next chapter.

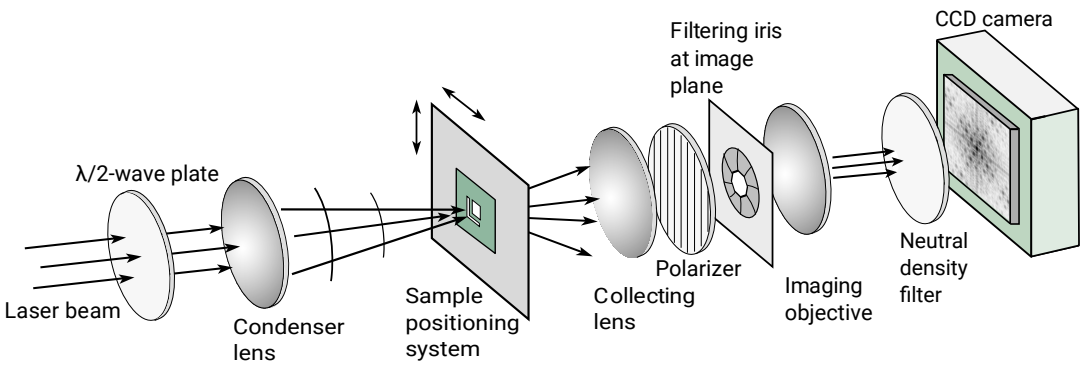

FIGURE 3.1: A simplified scheme of our visible light ptychography setup. The condenser lens is defining the wavefront curvature on the sample. The illumination probe is cropped by an iris placed in the imaging plane of the collecting lens. The imaging objective allows us to easily select imaging regime from real-space, near-field to the far-field regime.

3.1 Experimental setup

Our visible light setup was build mainly with focus on flexibility and easy option to test different arrangements of ptychography. A simplified schema is shown in Fig. 3.1. The initial polarization of the laser beam can be set by a half-wave plate and the beam is focused on the sample by a condenser lens. Selection of focal length of the condenser lens allows us to tune the wavefront curvature of the illumination probe. The sample is moved by a 3 axis nanoprecision piezoelectric actuator Smaract SLC-1730 with 26 mm travel range and roughly 100 nm position precision (as shown in Appendix B). The diffracted light was collected by long working distance microscopy lens from Olympus with NA=0.4, 20×. Finally, the beam was imaged by another system of lenses onto a CCD chip placed close to the Fourier plane. The used PIXIS 1024 CCD camera can provide up to 16 bit dynamic range at 100 kHz pixel-readout speed (i.e. 0.1 frames/s). However, for our linear ptychography experiments, we were using 2 MHz pixel-readout speed, therefore the dynamic range was limited roughly to 12-bit by the readout noise.

The pixel size (and achievable resolution) in the presented reconstructions determined by the front collecting lens with NA=0.4 to

$$\Delta x = \frac{\lambda}{2\mathrm{NA}} = 1.25\lambda \quad . \tag{3.1}$$

The dynamic range of the PIXIS CCD camera was extended by collecting multiple different exposures, each subsequent exposure containing only the over-saturated

region of the previous one. The final high dynamic range pattern $I(\mathbf{k})$ is equal to

$$I(\mathbf{k}) = \frac{\sum_i^N W_i(\mathbf{k}) I_i(\mathbf{k})}{\sum_i^N W_i(\mathbf{k}) t_i} t_N \quad , \tag{3.2}$$

where t_i is the i-th exposure time, $I_i(\mathbf{k})$ are background subtracted data. Finally, W_i denotes a weighting factor proportional to the inverse noise $1/\sigma$ in the measured regions and to zero elsewhere. The maximal dynamic range was extended from 16-bits to roughly 27-bits (10^8 counts) for the PIXIS camera with average readout noise 2 counts. Cooling of the camera was always set to -70°C in order to suppress dark noise. All noise properties such as the dark noise (dark current and background light), the readout noise and bias were automatically measured before each ptychography experiment.

3.1.1 Virtual probe

On contrary to the common ptychography setup used in EUV and X-ray, visible light allows us to use high quality optics behind the sample and therefore the pinhole can be placed behind the sample without removing any high spatial frequencies. In our case, an iris is placed ≈ 180 mm behind the collecting lens at the imaging plane, where there is a magnified virtual image of the sample (see Fig. 3.1). The virtual probe allows fine tuning of the illumination probe diameter in order to achieve the optimal oversampling and uniform illumination over the virtual illumination probe. Magnification of the lens used was 20×, therefore an iris with opening diameter 3 mm corresponded to 150 μm illumination probe on the sample. Moreover, a weak diffuser composed from ground glass was mounted into the iris in order to improve high spatial frequency content in the ptychography probe because sharper structure in the illumination probe helps to get a better reconstruction [45, 124]. Moreover, the diffuser reduces the dynamic range of the diffraction pattern and thus allows faster data acquisition and helps to avoid reflections between the used optics. Examples of different illumination probes from our setup are shown in Fig. 3.2.

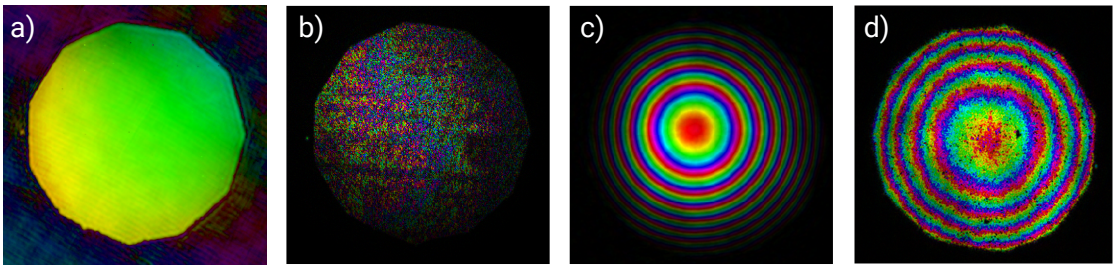

FIGURE 3.2: Examples of different probes in our visible light setup. Probe (a) is flat illumination wavefront cropped by iris, (b) is also flat illumination wavefront but with a weak diffuser inside the iris. Probe (c) is an example of curved illumination, it is again cropped by the iris but only loosely because the beam was small. Finally, the probe (d) was created by curved illumination (NA≈0.1) in combination with the iris and the diffuser in the imaging plane and it was used for the anisotropy experiment (Section 3.3).

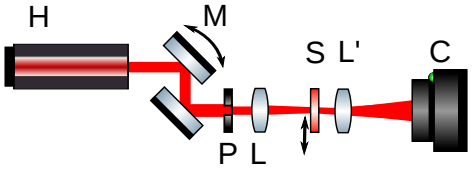

FIGURE 3.3: A schema of the optical setup for testing the OPRP method. The HeNe laser beam is deflected by a motorized mirror M, cropped by a pinhole P and projected on a sample S using lens L. The scattered light is collected by a microscope lens L' and demagnified on to a camera C that is placed near to the Fourier plane.

3.2 Orthogonal probe relaxation ptychography (OPRP)

In this section, we will show tests of our OPRP method that was introduced in Section 2.7. We proposed the OPRP method as a way to deal with intensity and wavefront shape variations of the illumination beam between the scanning positions. If the changes are faster than the exposure time, then it causes blurring and decoherence effects. The wavefront variation issue is not too common for the visible light experiments because the available laser sources are very stable, although as we will show here, our method was able to correct for even small beam movements caused by thermal bending of the optical table during the ptychography scan with HeNe laser.

In order to simulate the effects of unstable illumination, a testing setup (Fig. 3.3) with the beam deflected by an active mirror was built. The sample was illuminated by a HeNe laser at 633 nm wavelength. The beam was deflected by the mirror M

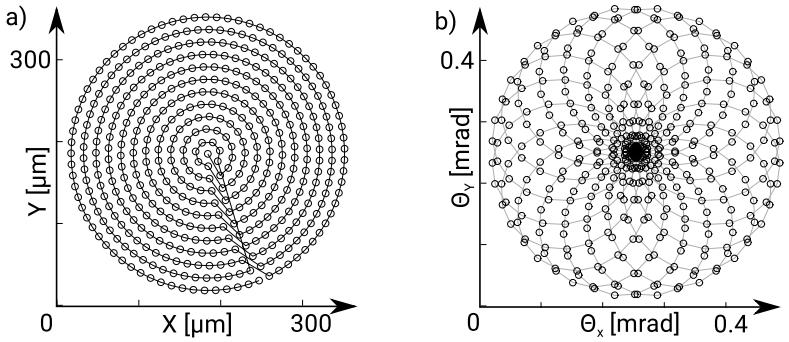

FIGURE 3.4: Figure a) shows the selected double spiral scanning path with all 463 positions. Figure b) shows the desired beam deflection angle.

by up to ±0.4 mrad. The beam displacement was comparable with the beam size at the iris position. The desired beam movements are shown in Fig. 3.4, however the actuator used had no position feedback, therefore the real positions may have been different due to hysteresis of the actuator. Note that knowledge of the beam position was not used anywhere in the reconstruction process. The deflection path was chosen to roughly simulate random movements with the Lorentz distribution. The path was intentionally selected to be a continuous smooth curve (see Fig. 3.4) in order to easily see the smooth evolution in the final reconstruction. However, the most easy to reconstruct distribution would be fully random because long term drifts are more difficult to sense and recognize. In order to deal with it, we propose to use double spiral scanning pattern (Fig. 3.4). This allows us to make the adjacent scanning positions more distanced in time.

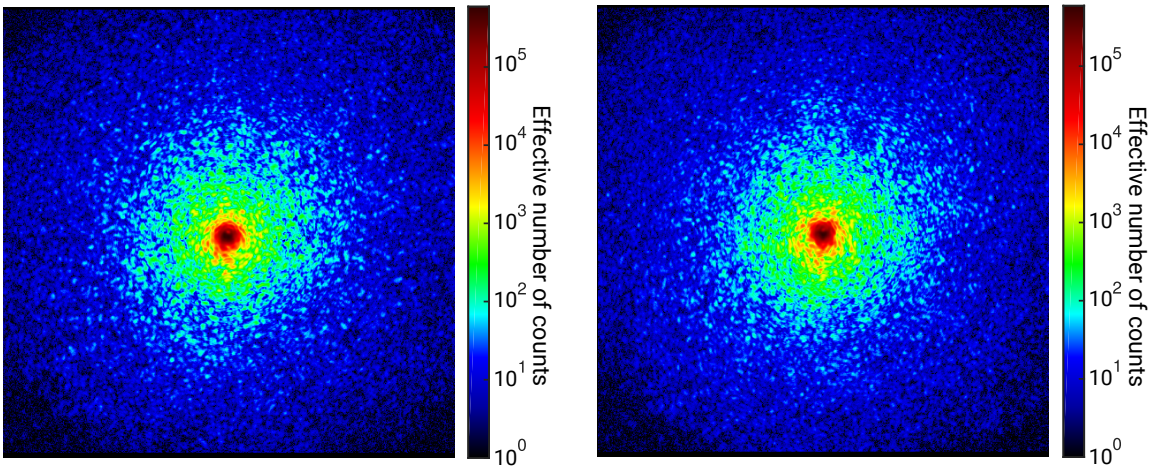

FIGURE 3.5: Two examples of the measured diffraction pattern used in this experiment. The maximal measured scattering angle is limited by the acceptance NA of the front collecting lens (see setup in Fig. 3.3). The exposure time was in range between 20 ms-1 s.

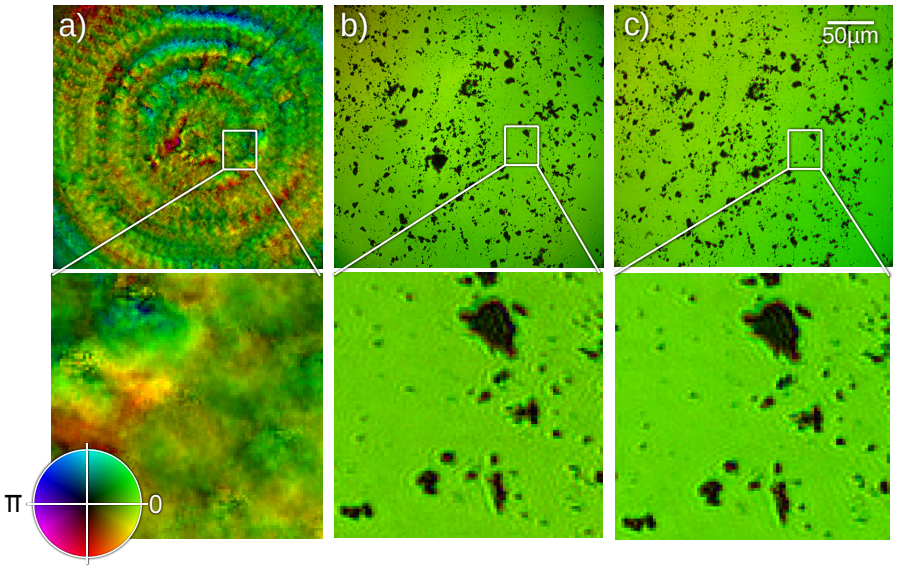

FIGURE 3.6: Reconstructions of graphite particles shown in complex colour scale. Image (a) shows a reconstruction of the actively disturbed dataset with the standard ePIE method. Image (b) shows the same dataset reconstructed using the OPRP method. Finally, the last image (c) shows ePIE reconstruction with steady illumination. The second row shows a zoomed region of the reconstruction. Image adapted from [6].

Two datasets were collected, one dataset with steady illumination and the second with an actively disturbed beam. Each ptychography scan consisted of 463 positions with 30 μm step size with illumination probe size about 140 μm, in order to get linear overlap of 80%. Carbon dust particles deposited on a glass slide were used as a simple but aperiodic test object to compare the reconstruction quality. An example of the diffraction pattern is shown in Fig. 3.5.

Figure 3.6 shows reconstructions with the actively disturbed beam (a,b) and the steady beam (c) and the eigen-probes are shown in Fig. 3.8. The sample reconstruction (a) was performed by the standard ePIE method, results from DM method were similar. Reconstruction (b) was based on the same dataset as (a) but it was performed by the OPRP method and finally, reconstruction (c) is based on an ideal dataset with steady illumination using the ePIE algorithm. The magnified region in the second row shows that the results with and without the beam movements are comparable.The small fringes around the objects are not related to the reconstruction process, we have observed them even in direct imaging without any processing. Therefore, we assume that they are related to astigmatism of our microscope. In order to fully assess the reconstruction quality, we have applied Fourier ring correlation (FRC) introduced in Section 2.2.5. The FRC comparison

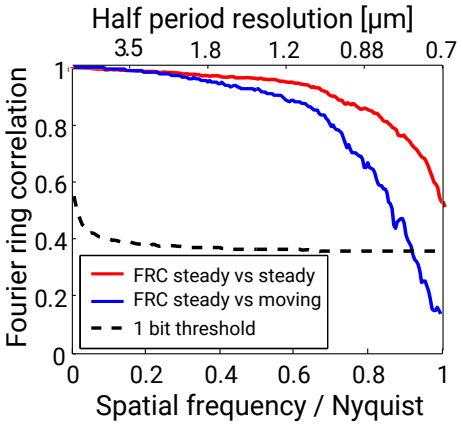

FIGURE 3.7: Fourier ring correlation (FRC) between two reconstructions from independent datasets with steady illumination (red line) compared to the FRC between the steady reconstruction and one reconstruction with disturbed illumination provided by the OPRP method (blue line). The black dashed line denotes the 1-bit threshold (Eq. 2.58).

of reconstructions is shown in Fig. 3.7. The red line is comparison of two different steady datasets and the blue line is a steady dataset versus the disturbed dataset. The difference was increased between the actively disturbed dataset and the steady one, but the beam movements were intentionally selected large in order to test the limits of the OPRP method.

Besides the optimal object reconstruction, the OPRP method provides also information about the illumination. OPRP method provides orthonormal basis U of the reconstructed probe. These orthonormal basis can give insight about the illumination changes, however note that each of the bases separately has generally no meaning, it is necessary to use the full time evolution of the illumination P given by

$$P = USV^* \quad . \tag{3.3}$$

The orthonormal basis U, singular values S and complex time evolution V shown in Fig. 3.8 are however useful for visualization.

The first two rows are showing orthonormal basis U for the steady dataset. It was unexpected to see some vectors for the steady dataset, however the spatial evolution V implies that the additional modes were needed to explain long term drift probably caused by temperature changes in lab. The last two vectors were not able to converge to any consistent solution, however their relative power is close to zero. The second two rows contain a decomposition of the reconstructed probe P from the actively disturbed dataset. The complex spatial evolution V

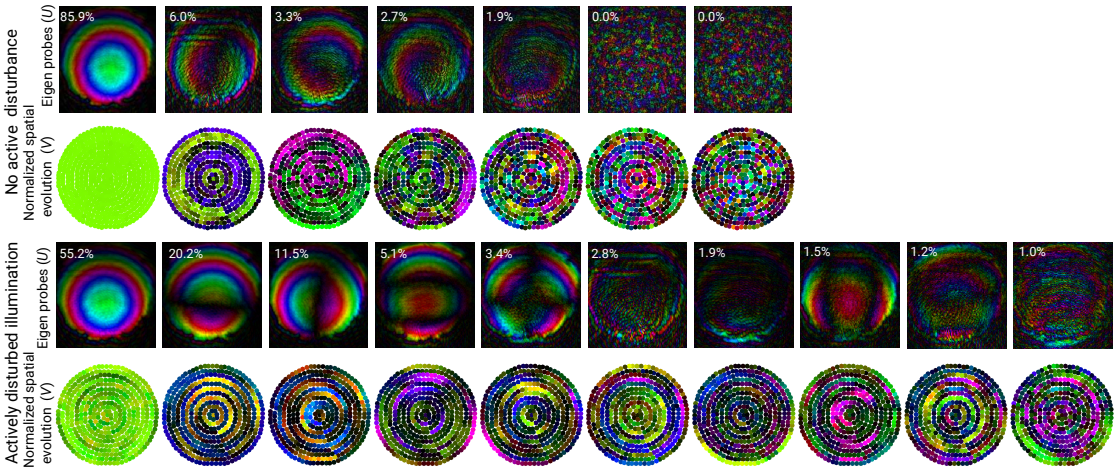

FIGURE 3.8: Orthonormal basis shown in the complex colour scale with the relative singular values $S_i / \sum S_i$. The first two rows contain basis produced by the OPRP method applied on the steady dataset and the second two rows are showing bases for the actively disturbed dataset.

is plotted for each scanning position in the complex colour scale in the last row. This was done intentionally to show that there are no features of the sample picked up by the decomposition. The evolution V shows that the modes were changing smoothly along the double spiral scanning path (see Fig. 3.4) as it was expected. The discontinuity of V values in the bottom right quadrant of the scanned path is caused by the double spiral shape of the scanning pattern. The presented datasets are available at DOI:10.5258/SOTON/381931. The animation of the actively moved illumination probe is available at https://www.osapublishing.org/oe/viewmedia.cfm?uri=oe-24-8-8360.

3.3 Anisotropic ptychography

The idea of anisotropy ptychography was introduced in Section 2.6.2. The method was tested on a thin anisotropic crystal composed of NMBA (4-nitro-4'-methylbenzilidene aniline). The refractive index of the NMBA crystal can vary from 1.5 to 2 at 800 nm [125] depending on the crystal orientation. Our crystal was created by fast cooling of NMBA between two glass slides in order to guarantee thickness below 10 μm and flatness of the surface.

The sample was illuminated by a lens with NA≈0.2 and the light was collected by lens with NA=0.4. Each of the ptychography scans consisted of 1600 diffraction patterns with 0.5 s exposure time per scanning position. Figure 3.9 shows

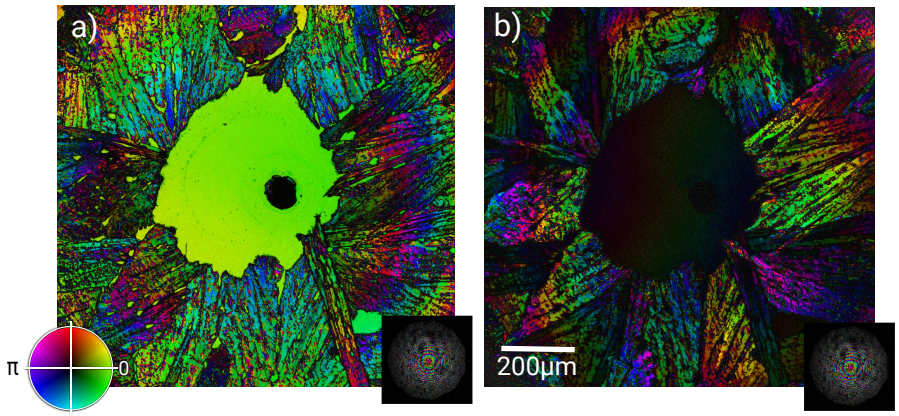

FIGURE 3.9: Reconstructions of the complex transmission function of the NMBA crystal sample. The image (a) shows a reconstruction with a polariser behind the sample parallel with the beam polarization and the image (b) shows the crossed polarization. The central region of the sample is glass.

reconstruction of the "isotropic" mode that is practically identical to the standard ptychography reconstruction. The reconstructions are clearly showing multiple regions of the crystal with different behaviour, which is caused by different orientation of the crystal axis. The central region is glass and the sample thickness is slightly changing in the radial direction and causing phase ramp up.

The results of the anisotropic correction are shown in Fig. 3.10. The anisotropy clearly shows different crystalline properties of the sample. The central glass has approximately zero anisotropy. The rest of the crystal exhibits at least weak anisotropic behaviour. The absolute calibration of the colourbar is not trivial and the correction factor values are not unique. We have tested on numerical simulations that any constant offset is self-adjusted to keep the average correction close to zero. Further, the low spatial frequencies reconstruction is not very reliable. Finally, the sample was already relatively strongly eroded at the time of the experiment and therefore the surface smoothness could have been a limiting factor. Despite this, the method is still able to provide interesting information that could be used to identify orientation of the crystals in combination with the phase measurement from standard linear ptychography. Moreover, this method could be used in combination with the nonlinear ptychography mentioned in the next chapter.

Figure 3.11 shows a map of the different crystalline regions on the sample. The sample was imaged with different angles between the input beam polarization and the analysing polariser and the intensity for each pixel was measured in realspace.

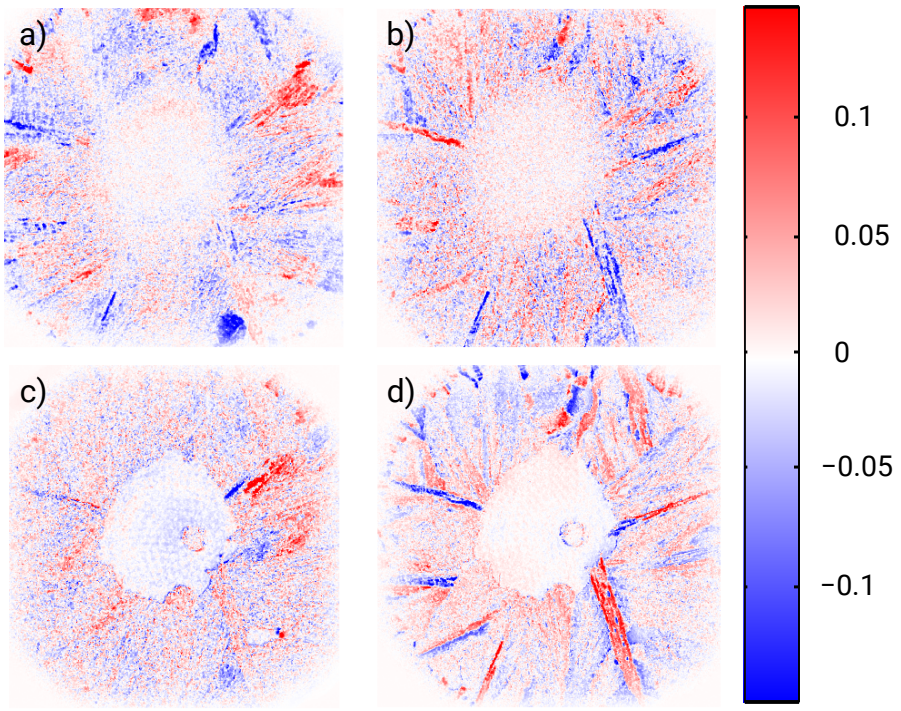

FIGURE 3.10: Reconstruction of the anisotropic phase correction terms. Image (a,b) shows vertical / horizontal correction factors for parallel polarization and (c,d) are vertical / horizontal correction factors for perpendicular polarization case.

Note that no CDI was used because only intensity was needed. Measured data were first filtered by SVD and downsampled. Then different regions (crystal orientations) were automatically recognized and separated using the K-means clustering algorithm (see e.g. Ref. [126]).

This polarization based method is sensitive to rotation of the crystal in the plane perpendicular to the beam, while the anisotropy ptychography can sense the rotation and properties change in the direction along the beam.

3.4 Conclusion

In this chapter, we have tested two methods that provide additional relaxation of the real space ptychography constraint. The orthogonal probe relaxation (OPR) allows significant variations in the probe intensity and wavefront and it provides reconstruction of time-evolution of the illumination probe. We expect that the main application of the OPR method will be for sources for unstable sources with characteristic time of the movements below the exposure time. This is nowadays

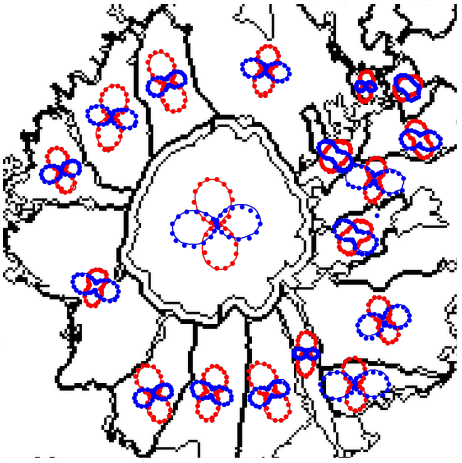

FIGURE 3.11: Polarization map of the NMBA crystal. Blue dots show intensity measurement for vertically polarised illumination and red dots for horizontally polarised beam. The measured points are shown in polar coordinates in dependence on the analysing polariser angle and fitted by theoretical curve. Properties of the directly passing beam can be seen in the central region.

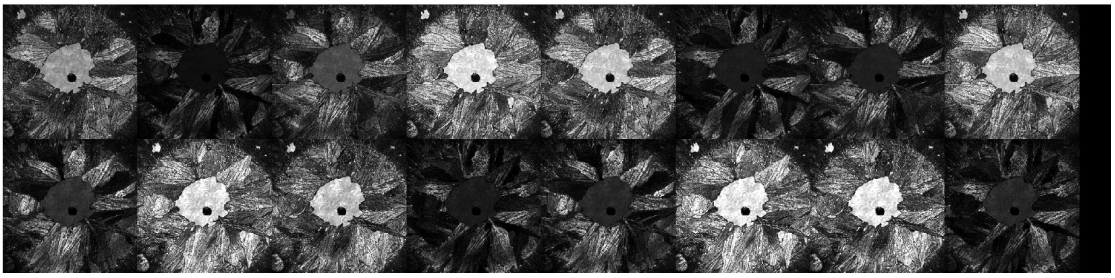

FIGURE 3.12: Examples of the transmission through the crystal with different angles of analysing polariser (columns) [0°,40°,. . .,320°] and different polarization angle of illumination beam (rows) [0°,90°]. These data were used after preprocessing for the clustering algorithm for automatic separation of the crystals with different properties.

a common issue for high harmonic generation (HHG), SXR lasers and free electron lasers. However, with further improvements in the coherence of synchrotron sources and therefore shorted exposure time per position, OPRP can find application even in X-ray synchrotron ptychography. We will show some more OPR results in the Chapter 5 with our HHG source. The second tested method allows us to include anisotropy into the ptychography method. It can be understood as the object space alternative to the OPR method and it allows us to recover local derivative of the refractive index in dependence on the illumination angle.

Chapter 4

Second Harmonic Generation (SHG) imaging experiment

Harmonic generation processes have wide applications in physics, biology and chemistry. Harmonic processes can be used for modification of high intensity light beams properties using frequency mixing, splitting or multiplication. Or it can be used for probing matter using harmonic generation based microscopy techniques such as the second harmonic microscopy [127–130] or the coherent anti-Stokes Raman spectrocopy (CARS) [131]. The interesting property of the harmonic processes is that they preserve the coherence of the generating beam and therefore the harmonic light, as we will show in this chapter, can be imaged using the CDI methods as well. CDI allows us to reconstruct the phase information that is lost if the standard confocal microscopy is used. Furthermore, the ptychography method allows us to separate the harmonic illumination probe that is created when the beam interacts with the sample from the SHG field corresponding to the sample properties.

In this chapter, our nonlinear coherent diffraction microscope will be introduced including some experimental results from nonlinear SHG active crystals. CDI SHG microscope could provide an alternative way to measure properties of the nonlinear nanocrystals such as attenuation, dielectric tensors, nonlinear tensors and orientation of the main axes without a need of difficult sample preparation.

We have tested SHG CDI on many samples, but only results from an organic crystal 4-nitro-4'-methylbenzilidene aniline (NMBA) and a periodically poled lithium

niobate (PPLN) crystal will be shown in this chapter. Moreover, both these samples are known for their good SHG properties due to their highly asymmetrical structure. NMBA has triclinic and PPLN has trigonal crystal structure, therefore the refractive indexes along the main axes are unequal. Properties of both samples are well known [125, 132, 133], therefore they were selected as good samples to test our imaging method.

4.1 Theory of light propagation through anisotropic medium

Propagation of light in an anisotropic sample or an anisotropic harmonic generation sample is significantly more complex than in the simple case of isotropic thin material that is usually considered in CDI experiments. The fundamental theory of light propagation through anisotropic media described in Refs. [21, 134, 135] is introduced in the following sections.

4.1.1 Basic properties of anisotropic media

Any plane wave propagating through anisotropic medium can be decomposed into a linear combination of two base vectors. Let $\mathbf{e}_1$, $\mathbf{e}_2$ be orthonormal vectors in direction perpendicular to the propagation direction with the corresponding propagation wave numbers k_1, k_2. The propagating wave can be written as

$$\mathbf{E}(\mathbf{r}) = A\mathbf{e}_1 \exp(\mathrm{i}k_1\mathbf{s}\cdot\mathbf{r}) + B\mathbf{e}_2 \exp(\mathrm{i}k_2\mathbf{s}\cdot\mathbf{r}) \quad , \tag{4.1}$$

where A, B are projection constants and $\mathbf{s}\cdot\mathbf{r}$ is the distance along the direction of propagation $\mathbf{s}$. The propagation wave numbers k are generally complex valued and thus already include attenuation σ. Values of $k_1, k_2, A, B, \mathbf{e}_1, \mathbf{e}_2$ can be calculated if the direction of propagation $\mathbf{s}$ with respect to the optical axis and the dielectric tensor ε_{ij} in its diagonal form are known. Therefore, in order to fully describe linear properties of the anisotropic medium, the diagonal of the complex dielectric tensor $\varepsilon_{1,2,3} + \mathrm{i}\sigma_{1,2,3}$ and rotation of the crystal axis with respect to the beam need to be known. In the most general case, this results in 9 independent variables. In order to fully characterize nonlinear anisotropic properties of a crystal, it is

necessary to know complex dielectric tensor for fundamental and harmonic wavelength (18 variables) and moreover the nonlinear optical tensor. This tensor can have up to 18 variables in the case of triclinic crystalline system and up to 6 for trigonal system [134, 135]. This simple estimation is just showing the complexity of material properties estimation in the case of anisotropic materials compared to simple isotropic case. It is very challenging to gain all the material properties without additional knowledge and approximations.

It is possible to find additional constraints on the anisotropic tensor ε. Using Poyinting's theorem for power flux conservation [21], it can be derived that the complex dielectric tensor is Hermitian

$$\varepsilon_{ij} = \varepsilon_{ji}^{*} \quad . \tag{4.2}$$

This relation is valid only if the contributions from currents can be neglected. Otherwise, ε_{ij} can be asymmetric (e.g. plasma in magnetic field).

Several important conclusions follow from the symmetry of the dielectric tensor:

- tensor ε_{ij} can be decomposed into the form $\mathbb{X}^{-1}\Lambda\mathbb{X}$ where Λ is a diagonal matrix of eigen numbers and $\mathbb{X}$ are eigen vectors
- eigen vectors $\mathbb{X}$ are orthogonal, therefore the dielectric principal coordinate axes are orthogonal
- eigen numbers ε_{ij} are positive.

The diamagnetic tensor μ can be approximated in the optical range by its vacuum value μ_0 and results in refractive index $n \approx \sqrt{\varepsilon/\varepsilon_0} \equiv \sqrt{\varepsilon_r}$. This allows us to constrain the eigen numbers ε_r values if we know maximum and minimum refractive index of many randomly oriented crystals

$$n_{\min}^2 \lesssim \varepsilon_r \lesssim n_{\max}^2 \quad . \tag{4.3}$$

The expected range of the refractive indexes for our NMBA sample extends from 1.5 to 2 at 800 nm and up to 2.5 at 400 nm [125]. The second test sample PPLN has refractive index from 2.21 at 800 nm to 2.43 at 400 nm [136]. If the dielectric tensor ε is known, the propagating wave vectors $\mathbf{k}$ can be obtained from the Maxwell equations

$$\mathbf{k} \times (\mathbf{k} \times \mathbf{E}) + \omega^2 \mu \varepsilon \mathbf{E} = 0 \quad . \tag{4.4}$$

This is a linear homogeneous system of equations and it has a non-zero solution $\mathbf{E}$ only if determinant of the system is zero. Using the relation $\mathbf{k} = \frac{\omega}{c} n\mathbf{s}$ leads to a transcendent equation

$$\frac{s_x^2}{n^2 - n_x^2} + \frac{s_y^2}{n^2 - n_y^2} + \frac{s_z^2}{n^2 - n_z^2} = 1/n^2 \quad , \tag{4.5}$$

where $n_i \equiv \sqrt{\varepsilon_{ii}/\varepsilon_0}$ and n^2 is the searched solution. There are two independent solutions of n^2 that can be expressed analytically but not in a compact form. The eigen vectors can be calculated numerically if the solution n of the Eq. (4.5) is inserted to the Eq. (4.4). Then the null space of the matrix (4.4) will be eigen vectors of polarizations $\mathbf{E}_x$. However, the solution can be also expressed analytically

$$\mathbf{E}_x = \begin{pmatrix} \frac{s_x}{n^2 - n_x^2} \\ \frac{s_y}{n^2 - n_y^2} \\ \frac{s_z}{n^2 - n_z^2} \end{pmatrix} \quad . \tag{4.6}$$

Each of the two solutions n of the Eq. (4.5) provides a different eigen vector $\mathbf{E}_{xi}$. Note that the eigen vectors $\mathbf{E}_{xi}$ are not perpendicular for a general biaxial crystal (e.g. the NMBA crystal)

$$(\mathbf{E}_{x1} \cdot \mathbf{E}_{x2}) \neq 0 \quad . \tag{4.7}$$

The vectors will be perpendicular only in the case of uniaxial crystals. Moreover, $\mathbf{E}_{xi}$ can be approximated to be perpendicular to the propagation direction $\mathbf{s}$ only for weakly birefringent material. The vectors $\mathbf{E}_{xi}$ are linearly polarized in the case of non-absorbing medium (n is real) and therefore the crystal is not optically active.

Further, it is possible to calculate the energy propagation vector $\mathbf{t}_1, \mathbf{t}_2$ (ray vectors). In the general case of biaxial crystals, there will be two directions leading to birefringence

$$\mathbf{t}_i = \mathbf{E}_{xi} \times (\mathbf{s} \times \mathbf{E}_{xi}) \quad , \tag{4.8}$$

It can be easily seen that if the eigen vectors $\mathbf{E}_{xi}$ are perpendicular to the direction of propagation $\mathbf{s}$ (i.e. weakly birefringence materials), the ray vectors $\mathbf{t}_i$ will be parallel with $\mathbf{s}$.

The corresponding displacement vectors $\mathbf{D}_{xi}$ can be simply calculated as

$$\mathbf{D}_x = \boldsymbol{\varepsilon}\mathbf{E}_x \quad , \tag{4.9}$$

where $\boldsymbol{\varepsilon}$ is the diagonalized dielectric tensor. It can be also obtained from the equation

$$\mathbf{D} = -\frac{n^2}{c^2\mu}\mathbf{s} \times (\mathbf{s} \times \mathbf{E}) \quad , \tag{4.10}$$

which immediately leads to $\mathbf{s} \cdot \mathbf{D} = 0$. Further, it can be also derived that $\mathbf{D}_{xi}$ are orthogonal base vectors ($\mathbf{D}_{x1} \cdot \mathbf{D}_{x2} = 0$) [135].

4.1.2 Propagation inside of anisotropic medium

The electric field vector propagating inside of a crystal $\boldsymbol{\Psi}$ needs to be first rotated to the coordinates of the optical axis of the crystal

$$\boldsymbol{\Psi} = \mathbb{R}(\phi)\boldsymbol{\Psi}_0 \quad , \tag{4.11}$$

where $\mathbb{R}(\phi)$ is a rotation matrix in the plane perpendicular to the beam. Further, the plane wave represented by a column vector $\boldsymbol{\Psi}(z) = (E_x, H_x, E_y, -H_y)$ is projected to the basis vectors and propagated through the crystal using the Berreman's 4×4 propagation matrix method for one-dimensional optical medium [137]

$$\boldsymbol{\Psi}(z + h) = \mathbb{P}(h)\boldsymbol{\Psi}(z) \quad , \tag{4.12}$$

where the matrix $\mathbb{P}(h) = \exp(\mathrm{i}k_0 h\boldsymbol{\Delta}(z))$ is called propagating matrix over distance h and k_0 is the vacuum wave number. Distance h is the projection of the propagation direction $\mathbf{s}$ to position vector $\mathbf{r}$ (i.e. thickness of the crystal). Finally, $\boldsymbol{\Delta}(z)$ is 4×4 matrix that is a function of the material permittivity and permeability tensors. If the material can be considered homogeneous over the distance h and the anisotropy of the permeability tensor can be neglected, the propagation matrix $\mathbb{P}(h)$ can be expressed in a simple form

$$\mathbb{P}(h) = \mathbb{X}^T\boldsymbol{\Lambda}(h)\mathbb{X} \quad , \tag{4.13}$$

where $\boldsymbol{\Lambda}(h)$ denotes a 2×2 diagonal matrix with diagonal elements $\exp(\mathrm{i}k_0 h n_i)$, $\mathbb{X}$ is 3×2 orthogonal matrix in directions of the eigen vectors $\mathbf{E}_{xi}$ and n_i are the complex refractive indices along the optical axis.

Finally, the output field $\boldsymbol{\Psi}_{\mathrm{out}}$ can be simply expressed as

$$\boldsymbol{\Psi}_{\mathrm{out}} = \mathbb{R}(\phi)^T\mathbb{X}^T\boldsymbol{\Lambda}(h)\mathbb{X}\mathbb{R}(\phi)\boldsymbol{\Psi}_0 \quad . \tag{4.14}$$

In addition to the propagation effects caused by Eq. 4.14, effects on boundaries including reflections need to be considered. Reflections R need to be calculated for each eigen vector $\mathbf{E}_{xi}$ (and each n_i) separately from the Fresnel equation

$$R = \left(\frac{n_0 - n_1}{n_0 + n_1}\right)^2 \quad , \tag{4.15}$$

where n_0 is 1.5 for glass and 1 for the air interface while n_1 is the refractive index of the sample. Double reflections can be neglected and the maximum losses are expected to be 6-13% of the input power. Fig. 4.1 shows reconstructed amplitude of a HeNe laser ($\lambda = 633\,\text{nm}$) passing through our NMBA crystal sample. The intensity attenuation was between 10-20% for the large flat regions. The central region of the NMBA sample is pure glass with no crystal.

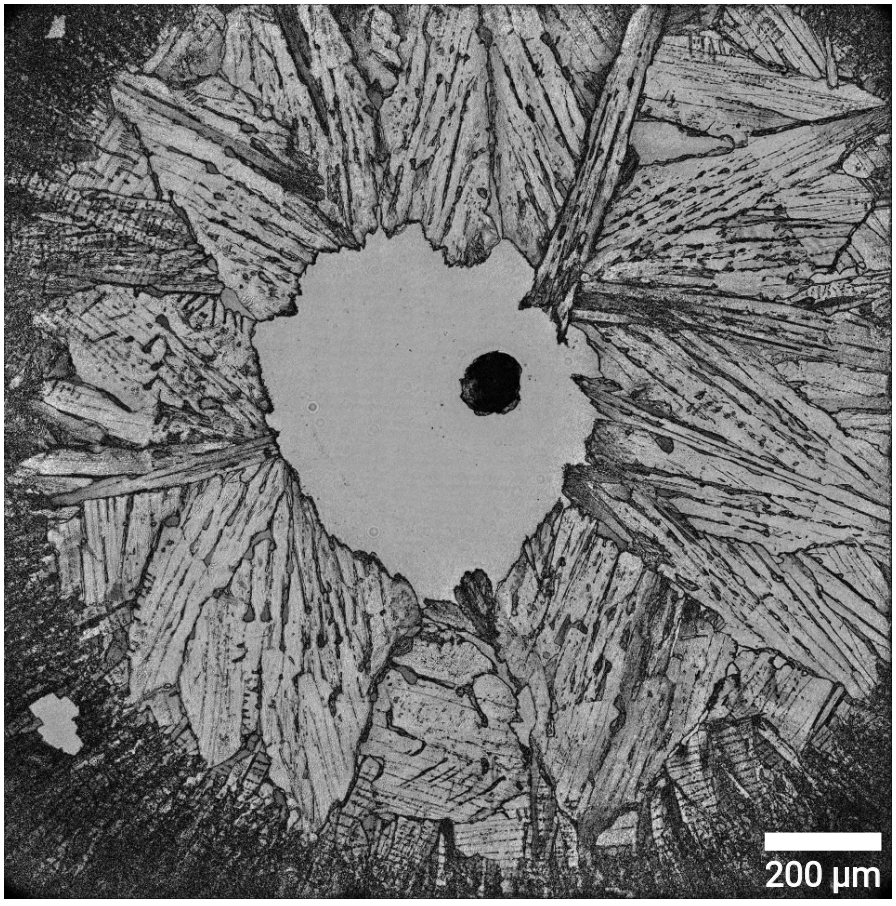

FIGURE 4.1: Amplitude reconstructed by the ptychography method using a HeNe laser ($\lambda = 633\,\text{nm}$) passing through the NMBA sample. The amplitude distribution shows that the transmission through most of the crystalline regions was close to the transmission through the central region that is from pure glass with intensity attenuation between 10-20%.

4.2 Second harmonic generation theory

The second harmonic generation process was first observed in 1961 by Franken *et al.* [138] and it is the simplest example of an optical nonlinear effect. It can be described by the Taylor expansion of the polarization P

$$\begin{aligned} \mathbf{P} &= \varepsilon_0 \boldsymbol{\chi}_L \mathbf{E} + \mathbf{P}_{\mathrm{NL}} \\ (P_{\mathrm{NL}})_i &= 2 d_{ijk} E_i E_k \quad , \end{aligned} \tag{4.16}$$

where $\boldsymbol{\chi}_L$ is the linear anisotropic tensor and d_{ijk} is the nonlinear tensor for the SHG processes. The tensors can have different symmetries depending on the crystalline structure of the material, however every SHG active material must posses non-centrosymmetrical structure. This also means that isotropic material cannot produce SHG [135] except the material boundaries, where the symmetry is disturbed.

4.3 Light propagation through SHG medium

Light propagating through the SHG medium still needs to follow the linear propagation equations shown in the previous section, but additionally there is the nonlinear polarization term $\mathbf{P}_{\mathrm{NL}}$. The nonlinear polarization term $\mathbf{P}_{\mathrm{NL}}$ (Eq. 4.16) as shown in Ref. [135] can be inserted into the Maxwell's equations

$$\begin{aligned} \nabla \times \mathbf{H} &= \boldsymbol{\sigma} \mathbf{E} + \frac{\partial \boldsymbol{\varepsilon} \mathbf{E}}{\partial t} + \frac{\partial \mathbf{P}_{\mathrm{NL}}}{\partial t} \\ \nabla \times \mathbf{E} &= -\mu_0 \frac{\partial \mathbf{H}}{\partial t} \quad , \end{aligned} \tag{4.17}$$

where $\boldsymbol{\sigma}$ denotes conductivity (losses) and $\boldsymbol{\varepsilon} = \varepsilon_0(1 + \boldsymbol{\chi}_L)$. The wave equation can be expressed from Eqs. 4.17 as

$$\nabla^2 \mathbf{E} = \mu_0 \boldsymbol{\sigma} \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \boldsymbol{\varepsilon} \frac{\partial^2 \mathbf{E}}{\partial t^2} + \mu_0 \frac{\partial^2 \mathbf{P}_{\mathrm{NL}}}{\partial t^2} \quad . \tag{4.18}$$

The electric field along the propagation direction (axis 3) polarized along j-th direction can be expressed [135] as

$$\frac{\mathrm{d}E_{3j}}{\mathrm{d}z} = -\frac{\sigma_3}{2}\sqrt{\frac{\mu_0}{\varepsilon_3}} E_{3j} - \mathrm{i}\omega_3 \sqrt{\frac{\mu_0}{\varepsilon_3}} d_{ijk} E_{1i} E_{2k} e^{\mathrm{i}(k_3 - k_2 - k_1)z} \quad , \tag{4.19}$$

where $\omega_n, \varepsilon_n, k_n$ denotes frequency, dielectric constant and wave-number for the n-th beam. In the case of second harmonic generation, the equation can be simplified using the following relations

$$\omega \equiv \omega_1 = \omega_2 = \frac{1}{2}\omega_3, \quad \Delta k \equiv k_3^{(j)} - k_1^{(i)} - k_1^{(k)} \quad ,$$

where $k_1^{(k)}$ is wave-number for the beam at frequency ω_1 polarized along the k direction. If the sample is fully transparent ($\sigma = 0$) and there is no input second harmonic light $E_{3j}(0) = 0$, the amount of the produced SHG is low compared to the IR power and the illumination is a plane wave, then the integrated SHG electric field along z from distance 0 to the crystal thickness L can be expressed as

$$E_{3j}(L) = 2\mathrm{i}\omega\sqrt{\frac{\mu_0}{\varepsilon_3}} d_{ijk} E_{1i} E_{1k} \frac{e^{\mathrm{i}\Delta k L} - 1}{\mathrm{i}\Delta k} \quad . \tag{4.20}$$

This equation results in quadratic growth of the SHG intensity in dependence on distance L (i.e. $E_{3j} \sim L^2$) if $|\Delta k| \ll 1$. Therefore, matching the fundamental wavenumber k_1 and the SHG wavenumber k_3 will maximize the SHG yield. This effect is called phase-matching and it is important for any harmonic generation process. More details about phase-macthing of harmonic processes are provided in Section 5.4.

The Eq. 4.20 is very simple, however it is based on assumtion of flat illumination wavefront that is often not well describing the SHG microscopy. The following formula that assumes converging Gaussian illumination spot [139]

$$E_{3j}(L) \sim d_{ijk} E_{1i} E_{1k} \exp\left(-\frac{\alpha_{2\omega}}{2} L\right) \int_0^L \frac{\exp[(-\alpha_\omega + \alpha_{2\omega}/2 + \mathrm{i}\Delta k)z]}{z - l_f + \mathrm{i}z_0} \mathrm{d}z \quad , \tag{4.21}$$

where the absorption parameters $\alpha^{(\omega)}, \alpha^{(2\omega)}$ denote the absorption coefficients of the fundamental and SH waves, respectively. l_f denotes distance from the incident surface of the sample to the focus and z_0 is the Rayleigh length of the Gaussian beam. For simplicity, we will use the Eq. 4.20 for explanation of the principle of nonlinear CDI. In a standard SHG microscopy experiment, there is only a single input beam and thus $E_{1i} = E_{1k}$ and $i = k$. The output electric field is at frequency 2ω.

Our nonlinear ptychography method [7] is based on the idea that the exit wave field ψ has to be separable into a part that shifts with the sample position and a

part that corresponds to the illumination probe

$$\psi_j(\mathbf{x}) = P(\mathbf{x} - \mathbf{x}_j)O(\mathbf{x}) \quad .$$

This can be easily shown to be valid if the Eq. 4.20 is rearranged

$$\psi_j(\mathbf{x}) \sim \left[E_{1i}^2(\mathbf{x} - \mathbf{x}_j)\right] \left[\sqrt{\frac{1}{\varepsilon(\mathbf{x})}} \frac{e^{\mathrm{i}\Delta k(\mathbf{x})L(\mathbf{x})} - 1}{\mathrm{i}\Delta k(\mathbf{x})} d_{ijk}(\mathbf{x})\right] \quad , \tag{4.22}$$

where the first square bracket contains the probe related terms and the second the object terms. We will call the second term *nonlinear conversion factor*. The nonlinear conversion factor contains mixed terms corresponding to the linear response of the crystal $\Delta k(\mathbf{x}), L(\mathbf{x}), \varepsilon(\mathbf{x})$ and the nonlinear tensor d_{ijk}. The relative phase shift $\Delta k(\mathbf{x}) \cdot L(\mathbf{x})$ can be estimated from two linear ptychography reconstructions with ω and $\omega/2$ illuminations. However, in order to get the quantitative values, it is still necessary to measure the local thickness using some external diagnostic or use ptychography in the reflection mode [28, 140].

For the phase-matching based processes, it is useful to define a constant called the "coherence length"

$$L_c \sim \frac{\pi}{|\Delta k|} \quad . \tag{4.23}$$

Note that L_c has no connection to the laser coherence. It is just a common parameter denoting the distance over which the harmonics add up coherently and thus increasing their intensity.

Virtual probe

We call the first term in the Eq. 4.22 the virtual illumination probe

$$P \sim E_{1i}^2(\mathbf{x}) \quad .$$

On contrary to the linear ptychography, the probe at the SHG frequency does not exist in front of the sample and it is only created by the nonlinear interaction. It can be easily seen that the intensity distribution of the virtual probe is proportional to the square of the linear intensity distribution and the phase curvature is twice larger. The virtual probe collects also other features of the system that are not moving with the sample, such as the diffuser that we use to enhance the high spatial frequencies in the illumination or more correctly after the illumination.

Examples of the linear and nonlinear probes taken in the same configuration and for the same sample are shown in the Fig. 4.2.

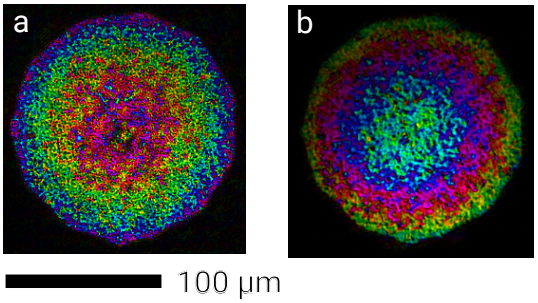

FIGURE 4.2: Two examples of the reconstructed probes in the complex colour scale. Probe (a) is a linear illumination probe at 400 nm, where the 400 nm was produced by a BBO crystal placed upstream the microscope. The probe (b) is the nonlinear illumination probe at 400 nm wavelength with 800 nm illumination wavelength.

4.4 Experimental results

The experimental data were collected using our visible light ptychography setup (see Fig. 4.4) introduced in the previous Chapter 3. Samples were illuminated by a Coherent MIRA femtosecond Ti:sapphire laser designed for pulse length 120 fs with 80 MHz pulse frequency and output beam power 3 W centred around 780 nm wavelength. Unfortunately, because the laser diodes were already beyond their designed lifetime, the output power was around 460 mW at the output which results in power of 220 mW at the sample due to unavoidable optics in the beam path. Moreover, the pulse length was estimated to be around 700 fs from the measured spectrum as it is shown in Fig. 4.3. This resulted in roughly 50× lower peak power compared to the optimal system and due to the quadratic dependency of the produced SHG light, the measured SHG intensity was roughly 3 orders of magnitude lower.

The polarized ultra-fast laser beam was loosely focused by a 25 cm focal length lens on a SHG active sample resulting in illumination NA≈0.01 in order to provide relatively flat wavefront illumination spot with FWHM roughly 90 μm. The resulting peak power density is $0.9{\cdot}10^6$ W.mm^{-2}.

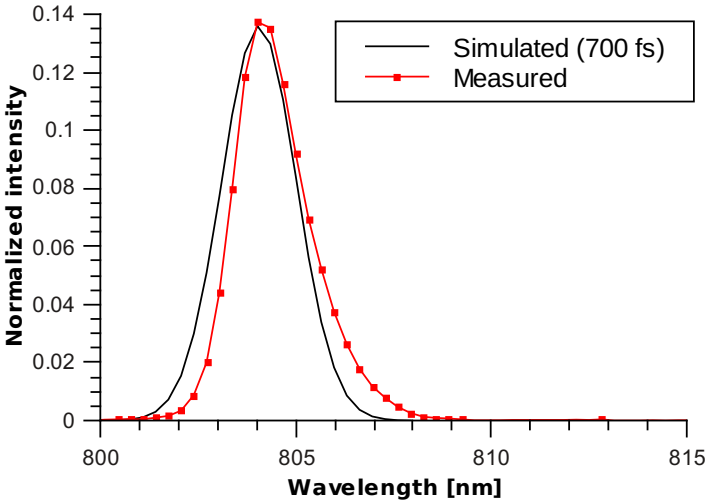

FIGURE 4.3: Measured spectrum of the MIRA laser (red) compared to expected spectrum for 700 fs pulse (black).

The high intensity IR light was removed by a narrow bandwidth filter (400±10 nm) placed in front of the PIXIS CCD camera. The filter also removed most of the background illumination in the already dark lab.

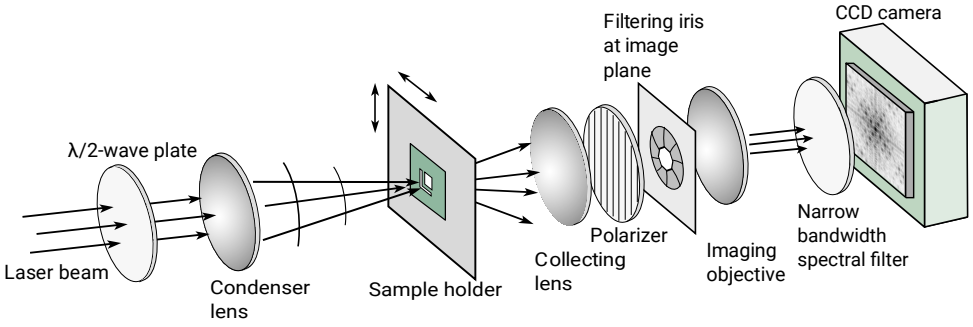

FIGURE 4.4: A simplified schematic of the SHG ptychography setup. The condenser lens increases the beam intensity on the sample. The illumination probe at the SHG wavelength is created inside the sample and then further defined by an iris and diffuser placed at the imaging plane of the collecting lens. The strong fundamental beam is removed by a narrow bandwidth filter placed on the camera.

The maximum resolution of the linear CDI experiments was limited by the used collecting microscope lens to NA=0.4. The real resolution however nonlinearly depends on the available IR flux and pulse length. It also depends on the sample SHG activity and transmission at fundamental and SHG wavelengths. Therefore, the signal was often limited by the maximum acceptable exposure time and limited stability of the used femtosecond laser (MIRA, Coherent). In the case of IR light (800 nm), the maximum achievable full period resolution was 2 μm and for SHG

(blue light, 400 nm) the resolution was 1 μm. Precise pixel size in the image was calculated for each reconstruction using a cross-correlation based method (2.6.2). The relative sample positions were assumed to be precise. The probe illumination diameter was 136 μm, therefore the linear oversampling of the dataset was around 1.5 for SHG and 3 for IR data. The measurement was also complicated due to instability of the used ultra-fast laser. The laser beam position was rapidly vibrating, causing blurring of the diffraction pattern, the laser intensity was varying and sometimes the mode-locking dropped off during longer scans. For these reasons, the scanning time had to be as short as possible.

The exposure time for the SHG data was mainly limited by flux and most of the SHG measurements were taken with maximum exposure time 10 s per position with slowest readout speed (100 kHz) in order to suppress the readout noise. The NMBA scans contained 900 scanning positions, the PPLN sample scans used 1185 positions. The step size was set to 30 μm. In combination with 136 μm probe diameter, it resulted in 80% linear overlap between the positions. The local thickness of the NMBA sample shown in Fig. 4.5 was measured by a white light interferometer.

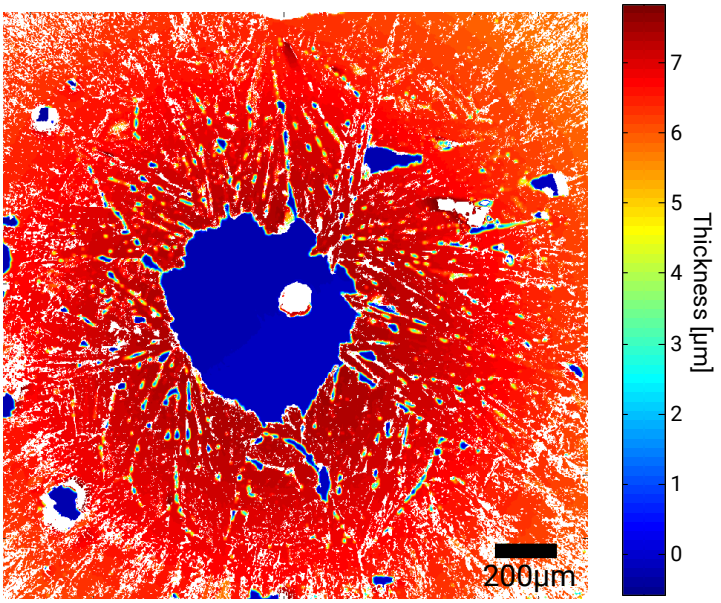

FIGURE 4.5: Thickness map of the NMBA sample measured by a white light interferometer. The white regions were removed due to errors. The thickness of the crystal is mostly around 7 μm.

The reconstructions of our sample for different polarizations are shown in the Fig. 4.7. An example of the measured diffraction pattern is in Fig. 4.6. The

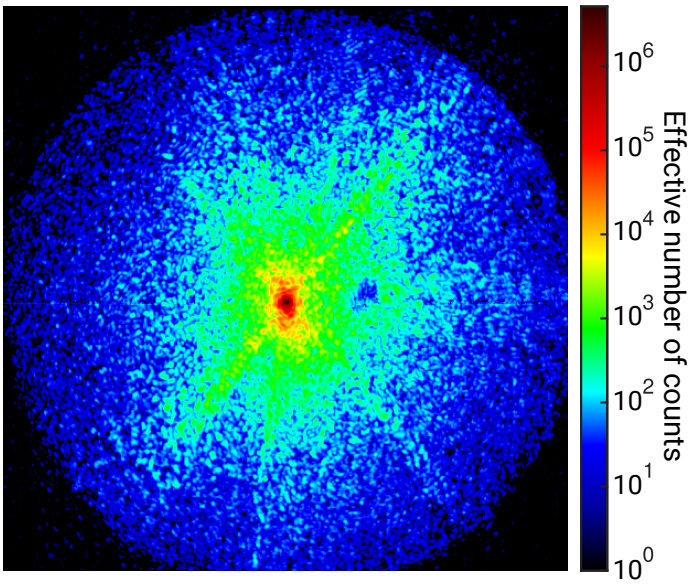

FIGURE 4.6: An example of a measured diffraction pattern from the NMBA crystal. The maximal measured scattering angle is limited by the acceptance NA of the front collecting lens (see setup in Fig. 4.4). The exposure time was in range between 20 ms-1 s.

central region is plain glass that is not introducing any phase shift or SH signal. The central region is surrounded by a ring of roughly 7 μm thin NMBA layer that is composed from many crystals with almost random orientation. Different orientations were shown in Fig. 3.11.

The central region of the cross-polarized signal from 800 nm wavelength (Fig. 4.7) shows that the polarization was not perfect because there was still light going through the glass. The passing intensity through crossed polarisers was below 5%.

Figure 4.8 shows comparison of the reconstructed relative phase shift for reconstructions with 800 nm, 400 nm illumination and a SHG reconstruction. The images were mutually aligned and the phase was shifted to get zero in the central region. Unfortunately, the complex structure of the object with significant local phase variations prevented us to use standard phase unwrapping algorithms such as the Goldstein unwrapping [141]. The blue reconstruction (b) shows roughly twice faster phase evolution than the IR reconstruction (a). However, the SHG reconstruction (c) shows even faster phase evolution than the blue reconstruction, which is caused by the phase-matching process. The effect was simulated using Eq. 4.21. The NMBA sample is highly attenuating at 400 nm wavelength. The attenuation length for 400 nm is 1.2 μm, the attenuation for 800 nm is negligible.

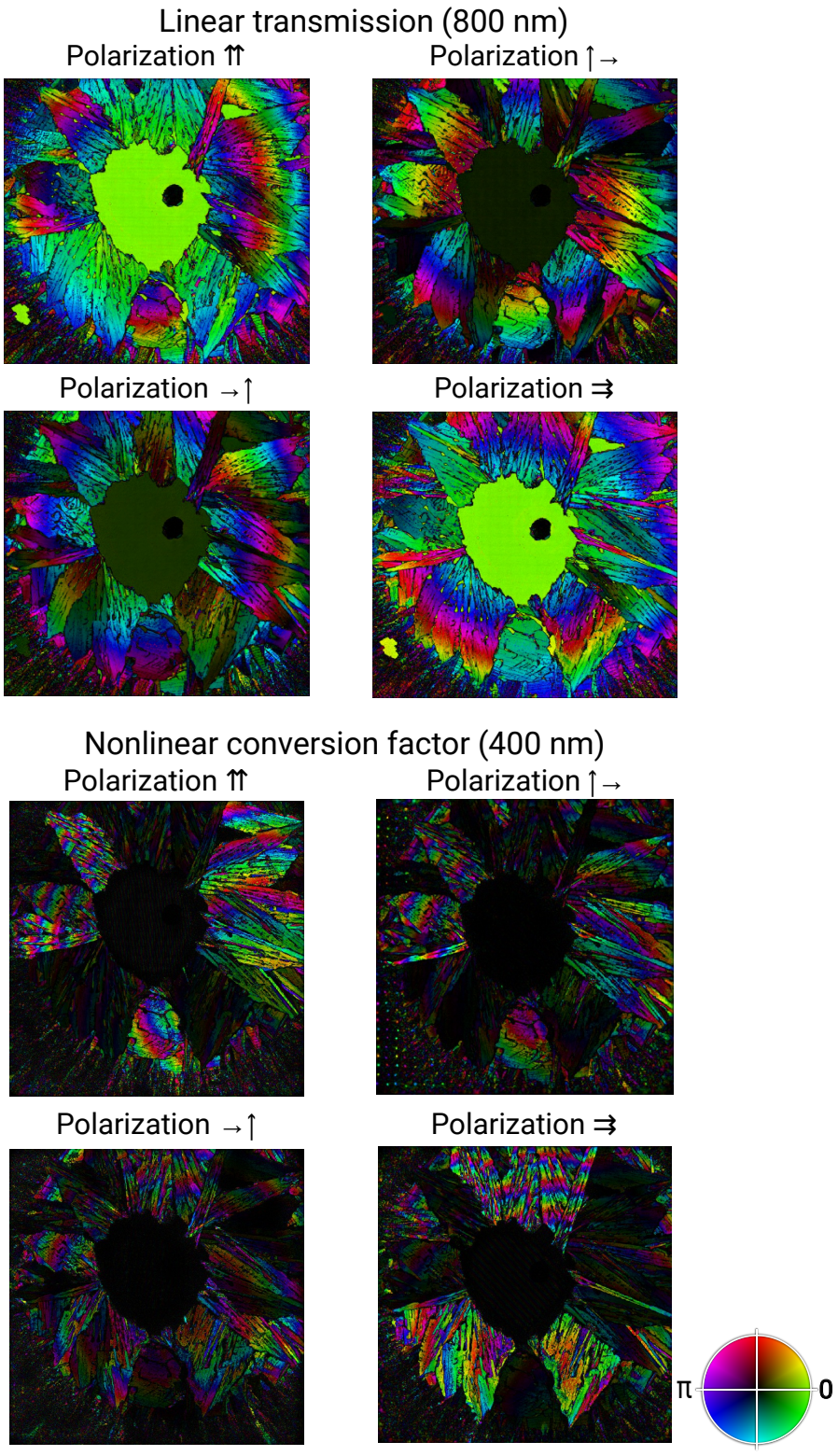

FIGURE 4.7: SHG and IR ptychography reconstructions for different polarizations of the input beam and the analysing polarizes in the complex colour scale. Note that these images are showing electric field and not intensity, which relatively amplifies the weak intensity regions.

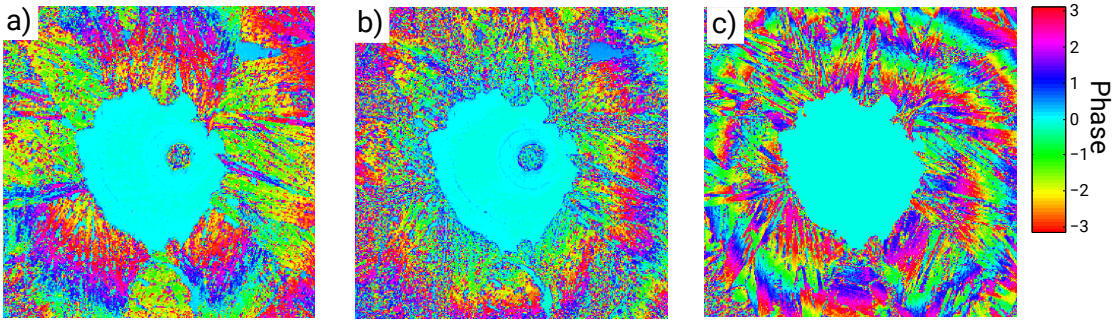

FIGURE 4.8: Comparison of phase evolution for linear ptychography at 800 nm (a), linear ptychography at 400 nm (b) and SHG ptychography (c). The phase change for the SHG (c) is varying significantly faster than for the blue reconstruction (b). Image (b) is slightly more degraded because it was taken later, when the sample was already more eroded.

Furthermore, the sample is highly birefringent. The difference between the refractive index for 800 and 400 nm is up to 0.9 [125], which results in the phase mismatch term up to 25 μm^{-1} and coherence length approximately 0.4 μm-10 μm in dependence on the crystal orientation.

In order to verify the measured phase shift values, we have simulated the produced SHG field. Figure 4.9 shows simulation of amplitude and phase-shift for different propagation distances though a harmonics generation sample and phase mismatch values Δk. In the case of small phase mismatch ($|\Delta k| \ll 1$), the phase ramp up corresponds to 400 nm light propagation through the sample without nonlinear effects. However, in the case of maximal phase mismatch that can be reached in the NMBA sample, the SHG phase can grow 2-3 faster than the phase-shift of the 400 nm light. This agrees well with the measured phase in Fig. 4.8.

Also note that due to large absorption of the 400 nm wavelength, the produced SH field is not growing quadratically even for ideal phase-matching ($\Delta k = 0$), but it reaches saturation after $\approx$3 μm distance from the beginning of the sample. It means that only roughly the last 3 μm of the sample participate on the measured harmonic signal.

Figure 4.10 shows the effect of crystal twinning. The highlighted regions are showing that there is no variation in the linear phase, while there is a strong variation in the SHG phase. It is caused by reverse sign of the nonlinear coefficient d_{ijk} due to reversal of the crystallographic axis, which results in π phase change in the SHG field.

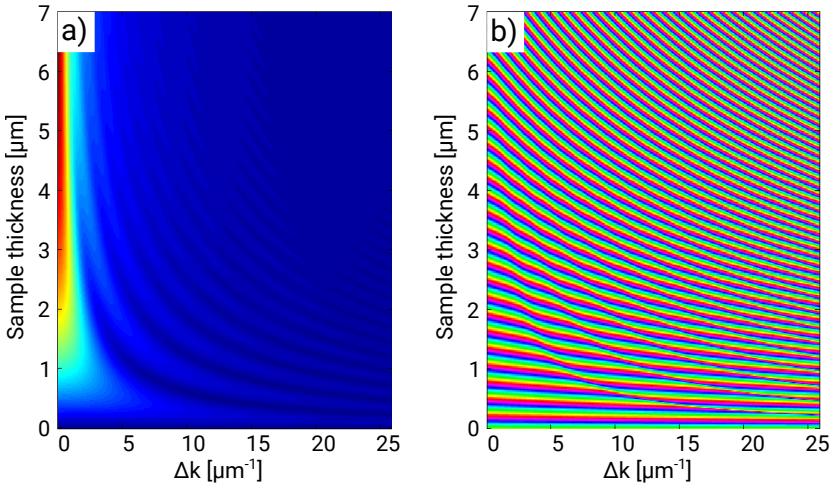

FIGURE 4.9: Simulation of the SHG phase and amplitude for different thickness and phase mismatch of the the NMBA sample. The sample thickness is between 5-7 μm. The phase mismatch Δk can be up to 25 μm^{-1}. Image (a) shows the modulus of the produced SHG field and image (b) is the corresponding phase.

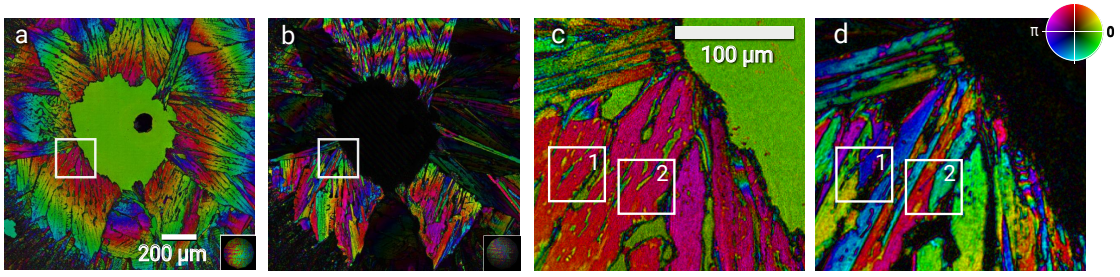

FIGURE 4.10: (a,b) IR and SHG reconstructions of the NMBA sample. (c,d) detailed images of the reconstruction show phase π jumps caused by crystal twinning. The highlighted regions 1,2 in image (c) show that there is no variation in the IR reconstruction but there is a significant phase change in the nonlinear reconstruction (d). Adapted from [7].

The second investigated sample was periodically poled lithium niobate (PPLN) provided by Covesion Ltd. PPLN is an optical crystal with highly nonlinear properties and hence widely used for harmonic generation or frequency mixing [142]. The disadvantage is that it is not possible to reach phase matching simply by the optimal crystal orientation. Therefore, it is necessary to use the quasiphase matching. The PPLN sample was manufactured with a periodic reversal change in the nonlinear coefficient d_{ijk} caused by strong electric field that changes the crystal orientation. The linear optical properties are not changed by this process, therefore the crystallographic change cannot be measured directly by an optical microscope. Lithium niobate was deposited on a glass substrate then polished so that the PPLN layer was reduced roughly to 80 μm thickness. The reconstructions of the PPLN sample are shown in Fig. 4.11. Figure 4.11(a) contains the phase

ramp subtracted complex transmission and 4.11(b) contains the phase ramp subtracted nonlinear conversion factor. The phase ramp up is caused by wedged shape of the sample and the phase change allows us to calculate that the thickness gradient was 3.78 μm/mm. The original reconstructions before the phase subtraction are shown in the upper inset and the illumination probe at scale is in the lower inset. The poled region is easily visible in the nonlinear reconstruction, however with the phase ramp removed, a small variation of the optical thickness is visible even in the linear reconstruction. Also there is optical thickness variation between the broad poled stripe and the unpoled area. These effects are partly caused by material properties change during the poling process and partly maybe due to the sample polishing process. Figure 4.11(b) also contains two diagonal dark stripes (i.e. regions with zero amplitude of SHG light). The intensity drops are parallel with the phase ramp up (see upper inset in Fig. 4.11(b)). It was caused by the phase matching process. At this region the produced SHG interfered destructively in the direction of the IR beam and it interfered constructively in the counter direction of the IR beam propagation. Separation of the two dark diagonal stripes is 597 μm. Using the measured sample thickness gradient, the thickness difference is 2.255 μm. If we estimate the phase-matching term Δk from the Sellmeyer equation for MgO-doped PPLN [143], the coherence length (Eq. 4.23) along the beam axis is equal to 1.125 μm. This corresponds to zero SHG intensity separation 2.25 μm along the beam axis. This also means that the observed SH light was produced only by the last $\approx 2\,\mu$m of the sample.

Figure 4.12 shows detail of the poled region. The image 4.12(b) allows us to see some imperfections of the poling process (wider lines). We have measured the phase in between these regions and the phase evolution is shown in Fig. 4.12(c). The measured PPLN poling period is 6.99 μm which corresponds well to the designed period of 6.94 μm. The π phase difference is similarly to the NMBA sample caused by opposite orientation of the crystal along the crystallographic axis. This reverses the sign of the nonlinear coefficient d_{ijk} and caused π jump.

The presented datasets are availible at DOI:10.5258/SOTON/386425.

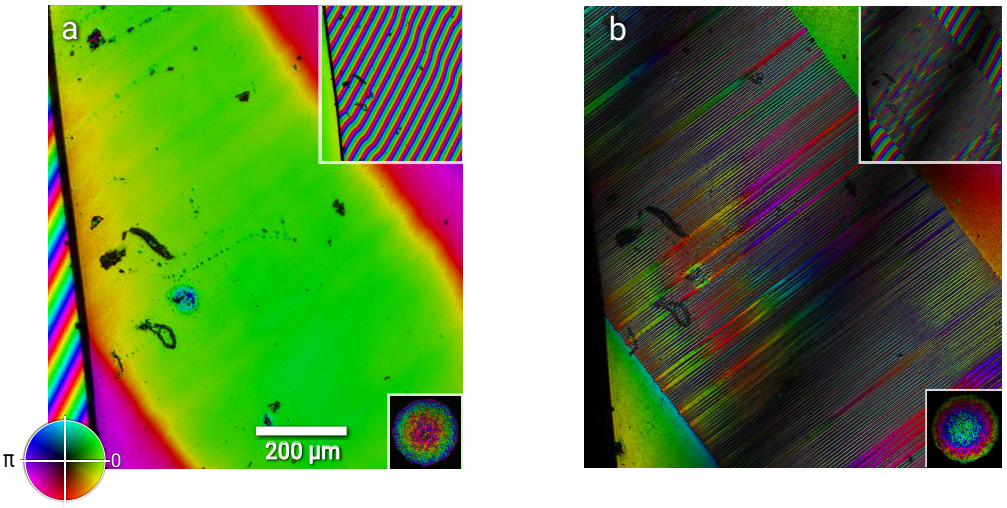

FIGURE 4.11: The reconstruction (a) shows linear complex transmission of the PPLN sample at 400 nm. The phase ramp was subtracted (original is in the upper inset). In the left most region of the reconstruction, the beam is passing directly through air. This allows us to accurately measure the thickness variation. The poled region can be seen as the region with slightly larger phase shift. The lower inset contains the illumination probe in scale. The details of the reconstructed probes can be seen in Fig. 4.2. The image (b) shows the nonlinear conversion factor reconstructed using the nonlinear CDI method. Significant variations in phase are caused by sign changes in the SHG tensor d_{ijk}. Adapted from [7].

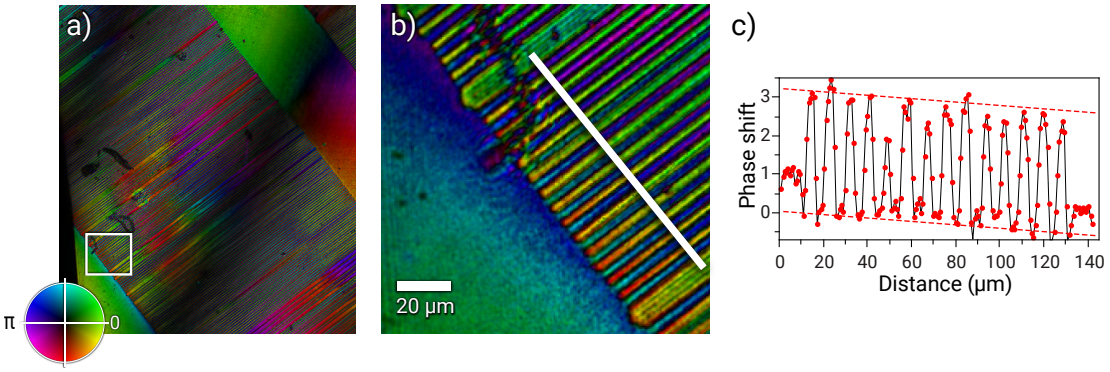

FIGURE 4.12: Image (b) shows detailed view of the poled region highlighted in (a). The reconstructed phase of the poled region was measured along the white line in (b). The phase evolution (c) shows the π phase variations caused by the poling process.

4.5 Conclusion

The presented nonlinear CDI experiment was only a proof of concept of a new method. Compared to the standard confocal SHG microscopy, it allow us to reconstruct the phase of the generated SHG light and thus obtain additional information. The phase reconstruction of the SHG field can be also achieved using

holography methods [127, 128], however it requires beam splitting and separated SHG generation for the reference wave, which leads to less efficient use of the available flux. Moreover, nonlinear ptychography allows us to get artefact-free reconstruction without requirement of a flat illumination wavefront. Finally, ptychography allows us to use the whole framework of methods proposed for linear ptychography such as the multi-slice ptychography [25] or our anisotropic ptychography and the orthogonal probe relaxation method.

In future, the main potential of the nonlinear ptychography is in combination with the higher harmonic processes such as Coherent anti-Stokes Raman spectroscopy (CARS) microscopy that is based on third harmonic generation (THG). THG is produced practically by any material because there is no requirement of non-centrosymmetrical crystalline structure. However, for practical microscopy, THG needs to use illumination wavelength above 1000 nm in order to use the produced harmonics in the wavelength sensitivity range of the standard silicon based detectors. Also the produced power will be generally lower compared to SHG and thus more tightly focused beam will be required. On the other hand, the nonlinear confocal microscopy of many samples is limited by the damage threshold, therefore expanding the illumination beam (e.g. going out of focus) will allow us to use more flux without sacrificing the resolution. A small illumination probe will lead to a large number of scanning positions with relatively low number of collected pixels per position. We estimate that the scatter pattern will be sufficiently described by $\approx 64 \times 64$ px compared to 512×512 px used now. Collection of these datasets would be inefficient with CCD sensors, therefore a scientific CMOS sensor would need to be used to guarantee sufficient acquisition speed and low readout noise.

The proposed nonlinear SHG ptychography can provide an estimation of rotation and thickness of small SHG active crystals. However, it would be very difficult to obtain all material constants of an unknown crystal without any prior knowledge. A general crystal has a large number of unknown variables:

Variable name	maximum number of unknowns
Complex refractive index (800 nm)	3+3
Complex refractive index (400 nm)	3+3
Rotation of axis	2
Thickness	1
Nonlinear tensor	up to 18

where rotation is generally different for every crystal and thickness varies for each reconstructed pixel. However, as we have shown in this chapter, this method can be used if some prior knowledge about the sample is added.

Chapter 5

High Harmonic Generation (HHG) used for CDI EUV imaging

Coherent diffraction imaging (CDI) methods were originally developed for the short wavelength regions, where the imaging optics are challenging to manufacture, expensive, and experience high losses [144, 145]. Nowadays, the quality of the X-ray optics is often not the main limitation and the main advantage of CDI methods is the ability to provide measurement of the phase shift and attenuation of the sample in these short wavelength regions. Phase measurement is crucial for hard X-ray imaging when the attenuation of most materials is low and majority of the interaction between material and photons is caused by elastic scattering. Most of the sample information is then contained in the phase that is lost in the standard lens-based bright-field microscopy. Some phase sensitivity can be achieved if the sample is imaged slightly out of focus leading to so called propagation based contrast [41]. It is also possible to use the Zernike microscopy [146] or the scanning transmission X-ray microscopy (STXM) that can also provide the local phase gradient. The advantage of CDI is the ability to provide diffraction limited artefact-free quantitative phase values that are, in the hard X-ray regime, almost directly proportional to the sample projected electron density.

Currently, the main drawback of X-ray coherent sources is their limited availability because the only sources with sufficient coherent flux are 3rd and 4th generation

synchrotrons and free electron lasers. However, the user beamtime of these facilities is limited and therefore there is a need for laboratory scale sources. The second potential drawback of the X-ray imaging is very low contrast for unstained biological samples and therefore it requires very high coherent flux density that can cause significant radiation damage in order to achieve reasonable resolution for weakly scattering specimens.

CDI methods have generally high requirements on the coherence and stability of the illumination beam. Theoretically, CDI promises diffraction limited resolution, however in the practical application the resolution is often limited rather by low signal or by systematic errors. This issue was already shown in the synchrotron based experiments [89–91, 114, 147–149], however it is even more critical for tabletop EUV experiments. Despite this, in the last years, simple CDI experiments with tabletop EUV sources such as high-harmonic generation or plasma discharge sources have been successfully conducted but only for highly scattering practically binary samples [28, 29, 56, 150–152].

There are several reasons why coherent EUV laboratory sources have not been yet used for high resolution imaging of real weakly scattering samples such as biological specimens. Firstly, the available photon flux density is very low compared to large facilities. It is more challenging to generate sufficient flux for high resolution imaging of weakly scattering and relatively thick biological samples compared to simple binary highly scattering samples. Furthermore, the coherence properties of the coherent laboratory sources are usually worse compared to the large facilities. This can be improved by additional filtering, which would however even further reduce the available photon flux. Finally, the limited stability of the coherent laboratory sources makes it more difficult to collect sufficient amount of data with constant illumination for large scale reconstructions through thicker specimens.

In this chapter, we will introduce our laboratory EUV CDI microscope and among other, show almost artefact-free high resolution ($\approx$50 nm) reconstruction of biological samples such as hippocampal neurons using the far-field ptychography. Further, high fidelity reconstructions with large field of view ($>$1 mm^2) using the near-field ptychography method [38] are shown in this chapter. The strict requirements of CDI methods on the overall beam quality such as high spatial and temporal coherence, related to the oversampling ratio of the diffraction pattern [48], and good temporal stability are relaxed by the methods presented in the chapter 2.

In Southampton university, we use an EUV source based on the high harmonic generation (Chapter 5.1-5.4). It is easier to produce coherent EUV light than X-rays because of the lower photon energies, however it is still extremely challenging to produce stable illumination with high coherent flux. HHG process can produce sufficient flux for imaging down to 13 nm wavelength [56], however in our experiment, we were limited by the available IR laser pulse energy and pulse length to wavelengths around 29 nm produced in argon gas.

5.1 Theory of HHG

Harmonic generation (HG) is a nonlinear process that uses a fundamental frequency ω beam to generate higher frequencies at multiples of the fundamental frequency. This process was first discovered by Franken [138] in 1961, who used crystalline quartz as the non-linear medium to generate second harmonic frequency using a focused ruby laser.

Low order HG is a result of the standard non-linear interaction of the light with a material and this process is relatively well understood. Higher order harmonic photons can be also generated, however, the efficiency of this process rapidly decreases for higher orders because of low probabilities of multi-photon interaction as it is a perturbation process. Generally, the HG process preserves the temporal and spatial coherence of the generating field and thus it allows to generate short wavelength coherent light.

5.2 Basic principle

If an atom is in an intense electric field, the electromagnetic force can be so large that is comparable with the Coulomb force of the atom. When the driving field intensity is relatively low, the electric field interaction with matter can be considered a small perturbation of the Coulomb potential of the atom. In this case, it is possible to approximate the polarization P by the Taylor expansion

$$\mathbf{P} = \varepsilon_0 \sum_{i=1}^{N} \chi_i \mathbf{E}^i \quad ,$$

and higher orders of the polynomial approximation are added if the intensity is increased.

However, in the case of the high harmonic generation, the intensity is so high ($\approx 10^{14}\,\mathrm{Wcm}^{-2}$), that the interaction with matter cannot be modelled by a simple polynomial approximation of the polarization tensor P and the response of the electrons should to be modelled by full quantum mechanical calculations.

If the driving field is strong enough, several ionization effects can be observed:

1. multi-photon ionization - several photons simultaneously interact with an electron and remove the electron from shell
2. over-barrier ionization - it occurs if the force caused by the field intensity exceeds the Coulomb force.
3. tunnel ionization of the electron - there is non negligible probability that the electron will tunnel out even for field lower than the Coulomb potential.

All these effects can participate on the HHG, however the tunnel ionization is preferred because it is most effective and leads to less plasma generation. The optimal intensity range for this process is called tunnelling ionization regime and the rate at which tunnelling occurs can be calculated using the ADK rate [153].

5.3 Semi-classical model – three step model

The three step model suggested by Corkum et al. [154] is a useful way to explain some of the HHG parameters. This model employs a combination of quantum mechanics and classical Newtonian physics. Response of each atom to the high intensity electric field is treated separately. The schematic of the 3-step model is shown in Fig. 5.1.

Firstly, the high electric field distorts the Coulomb potential and allows the outer most electron to tunnel out. The free electron is subsequently accelerated by the driving electric field. Under certain conditions, there is finite probability that the electron returns back to the original ion. During recombination the gained kinetic energy of the electron and ionization energy is radiated as a photon. This results

in series of sub-femtosecond coherent bursts of EUV photons separated by half of the period of the driving laser field.

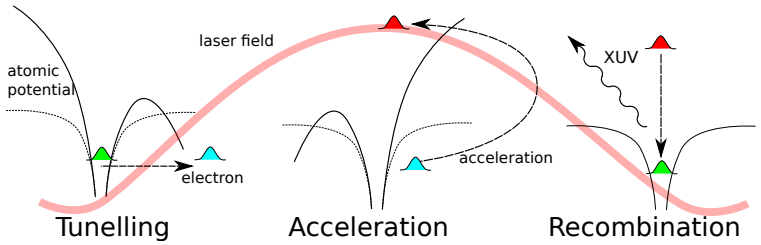

FIGURE 5.1: An illustration of the HHG "three step" model. At first, the electron tunnels out of the Coulomb barrier, then it gets accelerated be the EM field and finally it recombines back and the gained energy is radiated.

The energy that the electron gains during one cycle of acceleration is called the ponderomotive potential U_p and it can be obtained from integration of the electric field over one period

$$U_p = \frac{e^2 E_0^2}{4m\omega^2} \quad ,$$

where e, m are the charge and mass of the electron and E_0, ω are amplitude and frequency of the linearly polarized driving electric field. The probability that the electron returns back to the origin depends strongly on phase of the electric field at the moment of the tunnelling ionization.

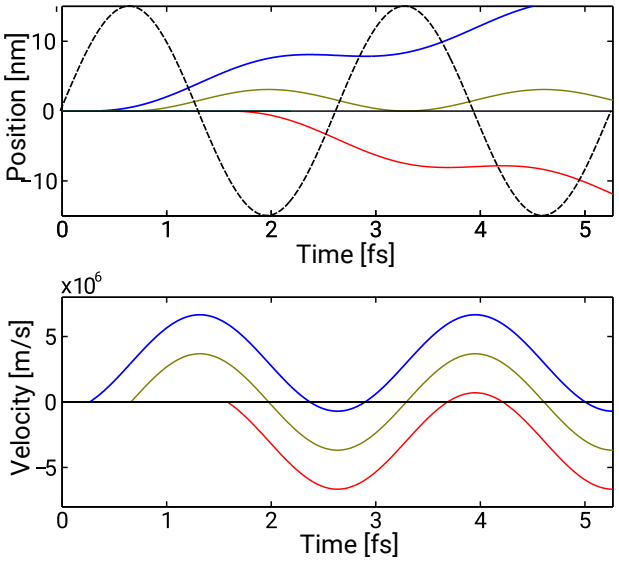

FIGURE 5.2: (top) Trajectories of electrons (full lines) that tunnelled out at different time between $0 \leq t_0 \leq T/4$ for sin driving field (dashed black line). (bottom) Velocity of the electron gained during the acceleration by the laser field.

The electron can return if the phase of the driving field during the tunneling event was between 0 and $\pi/2$ or in the opposite direction between π and $3\pi/2$ for a cosine laser driving field. Otherwise the electron is accelerated away from the atom and the parent atom stays ionized. The recombination appears each half-cycle and harmonics with opposite phase are generated. However only the odd harmonics are produced because the even orders interfere destructively. Because each electron returns at a different time, the phase-shift and energy (Fig. 5.3) of the generated electromagnetic field is different. This effect is increasing spectral line-width of the generated harmonics in the case of short pulses and high intensities [8].

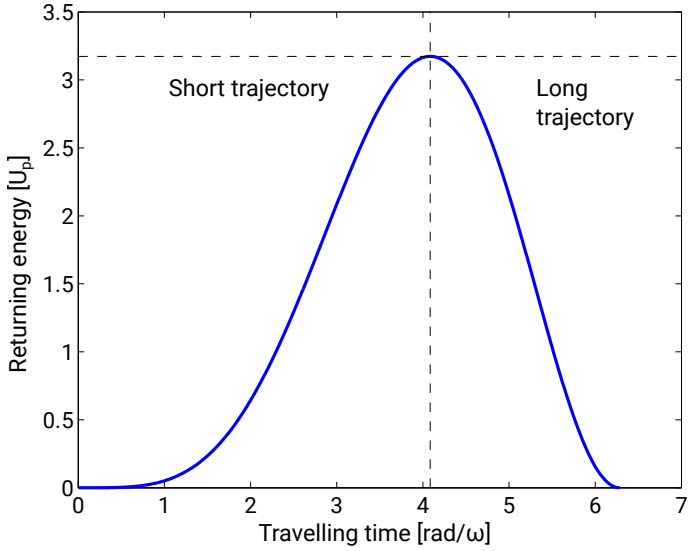

FIGURE 5.3: Dependence of the return energy on the time spent in continuum. Trajectories that return before the maximum energy are called "short" trajectories and the rest are "long trajectories". Only for cut-off are all trajectories equal.

Using the equation of electron motion in electric field, the maximum released photon energy can be estimated as

$$E_{\text{cutoff}} = 3.17U_p + I_p \quad .$$

Electrons with returning energy lower than E_{cutoff} can be results of two different trajectories. These trajectories are called long and short depending on the time of ionization of the parent atom. Both trajectories generate HHG light and the ratio between them depends on many factors [8]. The semi-classical model is able to well predict some experimental results such as the highest achievable harmonic (cut-off), but a quantum mechanical model needs to be used for more details

Ionization energy	Element	Atomic number
12.1298	Xenon	54
12.9676	Chlorine	17
13.5984	Hydrogen	1
13.6181	Oxygen	8
13.9996	Krypton	36
14.5341	Nitrogen	7
15.7596	**Argon**	18
17.4228	Fluorine	9
21.5645	Neon	10
24.5874	Helium	2

TABLE 5.1: Elements with the highest ionization energies from the neutral state.

about spectrum shape and behaviour. An illustration of HHG spectrum envelope is shown in Fig. 5.4.

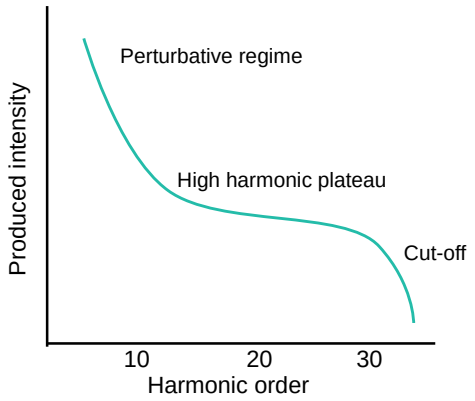

FIGURE 5.4: Expected envelope of the ideal HHG efficiency. It is showing fast decay for low harmonic orders followed by a flat region (plateau) and then the cut-off predicted by the 3-step model.

In order to increase the maximum produced photon energy (E_{cutoff}), either the ponderomotive potential or the ionization energy needs to be increased. The ionization energy can be increased if a different noble gas is used (see in Tab. 5.1).

The ponderomotive potential can be increased if longer wavelengths or higher intensities are used. However, higher intensity laser field can cause more problems during phase matching due to higher gas ionization and also the generated magnetic field (relativistic effects) cannot be neglected. Both these things reduce the probability of the electron returning to the original ion. The relativistic effects

lead to a condition

$$\dot{x}^2 = \frac{e^2 E_0^2}{m^2 w^2} \ll c^2 \quad \Rightarrow \quad \frac{eE_0}{m\omega c} \ll 1$$

The importance of the ionization can be expressed by the Keldysh parameter κ [155]

$$\kappa = \frac{I_p}{2U_p} \quad .$$

In the case of $\kappa > 1$, atomic processes such as ionization and multi-photon effects will dominate while for $\kappa < 1$, the tunnelling effect dominates. On the other hand, higher driving beam intensity should result in higher EUV intensity. Because HHG is a strongly nonlinear process, the generated EUV intensity is proportional to I^5 of driving laser intensity [156], although the dependence is more complicated if phase matching is included.

Another option to increase the cut-off energy is to use longer wavelength of the driving laser. This effectively increases the distance of the electron from atom during the HHG acceleration phase and allows to gain more energy. The drawback is lower probability of the electron returning back to the original atom which results in higher ionization and lower amount of the produced EUV photons. Experimentally, it has been shown that the efficiency of harmonics production scales with factor $\sim \lambda^{-6}$ of the HHG wavelength [157] (for Kr and Xe). Therefore, higher HHG energies can be achieved for the price of much lower efficiency. The conversion efficiency of the HHG process can be up to 10^{-6} for argon and 10^{-7} for neon [158].

Another effect that needs to be considered is for example different effective cross-section of the HHG tunnelling process for different gases which also significantly changes amount of the produced photons.

A more precise description of the HHG process is possible if the quantum effects are considered within the strong field approximation [156], however the simple 3-step model is sufficient to predict some of the main properties of the HHG process.

5.4 Phase matching

The 3-step model can help to predict the achievable photon energies, however it cannot account for the harmonic yield because propagation effects need to be considered. Harmonic processes need to fulfil phase-matching condition between the fundamental $\vec{k}_f$ and harmonic $\vec{k}_q$ wave vectors [21, 135]

$$\Delta\vec{k} = \vec{k}_q - q\vec{k}_f \approx 0 \quad ,$$

where q is the harmonic order. In the case of HHG, the main effects are [155, 159]:

- dispersion from neutral gas atoms $\Delta\vec{k}_d$
- dispersion from the created plasma $\Delta\vec{k}_p$
- geometrical effects such as Gouy shift $\Delta\vec{k}_g$ in the free space propagation geometry [155, 159, 160] or wave-guiding effects $\Delta\vec{k}_c$ in capillaries [161–164]
- atomic phase difference $\Delta\vec{k}_a$, which is caused by different path length for electrons in continuum [155] as shown in Fig. 5.2

For the standard low harmonic processes (SHG, THG, ...), the main contributions are dispersion in a non-linear material and the Gouy shift. Therefore, the second harmonic yield is usually maximized by optimization of the angle of the optical axis with respect to the beam. This changes the extraordinary and ordinary waves propagation velocity so that the SHG light is propagating at the same phase velocity as the fundamental beam .

The generated harmonic intensity is in the low depletion and low attenuation approximation proportional to

$$I(z) \sim \frac{\sin^2(\Delta k z/2)}{(\Delta k z)^2} z^2 \quad , \tag{5.1}$$

Therefore, the harmonic yield grows quadratically with the distance for very low dispersion or thin targets. On the other hand, the generated harmonic intensity goes to 0 at position $2L_c$ (L_c is the coherence length defined in Eq. 4.23) for weakly attenuating samples, therefore the phase mismatch Δk has to be minimized. In order to achieve high gain, it is preferable to have the coherence length L_c slightly longer then the attenuation length.

5.4.1 Phase matching in HHG

The phase matching in HHG can be described by four terms [155]

$$\Delta\vec{k} = \Delta\vec{k}_g + \Delta\vec{k}_p + \Delta\vec{k}_{\text{geom}} + \Delta\vec{k}_a \quad ,$$

where $\Delta\vec{k}_g$ (negative) is effect of neutral gas, $\Delta\vec{k}_p$ (positive) is caused by plasma dispersion, $\Delta\vec{k}_{\text{geom}}$ (positive) is caused by geometry effects, either by Gouy shift in free space propagation or by waveguides effects. Difference between capillary and free space propagation geometry is shown in Fig. 5.5. Finally, $\Delta\vec{k}_a$ is caused by atomic dipole phase and its sign is not fixed and depends for example on relative position of gas and beam focus.

Note that we will neglect the vector nature of the phase-matching term $\Delta\vec{k}$ in the rest of the chapter because the presented approximations are either valid only along the beam propagation axis ($\Delta\vec{k}_{\text{geom}}, \Delta\vec{k}_a$) or the phase-matching term is isotropic ($\Delta\vec{k}_g, \Delta\vec{k}_p$). A more complex model would be required to include full 3D propagation properties [159].

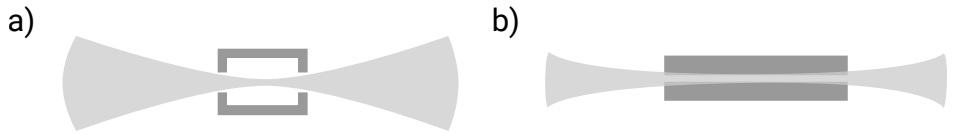

FIGURE 5.5: Examples of two common phase-matching geometries for HHG. Image (a) shows free space propagation geometry, when the beam propagates through a metal gas-cell filled by interaction gas. The second example (b) shows a capillary filled by gas that serves as a waveguide constraining the beam. The waveguide is enforcing identical beam properties along the capillary length.

5.4.2 Neutral gas dispersion

Dispersion in neutral gas Δk_g depends only on the gas composition and pressure, therefore it can be usually easily calculated as [165, 166]

$$\Delta k_{\text{g}} = -q\frac{P}{P_{\text{atm}}}(1-\eta)\frac{2\pi}{\lambda_f}(n_f - n_{\text{EUV}}) \quad ,$$

where q denotes the harmonic order, η is ionization fraction, P denotes the working gas pressure and P_{atm} is atmospheric pressure. n_f, n_{EUV} are refractive indices of the neutral gas for fundamental and produced harmonic wavelengths. Refractive index

of ions was approximated by 1, therefore they do not affect the phase-matching. The effect of the free electrons is included in the next section.

In the case of very high laser intensity, the refractive index is changed by

$$\Delta n = n_2 I \quad ,$$

where n_2 is nonlinear refractive index. This effect can result in self focusing of the beam that is changing the beam properties during propagation. In the case of low beam quality (M^2), this effect can result in lowered HHG beam pointing stability if the pulse length or beam power are varying.

5.4.3 Plasma dispersion

Dispersion in plasma Δk_p can be calculated for the fundamental beam as

$$\Delta k_p = -q\frac{2\pi(n_p - 1)}{\lambda_f},$$

where n_p denotes refractive index of cold plasma for the fundamental wavelength λ_f. The refractive index of plasma n_p can be expressed as [167]

$$n_p = \sqrt{1 - \frac{\omega_p^2}{\omega_f^2}} \approx 1 - \frac{1}{2}\frac{\omega_p^2}{\omega_f^2}$$

with the plasma frequency ω_p defined as

$$\omega_p = \sqrt{\frac{e^2 n_e}{m_e \varepsilon_0}}$$

where e, $n_e = \eta n_g$ and m_e denotes electron charge, free electron density, electron mass and η is ionization fraction of the neutral gas density n_g. The phase shift caused by the plasma can be neglected for the HHG EUV beam. The approximation of the plasma phase matching term Δk_p can be written as [168]

$$\Delta k_p \approx q\frac{\pi\left(\frac{\frac{e^2 n_e}{m_e \varepsilon_0}}{\omega_f^2}\right)}{\lambda_f} = q\eta n_e r_e \lambda_f$$

where r_e is the classical electron radius given by

$$r_e = \frac{1}{4\pi\varepsilon_0}\frac{e^2}{m_e c^2} \quad ,$$

Δk_p therefore depends on the gas pressure and the ionization fraction. Note that the ionization fraction can vary during the pulse duration if the driving pulses are too long.

The plasma dispersion can be used to compensate the neutral gas dispersion. However, the ionization cannot exceed the level, where the dispersion from neutral gas atoms is lower than the effects of plasma. In that case, the phase matching is not possible anymore. The critical ionization fractions for 800 nm fundamental wavelength can be estimated for capillaries roughly 0.5% in He, 1% in Ne and 4% in Ar [168]. The HHG process naturally produces plasma if the electron does not successfully return to the source atom and the HHG process is most efficient close to the gas ionization level as well. These pose severe limitations for short wavelength HHG production in helium and neon.

5.4.4 Geometric effects

Another important effect in the free space propagating geometry is the Gouy shift that can be adjusted by changing the divergence of the fundamental beam. In the free space propagation geometry that we use, the Gouy shift is used mainly to compensate for the atomic phase effects. The phase shift caused by the Gouy shift along the beam axis with the position z is equal to [159, 160]

$$\phi_{\text{Gouy}}(z) = \arctan\left(\frac{z}{z_R}\right) \quad ,$$

where z_R is the Rayleigh range. The local phase-matching Δk_{Gouy} depends on gradient of the Gouy shift

$$\Delta k_{\text{Gouy}} \approx q\frac{z_R}{z^2 + z_R^2}$$

and therefore, it varies with the position along focus. Note that the Gouy shift for the HHG beam is negligible compared to the fundamental wavelength. Close to the focus the dispersion term can be approximated as

$$\Delta k_{\text{Gouy}} = \frac{q}{z_R} \quad .$$

The contribution of the Gouy shift can be controlled by using a different focal length, changing size of the beam and thus the focused spot size and Rayleigh range or by placing the gas medium out of focus.

In the case of phase matching in capillary, the phase mismatch effect can be written as [161]

$$\Delta k_{\mathrm{cap}} = \frac{q u_{11}^2 \lambda_f}{4\pi a^2} \quad ,$$

where a is waveguide radius and u_{11} is the position of the first zero of the Bessel function J_0. The phase shift effect is also always positive as the Gouy shift but it stays constant during the propagation in capillary and therefore phase-matching can be achieved over longer distances.

5.4.5 Atomic phase effects

The atomic phase effects originate from different phases acquired due to different trajectories of accelerated electrons before they produce the EUV photons. The atomic phase and the Gouy shift are dominating effects if the neutral gas pressure is sufficiently low. The atomic phase depends on intensity of the driving field and on the trajectory of the electrons as it is shown in Fig. 5.6. The phase shift is therefore changing through the focus and if the atomic phase effects are not negligible, the phase matching depends on position of focus relative to the medium and also on the radial position relative to the optical axis.

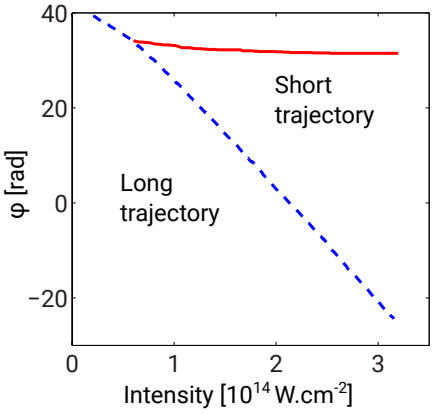

FIGURE 5.6: An example of the atomic phase dependence for the short and long trajectories. Adapted from [8].

If the atomic phase dependence on beam intensity is approximately linear and the intensity along the optical axis is estimated from the Gaussian beam as

$$I(z) = \frac{I_0}{1 + \left(\frac{2z}{z_R}\right)^2} \quad ,$$

then the dispersion term Δk_a can be estimated as [8]

$$\Delta k_a = -\alpha_q \frac{q\lambda_f}{2\pi}\frac{dI}{dz} = \alpha_q I_0 \frac{4q\lambda_f \pi}{z_R^2\left[1 + \left(\frac{2z}{z_R}\right)^2\right]^2} \quad ,$$

where $\alpha_q \approx$ 1-5$\cdot\ 10^{-14}$cm^{2}W^{-1} for the short trajectory and $\alpha_q \approx$ 20-25$\cdot 10^{-14}$cm^{2}W^{-1} for the long trajectory [8].

In our experiment, argon gas is used as the nonlinear medium. We use a ≈3 mm thick gas cell (see Fig. 5.5). The beam beam divergence is relatively low (NA$<$ 0.02) in order to increase beam spot size, avoid too high gas ionization in the focus and limit the effect of the Gouy shift.

5.4.6 Reabsorption in generating medium

The generated harmonics are relatively strongly attenuated in the generating gases. Moreover, the nonlinear part of the refraction index may not be negligible and therefore it can cause self-focusing of the driving beam. From this reasons, it is advantageous to keep the interaction path as short as possible because there is only small additional signal gain for interaction length longer than 3$\times$ the attenuation length (Fig. 5.7). The recommended coherence length should be according to Ref. [169] at least five times longer than the medium length, however in practice it is usually comparable with the medium length. Figure 5.7 shows that the absorption length of the generated EUV is significantly changing in dependence on the used gas.

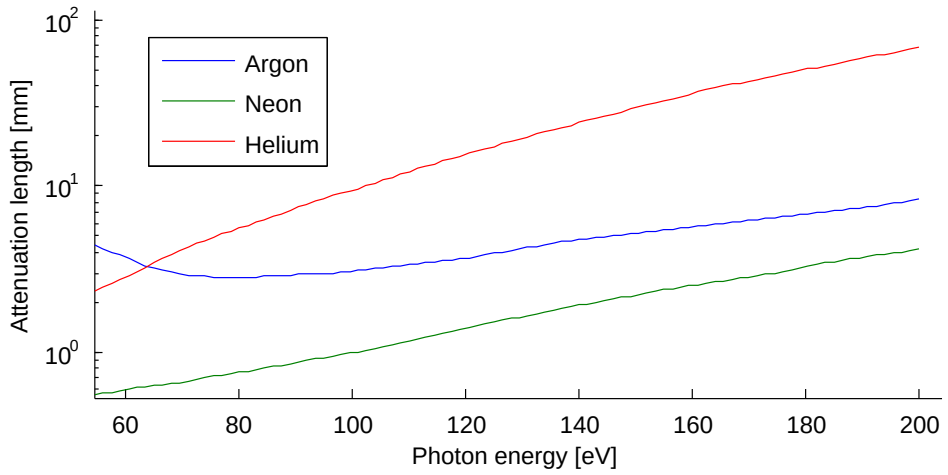

FIGURE 5.7: Attenuation length of different rare gases used commonly for HHG for pressure 100 mbar [9]

.

5.5 Experimental setup for HHG imaging

Laser system

A simplified schema of the laser system is shown in Fig. 5.8. The ultrafast seeding pulses with 30 fs length and 80 MHz repetition rate are generated in the Tsunami Ti:Saphire oscillator. Only one pulse per millisecond is selected by the Spitfire regenerative chirped pulse amplifier and the pulse energy is increased from 37 nJ to up to 3 mJ. The amplifier Spitfire is a regenerative amplifier and therefore the resulting beam is supposed to have relatively good pointing stability. The seeding pulse is chirped and stretched and then the pulse passes 10 times through a cryogenically cooled Ti-Sapphire gain crystal that is pumped by Evolution laser with 17 W of power. The number of passes needs to be optimized in order to saturate the gain crystal and compensate for higher order dispersion in the system. After compression, the system is designed to reach up to 40 fs pulses with 1 kHz repetition rate. However, the practically available parameters in our experimental setup are rather 50 fs pulse length and roughly 1.8 W power (1.8 mJ pulse energy) at the gas cell position. The peak power density at the focus position is $2.3{\cdot}10^{12}$ Wmm^{-2} if a 75 cm focal length lens is used.

One disadvantage of the regenerative amplifier Spitfire is relatively slow feedback of the cryosystem used to cool the gain crystal which results in periodic beam position, intensity and pulse length fluctuations with period roughly 10-20 s. In order to correct the beam movement, a beam stabilization system was developed

(see Appendix A), however the effect of the pulse length variation still has non negligible effect on the HHG beam pointing stability.

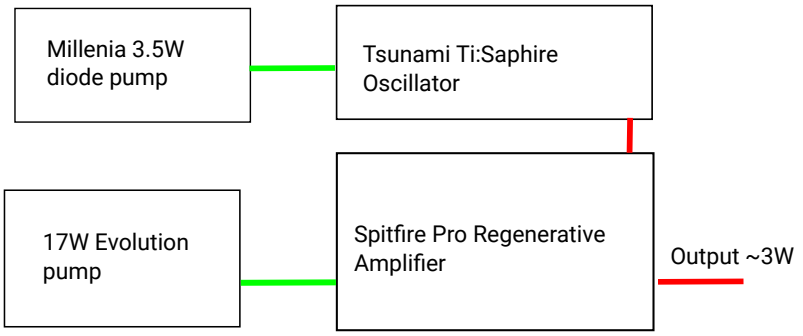

FIGURE 5.8: An illustration of the laser amplifying system. The Millennium laser is pumping the Tsunami Ti:Sapphire oscillator, which generates mode-locked laser beam centred at 780 nm wavelength with 3 W average power and 80 MHz rate. Spitfire amplifier selects only one pulse per millisecond and further amplifies it to pulse energy up to 3 mJ and pulse length ≈50 fs.

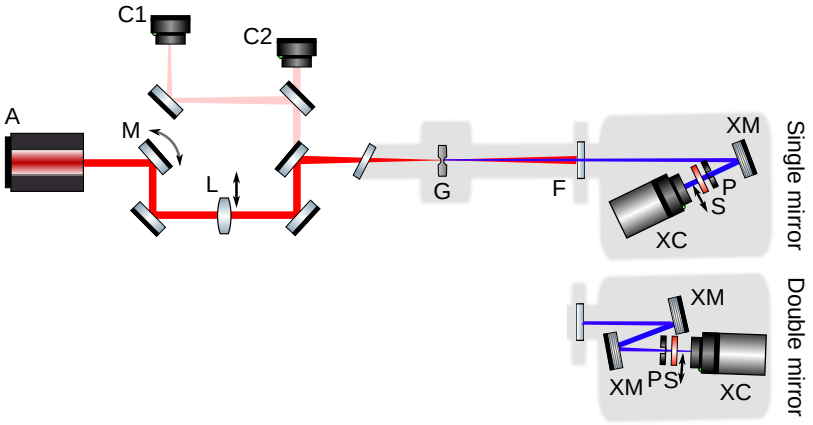

FIGURE 5.9: A simplified schema of our current HHG imaging setup. The IR laser (A) beam is controlled by a custom made active stabilization system using a stabilization mirror (M), a stabilization lens (L) and two CMOS cameras (C1,C2). Beam is focused by a 75 cm lens (L) onto a gas-cell (G), where EUV is generated and the driving IR beam is filtered out by a single aluminium 200 nm thick filter (F). The EUV beam is further spectrally filtered and focused by one or two multilayer mirrors (XM) onto a pinhole (P). The spatially modulated illumination propagates on a sample (S) and is collected by an EUV sensitive camera (XC).

Imaging beamline

The amplified IR beam is further redirected to different beamlines. The coherent diffractive imaging beamline is placed roughly 5 m from the laser amplifier. The

schema of the imaging beamline in shown in Fig. 5.9. The driving laser beam is focused down to a gas cell by a 75 cm focal length lens. The focal length was selected in order to optimize the Gouy shift and thus achieve good phase matching of the fundamental and HHG beam. The focal spot size depends on the focal length of the focusing optics and the beam quality. The beam quality was measured to be $M^2 = 1.8$ [170] and the focal spot diameter is FWHM=70 μm. The beam diameter and thus the effective NA and Gouy shift were finely tuned by an iris placed roughly 2 m upstream the gas cell. The ≈3 mm long gas cell was manufactured from an aluminium pipe. The gas outlets were drilled by the ultrafast IR laser beam in a thin copper tape in order to minimize the amount of leaking gas. The gas cell was filled with argon gas at 80-120 mbar pressure. Gas pressure and relative distance of the beam focus to the gas cell were optimized in order to preserve circular shape of the generated HHG beam. We have empirically tested that close to a circularly symmetric HHG beam shape (see Fig. 5.10) provides the best coherence with still relatively high flux, which is needed in order to get high coherent photon flux density on the sample.

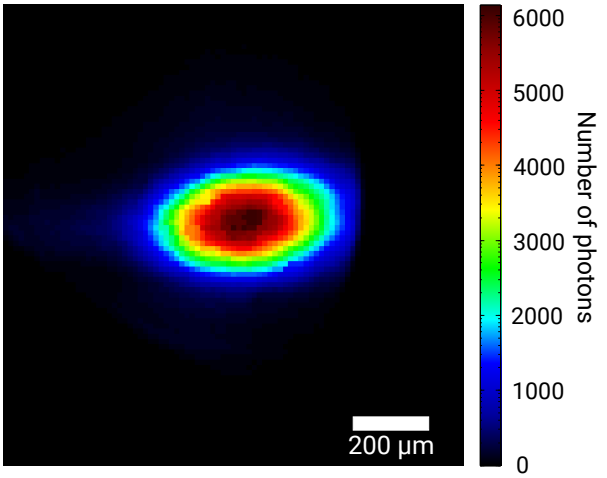

FIGURE 5.10: An example of the HHG EUV beam spot measured at the camera placed 10 cm behind focus of the EUV beam. The exposure time was 50 ms.

The background pressure along the EUV beam path should be below 10^{-3} mbar, because EUV light is attenuated even by the residual gas. Because the driving laser beam needs to be blocked, we are using a single 200 nm thick aluminium filter placed 1.2 m downstream the gas-cell. The single aluminium foil should have theoretically sufficiently attenuate the IR light, however if the laser power density on the filter is too high, micro-pinholes can be created in the filter and low amount of IR light can be transmitted into the imaging chamber. IR transmission

of a single defectless 200 nm aluminium foil including the reflection losses is in the IR range below $1.3{\cdot}10^{-12}$ [171]. The efficiency of the HHG generation is approximately 10^{-6}, therefore the extinction ratio is sufficient. Moreover, reflectivity of the aluminium is 87% for 800 nm wavelength, therefore the absorbed beam power is relatively low. In combination with good heat conductivity 205 W/mK, aluminium is suitable filtering material. Aluminium filters are transmissive in EUV range from 17 nm to 80 nm (see Fig 5.11). Additionally, the EUV light is also filtered by the argon gas that is strongly absorbing wavelengths above 30 nm. The high pressure argon gas in the gas cell is also creating a natural spectral filter that has optimum around 29 nm. That is the reason, why the 27th harmonic (29.6 nm) was selected for imaging.

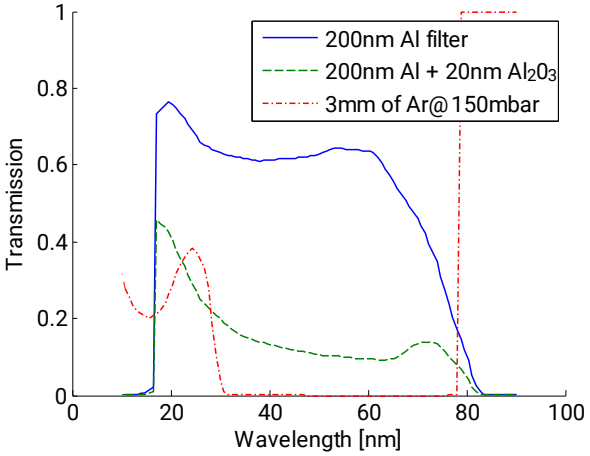

FIGURE 5.11: Transmission of the a single 200 nm aluminium filter with and without oxide layer and transmission of Ar gas [9].

EUV CDI microscope

Different mirror configurations can be used for HHG focusing and filtering as it is shown in Fig. 5.9. The double mirror setup has the advantage of better spectral filtering and smaller spot size in focus due to lower astigmatism that is determined by the folding angle of the focusing mirror. The single mirror setup has the advantage of higher flux because it avoids high losses caused by each reflection of the EUV beam. The peak reflectivity of the used B_4C/Si multilayer mirror is around 30%

Focal length of the used spherical multilayer mirror was 25 cm. The spot size in CLC (circle of least confusion) was limited by the astigmatism caused by the

folding angle of 7°. The spot size was estimated from the ptychography reconstructions to roughly 13 μm FWHM. The beam was spatially filtered by a small pinhole (7-10 μm) placed precisely at the CLC position. The pinholes were cut by focused ion beam (FIB) into a 50 nm thick silicon nitride foil covered by 100 nm of gold. The imaged sample was placed between 50-100 μm downstream the pinhole. A distance of 40 μm was found to be the minimal distance before the sample and pinhole started to touch because they are not perfectly parallel. The sample to probe distance needs to be also kept as low as possible, preferably close to the near-field regime, otherwise the high frequencies from the probe diffraction could not be reconstructed by the ptychography method (see Section 5.7). The acceptable limit depends on the oversampling and resolution of the collected data. From the basic geometry constraints that are required to provide at least sampling at the Nyquist frequency of the propagated exit-wave $\psi = P \cdot O$ at the detector plane, the maximal pinhole (or diffuser) to sample distance should be roughly

$$z \approx 0.25 \frac{D}{\mathrm{NA}} \quad , \tag{5.2}$$

where D is size of the constructed region around probe and NA is the numerical aperture determined by the detector size. For example for $E = 42$ eV, resolution 50 nm, FOV= 5 μm, the pinhole sample distance should be below 25 μm. This was not fulfilled in our HHG experiments, where the sample to pinhole distance was between 50-100 μm due to experimental constraints.

The pinhole and the sample mount were both mounted on a shared stage that was positioned by low precision closed loop Newport servomotors with repeatability roughly 2 μm. The distance and leverage between the sample and pinhole were minimized in order to avoid thermal drifts and vibrations. The sample was further mounted on a 3D nanoprecision piezoelectric stage Smaract SLC-1730 with 26 mm travel range and for our sample position the position reliability was better than 100 nm. More information about the precision and repeatability of these stages is shown in Appendix B. The distance between the sample and pinhole was estimated using a USB Dinolite microscope placed in vacuum.

The sample was illuminated by structured illumination resulting from near-field propagation of the cropped EUV beam. The scattered light from the sample was collected by an EUV sensitive back-illuminated CCD camera (Andor DX434). The used camera has 1024×1024 pixels with 13 μm size. The quantum efficiency

at 30 nm is 30%. However, it is possible to get the quantum efficiency around 80% with more modern CCD chips. Cooling of the camera was set to -50 ℃ during the experiments, in order to prevent accumulation of the dark (thermal) current that leads to additional shot noise. The Andor DX434 camera supports only pixel readout speeds of 31.25 kHz, 62.5 kHz, 500 kHz and 1 MHz. Readout speeds 62.5 kHz, 500 kHz were used for most of our experiments. Dynamic range of the Andor camera is relatively low, the pixel well depth is 11000 counts which corresponds to $2{\cdot}10^3$ photons at 42 eV. This may be sufficient to get high resolution reconstruction of simple highly scattering binary objects, however it is not enough for any real samples. In order to deal with it, HDR imaging by exposing the same region of the sample using different exposure times was implemented. The acquisition time was reduced by readout of only the over-saturated regions in the subsequent exposures as it was shown in Section 3.1. This allowed to extend the dynamic range by 3 orders of magnitude. Further extension was limited by artefacts in the high frequency region caused by moving large charge across the CCD. The acquisition time was increased by roughly factor of 2 compared to a single exposure. However, further extension of the exposure time could be achieved using a fast dynamic beamblock that would prevent the strong over-saturation for the longest exposure. Collected diffraction patterns were background subtracted and corrected for the curvature of the Ewald's sphere (Section 2.1.4).

5.5.1 EUV illumination

The EUV beam was cropped by a pinhole placed in the circle of least confusion (CLC) position. Near-field propagation of the cropped beam created a complex structured illumination on the sample. An example of the probe propagation is shown in Fig. 5.12.

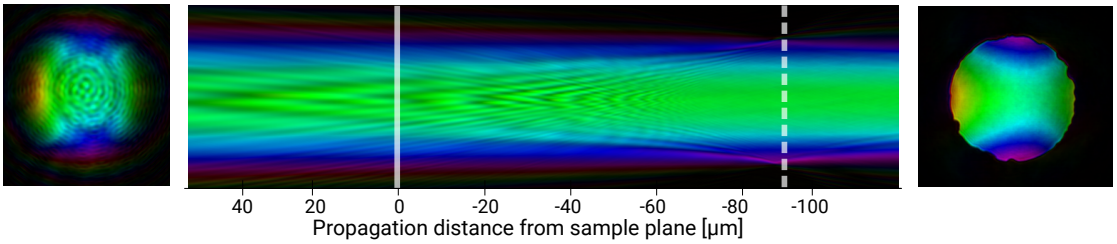

FIGURE 5.12: An example of numerical refocusing of the reconstructed complex electric field on the sample (left) into the beam cropping aperture with 10 μm diameter (right) positioned 92 μm upstream the sample.

The EUV photon flux on the sample was usually around $6{\cdot}10^8$ photons/s. The beam was only partially spectrally filtered by the single multilayer mirror, therefore the illumination could not be considered monochromatic. The spectrum had to be known if the PolyCDI method for the ptychography reconstruction (Section 2.6.1) is applied. The main reason why we have used only a single multilayer mirror is the already mentioned low reflectivity in the EUV wavelength region. The polychromatic spectrum increases the total photon flux density incident on the sample roughly by factor of 5 [17].

Note that the used PolyCDI method [51] is precisely valid only for achromatic samples and flat illumination wavefront. If these conditions are not fulfilled, full multicolour ptychography [52] should be used. However, biological materials can be considered achromatic in the used wavelength range and illumination wavefront is relatively smooth, therefore the approximation can be used.

The precise spectrum depends on the spectral characteristics of the mirror, geometrical factors such as folding angle and also on the phase-matching process. Therefore, the spectrum should be measured in the used experimental setup. We have spectrally characterized the beam by analysis of a diffraction pattern from the Young's slits. There have been already demonstrated several methods for spectrum reconstruction from the measured diffraction pattern that were based on the Fourier transformation [172, 173] and also on maximum entropy method [174, 175].

We have used another method based on the Tikhonov regularization [110] with positivity constraint. Known parameters of the slits i.e. width b, separation d and distance from camera Z can be used to calculate the Young's slits pattern for a single wavelength λ

$$I(x) \propto \left[1 + \gamma \cos\left(\frac{2\pi d(x-\mu)}{Z\lambda}\right)\right] \operatorname{sinc}^2\left(\frac{\pi b x}{Z\lambda}\right) \quad , \tag{5.3}$$

where γ denotes visibility and μ is offset of the centre of the pattern. Incoherent diffraction pattern in the far-field is proportional to an incoherent sum of multiple monochromatic patterns. Therefore, if the single monochromatic pattern is known I_λ and the incoherent sum is measured $\hat{I}$, the spectrum S can be recovered from a system of linear equations

$$\mathbf{I}_\lambda S = \hat{I} \quad . \tag{5.4}$$

This system is over-constrained if the number of bins of the spectrum is lower than the number of measured pixels, however it is highly ill-posed and therefore it cannot be solved directly by the linear least-squares method. We have used Tikhonov regularization [110, 112] with a nonlinear regularization term. The equation 5.4 can be solved as

$$\left(\mathbf{I}_\lambda^T \mathbf{I}_\lambda + \alpha \frac{\mathbb{I}}{\mathbf{S}_+}\right) \mathbf{S} = \mathbf{I}_\lambda^T \hat{I} \quad , \tag{5.5}$$

where $\mathbb{I}$ denotes identity matrix and $\mathbf{S}_+$ denotes positive part of $\mathbf{S}$. Because the system is nonlinear, it needs to be solved iteratively. The optimal regularization parameter α was selected using the L-curve criterion [113]. A reconstruction of the spectrum is shown in Fig.5.13 b).

The illumination bandwidth around the central wavelength $\hat{\lambda}$ can be expressed as

$$\frac{\Delta\lambda}{\hat{\lambda}} = \frac{1}{\hat{\lambda}} \sqrt{\frac{\sum_i S_i (\lambda_i - \hat{\lambda})^2}{\sum_i S_i}} \quad , \tag{5.6}$$

which results into effective bandwidth 20%. Therefore, the highest spatial frequency $d_{\min}$ that is not affected by the bandwidth incoherence effects is equal to [48]

$$d_{\min} = \frac{\Delta\lambda}{\lambda} \frac{W}{2} \quad , \tag{5.7}$$

where W denotes the size of the illumination. Using a $10\,\mu$m pinhole, this theoretically limits the achievable CDI resolution to mere $1\,\mu$m. In figure 5.13, one can see that the measured diffraction pattern deviates from the monochromatic simulation already for low angles $\approx 30\,$mrad, however there are still sharp speckles over the whole collected region. This is mainly possible because the spectrum is polychromatic rather than broadband. Moreover, each of the spectral peaks is very narrow thanks to relatively long pulses of our IR driving laser. The diffraction-limited resolution of our Young's slits spectrometer is 0.5 nm, however we expect that the peaks are significantly more narrow, because we have observed sharp speckles for NA$>$ 0.4.

Figure 5.14 shows the visibility of the Young's slits diffraction pattern for different separations between the slits. Visibility is defined as

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \quad , \tag{5.8}$$

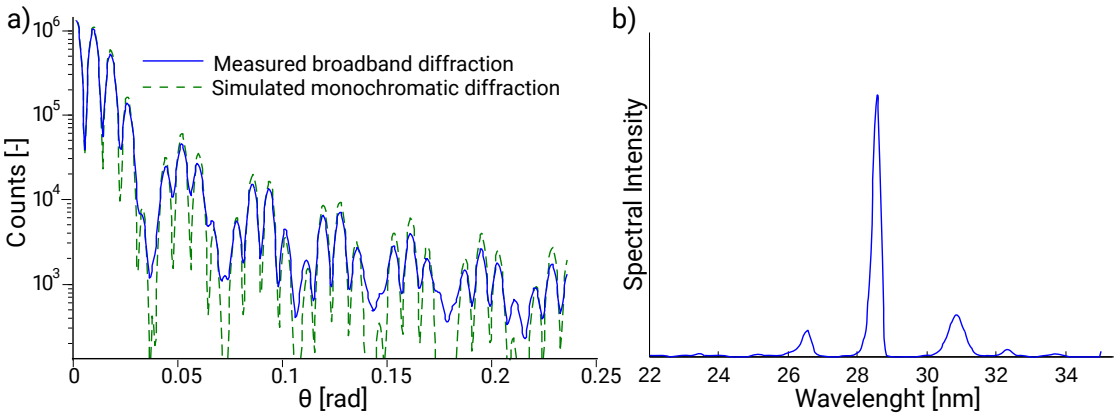

FIGURE 5.13: (a) Measured diffraction from 4 μm separation slits and simulated monochromatic pattern with identical spatial coherence and other experimental parameters. The reconstructed spectrum is shown in (b). Figures adapted from [10].

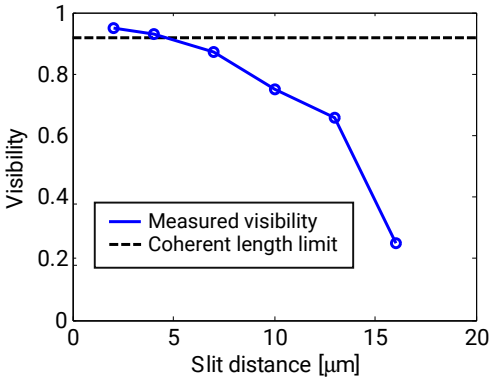

FIGURE 5.14: Measurement of visibility of the Young's slits diffraction patterns for different slit separations. Figure adapted from Ref. [11].

where $I_{\max}$ denotes maximum of the slits diffraction pattern and $I_{\min}$ is the nearest minimum. If we use an estimation that visibility $V = 0.88$ corresponds to one coherence length [41], then the lateral coherence length of our beam in focus is roughly 6 μm. Therefore, according to the CDI requirements presented in Ref. [48], the maximum pinhole size should be 3 μm, however as we will present in this chapter, ptychography with $\approx 70\%$ visibility and polychromatic illumination can be still used with the right relaxation methods.

The HHG CDI setup presented in this chapter was built in cooperation with Peter Baksh from University of Southampton and the presented data were also collected by Peter Baksh. However, I prepared all the data acquisition and automatisation scripts, created the laser stabilization system and the ptychography reconstruction code.

5.6 Ptychography reconstruction of high contrast test patterns

Our high contrast test sample was produced by H.S. Kim from Aachen university using focused ion beam (FIB) milling. A scanning electron microscope (SEM) image of the sample is shown in Fig. 5.15(a). The sample is an aperiodic grid that was covered by 100 nm gold layer. The transmissive region has a size $35 \times 35\,\mu$m and contains up to 50 bars, however their distance and shape is changing significantly. Results in this chapter were published in Ref. [10].

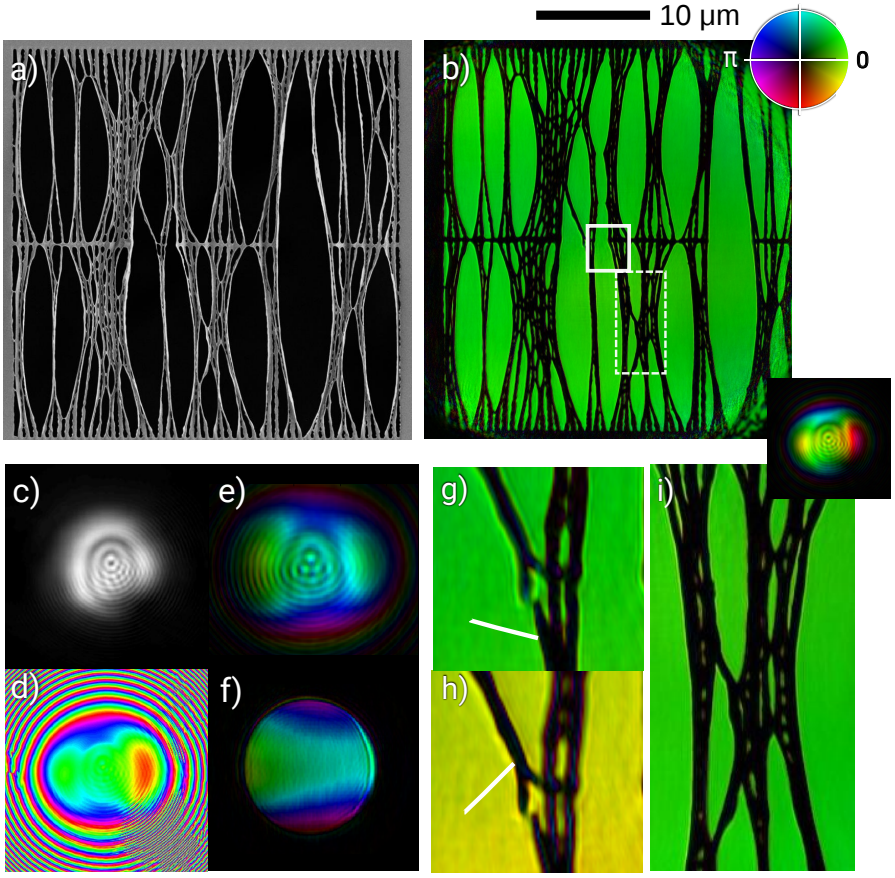

FIGURE 5.15: Image (a) shows a scanning electron microscope (SEM) image of the aperiodic grid sample, (b) is a ptychography reconstruction in complex colour scale and the inset contains the probe reconstruction at scale. (c,d) are amplitude and phase of the reconstructed probe and (e,f) contain the reconstructed complex amplitude of the probe at the sample position and back-propagated 95.0 μm to the filtering pinhole. (g) shows a magnified region of the sample with clear defocus effect and (h) is the same region after numerical refocusing ($-1.9\,\mu$m). (i) An example of a region that is entirely in focus.

The sample was illuminated through a 7 μm pinhole placed at the focus of the EUV beam. The scanning step size was 2 μm, therefore the linear overlap between

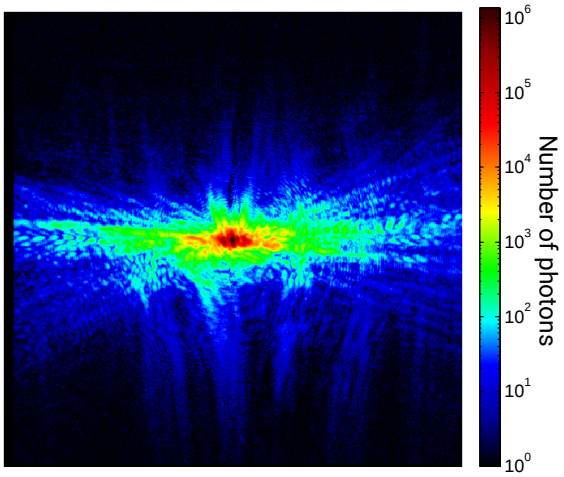

FIGURE 5.16: An example of the diffraction pattern intensity in logarithmic color scale. The signal is photon limited in the vertical direction and NA limited in the horizontal direction.

adjacent positions was ≈70%. We have collected 578 high dynamic range diffraction patterns with the maximum exposure 0.5 s and the minimum ≈10 ms with ratio between the subsequent exposures equal to 5. The total exposure time was less than 6 minutes. The pixel readout speed of the camera was 62.6 kHz. The distance between the sample and camera was estimated to 25 mm resulting in the pixel size in the reconstruction of 60 nm. The precise pixel size 56 nm was refined during the reconstruction process by the gradient method (Section 2.6.2).

The data were mapped onto the Ewald sphere and reconstructed using the ePIE based OPRP method with polychromatic and lateral incoherence correction. Relaxed object positivity and relaxed probe support constraints were used. Linear oversampling of the diffraction pattern was in this case roughly 5, therefore the data pixels were 2×2 binned before the reconstruction.

The final ptychography reconstruction is shown in the Fig. 5.15(b). The reconstructed shape corresponds well to the SEM image. However, there is a small difference around some of the sharp edges (Fig. 5.15(g)). There are two possible reasons that could have caused this: insufficient flatness of the sample results in position variation larger than the depth of field or non zero intensity of the diffraction at the edges of the diffraction pattern (see Fig. 5.16) could result in fringing and small artefacts in the horizontal direction [176]. The insufficient flatness of the sample is a limitation if the flatness variations exceed the depth of focus (DoF) (Eq. 2.37)

$$\mathrm{DoF} = \frac{\lambda}{\mathrm{NA}^2} \quad . \tag{5.9}$$

The DoF was estimated to be 200 nm in our experimental geometry. Therefore the sample to pinhole distance over the scanned regions has to vary less than 200 nm. This poses a serious limitation for large field of view scans, in particular in combination with high collected NA. However, because our sample is relatively thin (≈150 nm) and most of the variation is expected due to bending of the sample, the reconstructed object can be numerically refocused into the local sample position using the ASM method (see Fig. 5.15(g,h)). Note that it would not be possible if scattering from previous layers of the sample could not be neglected. In that case, the multislice ptychography method [25] would need to be used.

Reconstruction of the illumination probe at the sample and after back-propagation 95 μm at the pinhole position using the angular spectrum method (ASM) is shown in Fig. 5.15(e,f). The left side of the probe is slightly out of focus, which can be explained by slightly non perpendicular pinhole to the beam. The saddle-shaped wavefront caused by the astigmatic beam is clearly visible behind the aperture.

5.6.1 Reconstruction quality estimation

The reconstruction quality is evaluated in Fig. 5.17 using all the methods introduced in Section 2.2.5. The first image (a) shows an example of the knife edge test. The reconstructed amplitude (stars) is slightly out of the imaging plane (see for example Fig. 5.15(g,h)) and the circles are showing locally numerically refocused reconstruction. The refocused points were fitted by the complementary error function (ERFC) and the 10-90% distance was used to estimate resolution. This was repeated in multiple regions and the average horizontal half-period resolution was 58±5 nm, which indicates that the resolution was mainly limited by the experimental geometry (i.e. low collected NA).

The second method used for the quality estimation is the PRTF in Fig. 5.17(b). As it was mentioned, PRTF was not designed for ptychography but it gives reasonable results for highly scattering samples. We have done PRTF separately for the vertical and horizontal directions because of significant anisotropy of the features on our sample. The error bars were estimated from the variation of the PRTF for different scanning positions. The vertical resolution is flux limited and in the case of threshold $T = 0.5$, the resolution is 90 ± 20 nm. In the horizontal direction, the resolution is limited by the detector size and the resolution is equal to the pixel size. However, PRFT rather overestimates the real resolution and the real resolution is

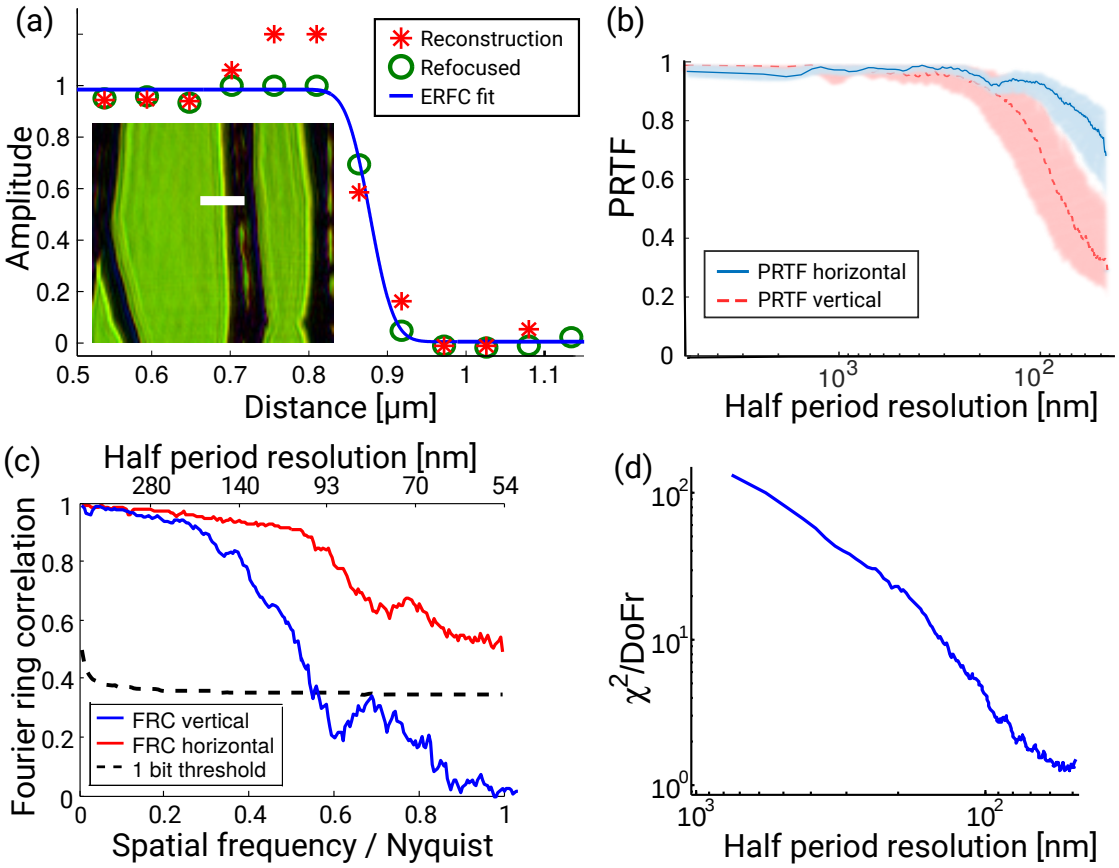

FIGURE 5.17: Examples of different methods for estimation of the reconstruction quality. Image (a) shows the knife-edge method using ERFC fit of the amplitude, (b) phase retrieval transfer function (PRTF) [12, 13] of the reconstructed exit-waves, (c) Fourier ring correlation (FRC) [14] between two independent reconstructions, (d) Goodness of fit χ^2 normalized by degrees of freedom (DoFr) should be close to one for different spatial frequencies for a good fit by the optimal model according to the Morozovs discrepancy principle [15]. Figures were adapted from [10].

usually worse (see e.g. [29]). Therefore, the FRC method was used to compare two independent reconstructions (Fig. 5.17(c)). The 1-bit threshold [60] is giving the resolution limit in the vertical direction of roughly 100 nm and in the horizontal direction it agrees with PRTF and knife-edge that the reconstruction quality is limited by the experimental geometry. Finally, figure 5.17(d) shows the standard goodness of fit criterion using the noise level calculated from the known shot noise of the camera and the Poisson noise of the detected signal. The error distribution shows that most of the average misfit was in the low spatial frequencies region, while the high spatial frequencies were fitted significantly better. The possible explanations are:

1. Systematic errors e.g. from HDR stitching, beam movements, intensity fluctuations and coherence estimation are dominating in the low spatial frequency region

2. Ptychography is constraining more strongly the low spatial frequencies which leads to higher misfit due to different systematic errors

3. The signal to noise ratio (SNR) was lower for the high spatial frequencies due to non negligible readout noise, so the "poor fit" caused by the systematic errors was hidden in the higher noise.

On the other hand, thanks to the ptychography over-constraining, the systematic errors have limited effects on the final reconstruction.

Illumination stability

The illumination stability can be assessed from the OPR eigen probes shown in Fig. 5.18. The relative intensity of the eigenprobes drops fast enough so that the OPR method can be used. The second row in the Fig. 5.18 shows the eigenprobes numerically back-propagated to the pinhole position. From the back-propagated probe, it can be seen that first 5 modes are mostly related to the beam movements because they contain intensity inside the pinhole support. On the contrary, the last 2 are related to the sub-pixel shift corrections because they contain mostly only intensity around the edge regions. The last row shows Fourier space incoherent modes [16] that can be easily recovered as well. The incoherent modes are related to the coherence of the illumination beam. Thanks to relatively small pinhole ($7\,\mu$m), the visibility of the diffraction pattern was relatively high (see Fig. 5.16) and 5 incoherent probes were sufficient to describe the beam properties.

5.7 Large field of view reconstruction of a polysphere sample

Most of the CDI imaging with laboratory sources was focused on high resolution imaging (e.g. Refs. [39, 56, 176–178]), however it was always limited to very small and highly scattering samples. Reaching high resolution and large field of view for

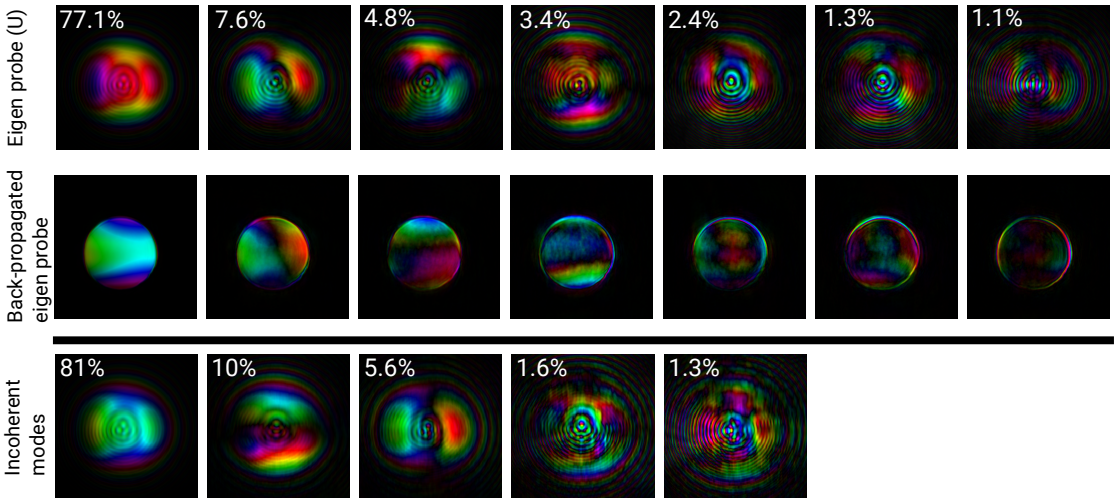

FIGURE 5.18: Reconstructed probe modes for the aperiodic grid sample dataset. The first row shows the OPRP eigenprobes with their relative weight. The second row are the eigenprobes back-propagated 95 μm to the pinhole position. The last row shows the incoherent probes used for the Fourier space correction [16]. Note that only evolution of the first incoherent mode was recovered by the OPRP method.

not fully artificial testing samples is not trivial. Due to long scanning time, beam stability is limited and thermal drifts are not negligible. Also, sample position errors and insufficient flatness of the samples over large field of view are serious limitations.

In order to test the limits of our imaging setup, we have imaged randomly distributed polymethyl methacrylate (PMMA) spheres with 400 nm diameter in average deposited on a 50 nm silicon nitride membrane. The sample was simulating extended sample with feature size similar to biological cells [12]. The total number of collected scanning positions was 976, corresponding to field of view $\approx 105 \times 105\,\mu$m. The illumination probe diameter was 10 μm and the scanning step was 3 μm, resulting in 70% linear overlap. The total exposure time was 0.6 s per a position leading to roughly 11 minutes of exposure time. The total acquisition time was limited by the EUV camera readout to 3 hours. The pixel size was 54 nm and the average deposited dose to the sample was 10^6 Gy. The probe to sample distance was estimated 95.2 μm. The reconstruction was performed with the ePIE based OPRP method with position correction and a weak positivity constraint.

The reconstruction is shown in Fig. 5.19. Due to low penetration of EUV radiation at 30 nm wavelength, the polyspheres create rather a binary sample on contrary

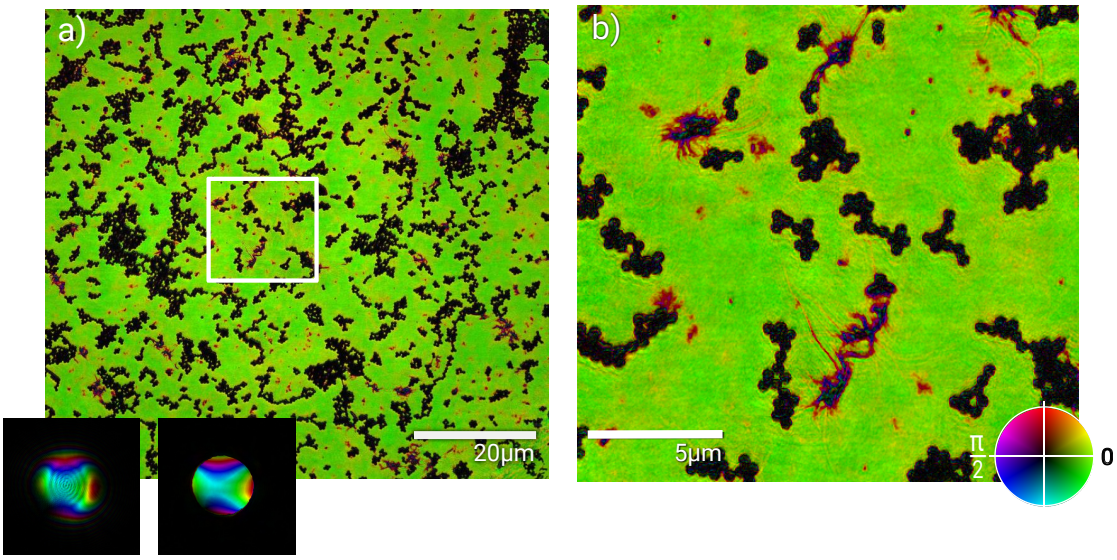

FIGURE 5.19: Ptychographic reconstruction of 400 nm PMMA spheres deposited on a 50 nm thick silicon nitride foil. (a) full reconstruction in the complex colour scale, the reconstructed probe and back-propagated one are in the lower inset up to scale. (b) a magnified region of the reconstruction.

to imaging at 13 nm shown in Chapter 6. However, the edges of the spheres are smoother than in the case of the aperiodic grid sample, so the scattering in the high spatial frequencies was limited. Moreover, the sample contains many regions that are weakly scattering and the diffraction pattern is dominated by the probe scattering. In the magnified region (Fig. 5.19(b)), it can be seen that the sample had some contamination that was partially transparent as well.

The slight "halo effect" that is visible around the PMMA spheres was probably caused by multiple effects:

- Hard edge in the Fourier space - the discrete Fourier transformation is not able to correctly account for features with high signal at the edge of the detector. This may cause artefacts around sharp edges in the reconstruction.
- Object size was close to the depth of focus limit for the aimed resolution. This may result in lensing effect of the polyspheres that cannot be considered as thin objects any more.
- Other approximations used in the ptychography such as the imperfect correction of the coherence properties of the HHG beam.
- Insufficiently sampled scattering from the illumination probe.

One of the suspected options that could have cause the halo effect was the poly-CDI approximation. In order to eliminate potential issues caused by the polyCDI approximation, a pseudo-multicolour ptychographic analysis was performed (see Section 2.6.2). Because the illumination and signal quality was not sufficient to perform full three colour reconstruction using the ptychography multiplexing algorithm (PIM) method [52], we have used the assumption of the sample achromaticity to constrain the object reconstructions. Therefore, only the object reconstruction for the central wavelength that contains more than 70% of power was used and the object reconstructions for the other wavelengths were obtained by FFT based sub-pixel rescaling. The rest of the reconstruction was performed similarly to the PIM algorithm. Reconstructions of the probes for each wavelength (Fig. 5.20) are showing that the spatial intensity distribution at the focus position was slightly different. This would mean that the divergence of each harmonic wavelength differs due to the phase-matching process. Similar behaviour was already reported in Ref. [175].

Another possible explanation is that the multiple probes have similar behaviour as the orthogonal incoherent modes in the mixture state algorithm [16] and the side wavelengths probes pick up also some spatial incoherence properties from the central wavelength. It was already noted that the mixture state algorithm can help with reconstruction of broadband illumination [50]. On the other hand, we have not observed any changes in the reconstruction.

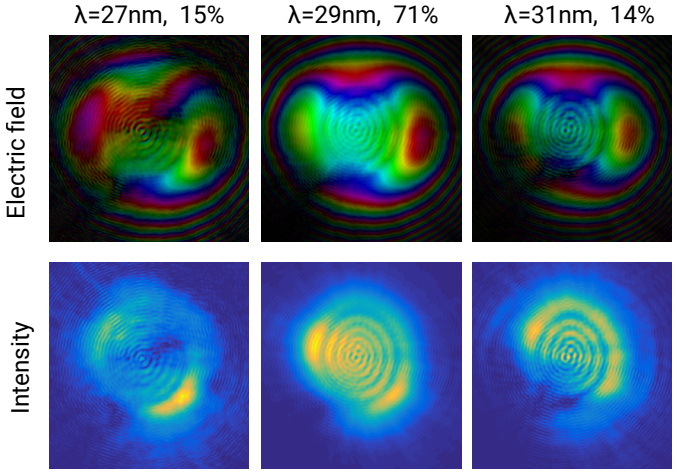

FIGURE 5.20: Reconstruction of multiple colours of the illumination probe with their relative power denoted in title. The intensity distribution for the minor wavelengths is not well centred on the iris.

Another approximation that may not be fully valid, is the maximum probe to sample distance requirement. If the illumination is created by a sharp pinhole, it

contains all spatial frequencies, in contrast to the case when the probe is only projected on the sample. If the probe to sample distance is too large, the high spatial frequencies can "escape" from the reconstructed region of the probe, however they are still collected by the detector. The limiting distance for this approximation was discussed in Sec. 5.5.

There are two main issues related to the pinhole defined beam: firstly, due to the finite pixel size, the oversampling of these high frequencies is below one and without additional knowledge, it is not possible to reconstruct them. The second issue is that in the case of a low contrast sample, the diffraction intensity from the high spatial frequencies of the illumination probe that are not possible to include into the probe, gets stronger than the sample scattering signal and it may cause artefacts. These artefacts are however partly reduced thanks to the over-constraining in ptychography and partly thanks to the use of multiple incoherent probes which pick up some of these unassigned frequencies. However, it may still result in lower reliability of the small features for weakly scattering samples. Comparison between the reconstructed probe intensity in Fourier space and the average diffraction pattern in Fig. 5.21 shows that the highest spatial frequencies of the probe were lost. In order to limit effect of the non-reconstructed probe frequencies, it is possible to use the mixture state algorithm [16], that will at least partly pick up the frequencies into one of the modes (see Fig. 5.21(b)) and thus limit their effect on the final object reconstruction.

Note that this is not an issue for numerical simulations, because the FFT based near-field ASM method propagates the light with the periodic boundary constraint, therefore the high frequencies are not lost, but only wrapped around as it was discussed in Section 2.1.2.

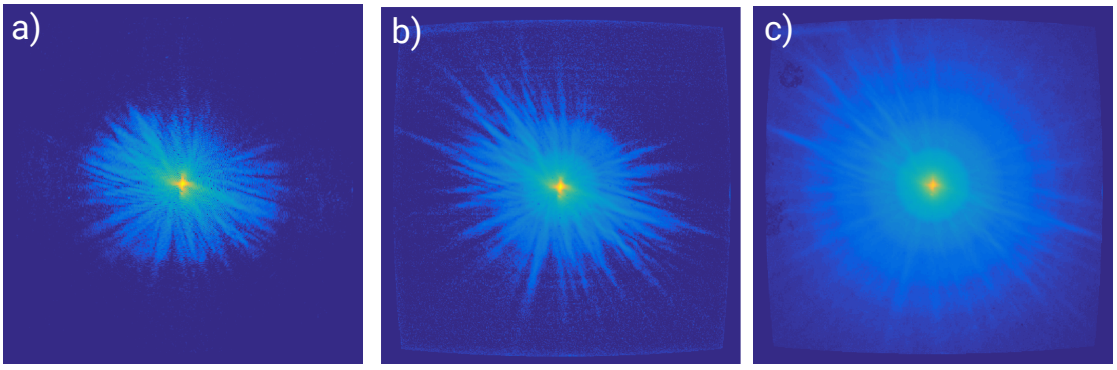

FIGURE 5.21: Comparison of the diffraction pattern from the first reconstructed probe mode (a) with incoherent sum of all probe modes (b) and the average diffraction pattern (c) in logarithmic colour scale.

5.8 Ptychographic imaging of Hippocampal neuron samples

The two main advantages of the EUV imaging at 30 nm is high resolution and good natural contrast for thin biological samples. The main drawback is that the attenuation length in polycarbonates is 50 nm and therefore the maximum sample thickness should be below 100 nm. It also means that there is only limited overlap with X-ray CDI microscopy because of its low sensitivity and high penetration.

One of the promising samples, where the high resolution and sensitivity is advantageous, is label-free imaging of neuronal cell structure. EUV ptychography can be used to perform morphological analysis of fine neuronal structures, such as synapses, during degenerative conditions, eg. Alzheimer's disease. It can help to understand the disease progression. EUV ptychography could help to study growing and evolution of the neurites and investigate how their structure depends on their environment. Since it is a label free method, it can be also used to check the reliability of label based fluorescence microscopy.

Our neuron samples were prepared by Jo Bailey from Southampton university. The neurons were isolated from E17 mouse embryos and they were cultured for 7-14 days on a 50 nm thick silicon nitride foil coated by a thin layer of poly-D-lysine. After that, the neurons were fixed in methanol. There was not any staining of the sample. There were also attempts to grow the neurons on a 30 nm thick support because transmission of the 50 nm silicon nitride is 16%, while 30 nm foil has 35% transmission, however due to different electro-chemical properties of the thin foils, it was not successful. More details about the sample preparation can be found in Ref. [179].

The reconstructed variable illumination is shown in Fig. 5.23. The illumination was very stable in this scan because it can be practically fully described by the first 4 modes, while the last 2 modes were used to correct for subpixel shifts.

High resolution reconstruction of 14 days in vitro neurons sample are shown in Fig. 5.24. The pixel size was 42 nm, however the real resolution was around 50 nm in highly transmissive regions and between 80 nm to $\approx$1 μm in the weakly transmissive regions. The scan contained 205 diffraction patterns with 15 s exposure

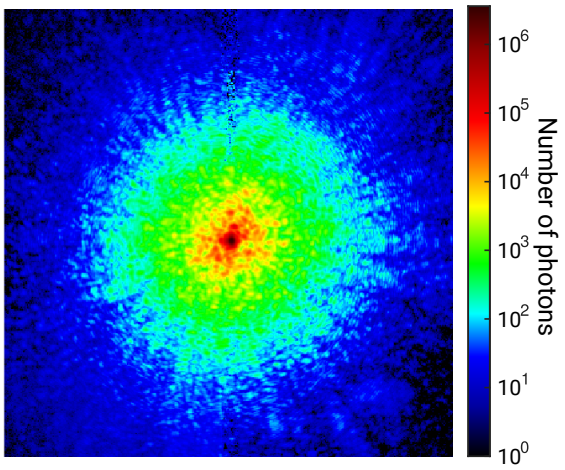

FIGURE 5.22: An example of a collected diffraction pattern from the neuron sample.

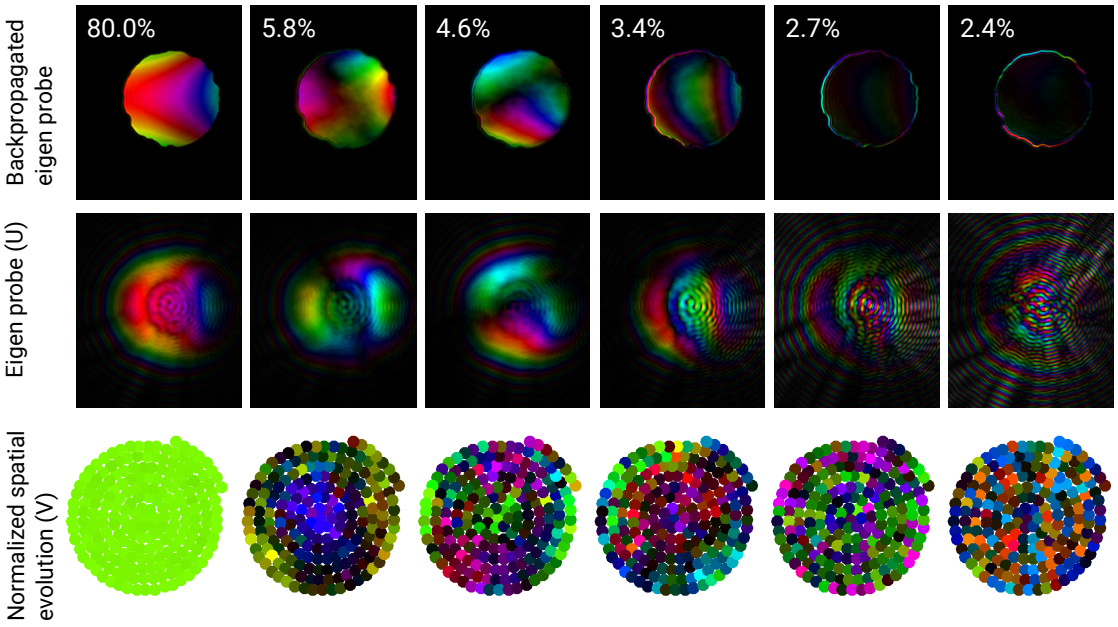

FIGURE 5.23: An example of reconstruction of the eigen probes and their complex time evolution for the neuron sample. Modes containing the beam movements (1-4) are clearly separated from the modes containing subpixel shift (5-6) in the pinhole plane.

time. The magnified region in Fig. 5.24(b) shows high contrast and reconstruction quality that would be very difficult to be achievable by any other microscopy techniques.

The complex reconstruction of the neurons can be used for some basic analysis. Knowledge of amplitude and phase shift can be used for estimation of the material constants δ, β and local sample thickness. The object complex transmission can

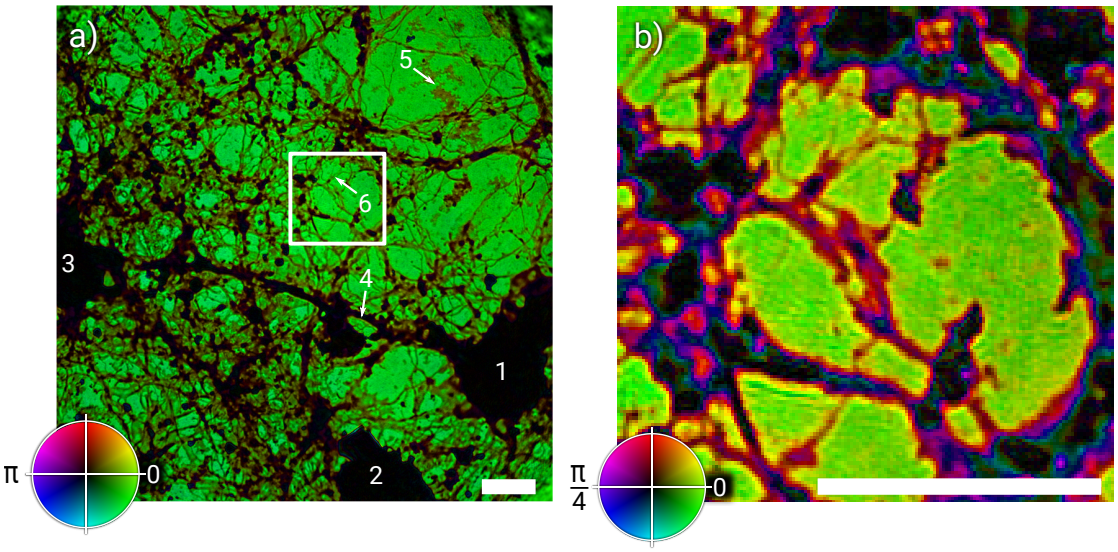

FIGURE 5.24: A reconstruction of 14 days in vitro neurons and their connections. 1-3 are neuron cell bodies (cell soma), 4 is axon connection, 5 are probably cell residuum after fixation and 6 is an example of a growth cone and filopodia. (b) contains an enhanced contrast subregion, where the dendrites on the poly-D-lysine layer are clearly visible. The illumination probes are shown in Fig. 5.23. Scale bar is 5 μm, pixel size is 42 nm.

be expressed as

$$O(\mathbf{x},\lambda) = \exp\left[\frac{2\mathrm{i}\pi}{\lambda}(1-\delta(\mathbf{x},\lambda)-\mathrm{i}\beta(\mathbf{x},\lambda))T(\mathbf{x})\right] \quad , \tag{5.10}$$

where complex refractive index is $n = 1-\delta-\mathrm{i}\beta$ and T is average thickness over the pixel area. If we assume that the δ,β depends on the wavelength only weakly and that the harmonic wavelengths are close to each other and the central wavelength is significantly stronger than the other harmonics, then $\delta(\lambda),\beta(\lambda)$ can be approximated by the value for our central wavelength 29 nm.

The simplest analysis is estimation of thickness from the Beer-Lambert law using the assumption that the neuron structure is composed mostly from proteins (Fig. 5.25(b,e)). The maximum thickness that can be reliably estimated is limited by attenuation to roughly 200 nm, therefore all higher values were cropped. The precision of the thickness measurements depends on validity of the assumption of the sample composition and also on the reconstruction quality. Our estimation is that the precision is better than 10 nm. The thickness of the neurites varies between 150 nm and 20 nm. It is also possible to use the phase for thickness approximation, however it is less reliable due to requirement of 2D phase unwrapping.

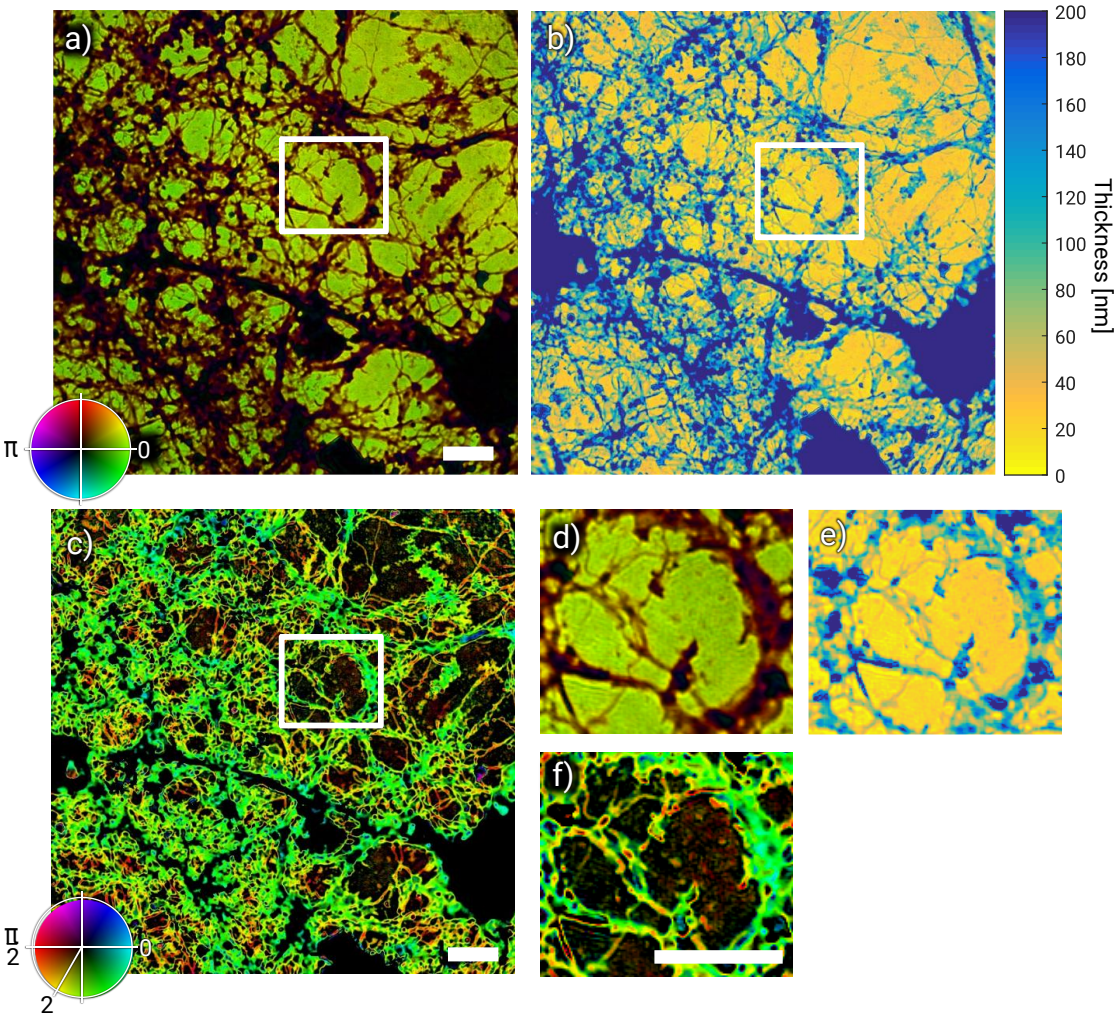

FIGURE 5.25: Basic analysis of the neurons reconstruction. (a) is the original reconstruction in the standard complex colour scale, (b) is thickness estimation from transmission measurements. Thickness above 200 nm cannot be measured due to low transmission. (c) shows δ/β material analysis plotted in complex colour scale, where colour corresponds to δ/β and brightness to estimation of reliably of the pixel. Images (d-f) contain magnified region of images (a-c). Scale bar is 5 μm.

Phase unwrapping cannot be done reliably for this sample due to large number of phase residua in the low transmission regions.

Another information that can be extracted is the material constant given by the δ/β ratio. In order to avoid phase unwrapping, we have removed low transmission regions from our measuements. Therefore, only the less absorbing regions with phase change sufficiency lower than 2π can be seen in Fig. 5.25. Moreover, reliability of the δ/β ratio is changing significantly in dependence on the local transmission. If the sample is thin, then both the phase shift and attenuation are low and their ratio has low reliability. If the attenuation is high, then the signal

through given pixel is low and the reliability is low as well. Therefore, the δ/β ratio in the Fig. 5.25(c,f) is shown in a complex colour scale, where brightness expresses the reliability (i.e. dark regions are unreliable) and the colour corresponds to the δ/β ratio.

Note that compared to the water-window SXR microscopy, the δ/β ratio is changing only weakly for different biological elements at 30 nm wavelength (see Tab. 5.2), therefore the material sensitivity is limited. Additional element sensitivity could be theoretically reached if the refractive index variation for multiple wavelengths is reconstructed as it was attempted in Ref. [151]. However, current quality of the HHG datasets is not sufficient for reliable multicolour imaging.

	δ	β	δ/β
C	0.172	0.0706	2.43
O	0.100	0.133	0.75
H	0.488	0.0696	7
$C_{16}H_{14}O_3$	0.0986	0.0464	2.12

TABLE 5.2: Estimations of the complex refractive index $n = 1 - \delta - i\beta$ for basic biological elements and an example of polycarbonate from the CXRO database [9]. The density was assumed 1.1 g/mm^3.

5.8.1 Reconstructions of young neurons

One of the primary goals of the neurons sample investigation is to understand more the development of young neurons from individual cells into complex neural networks. In Fig. 5.26, we show some preliminary results of neurons that were cultured for 7 days and then fixed in methanol. Despite the fact that EUV microscopy cannot provide in-vivo images due to low penetration in water, still it should be possible to gain some information from multiple reconstructions of different samples cultured for different periods. The reconstructions of the young neurons are clearly showing far less connections and much sparser features, which makes these samples even more weakly scattering than the samples presented in previous chapter (Fig. 5.24). On the other hand, the dendrites are thinner and therefore more transparent, which makes these samples more suitable for EUV CDI microscopy.

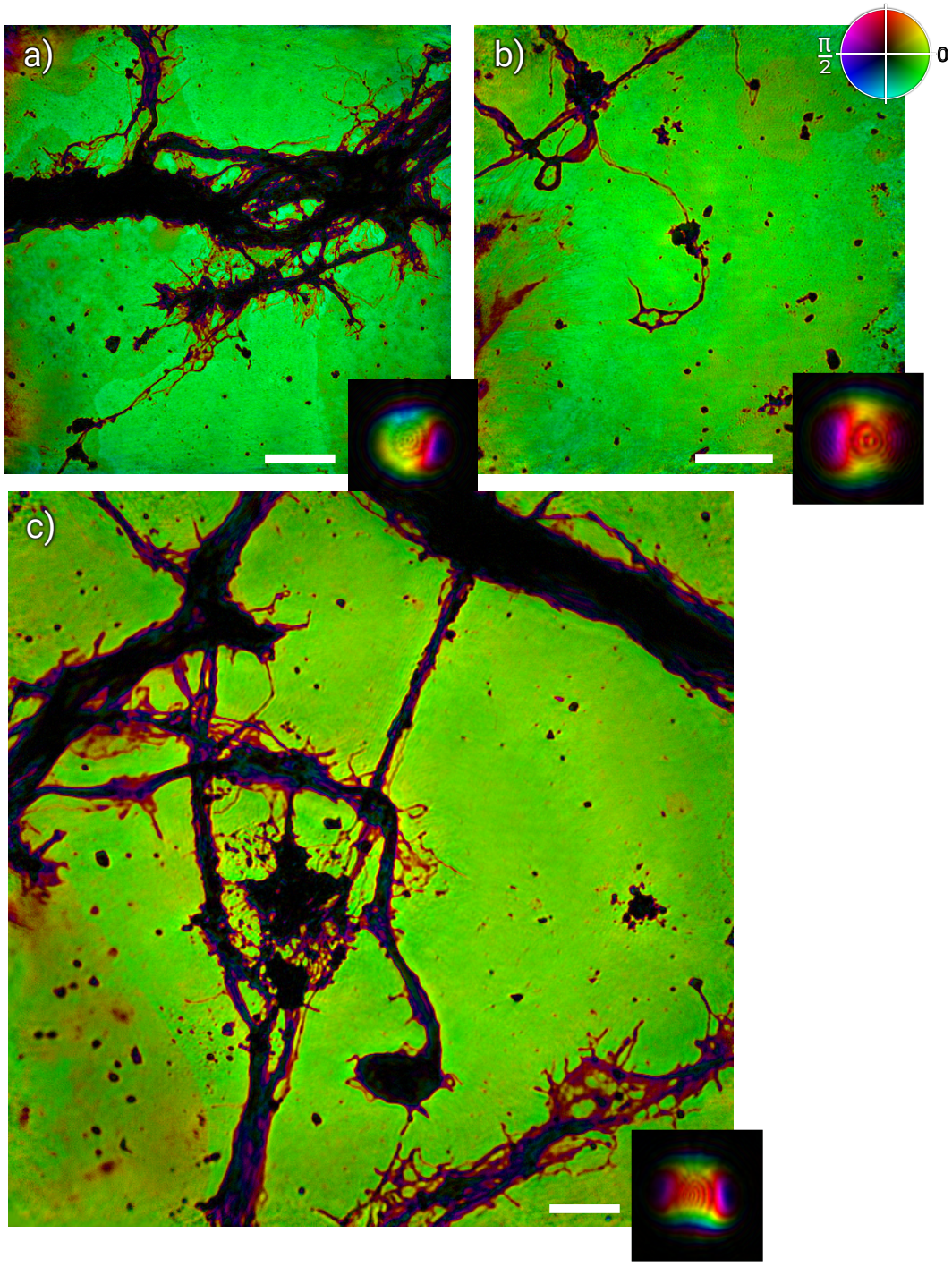

FIGURE 5.26: Examples of reconstructions of 7 days old neuron cell cultures. (a,b) are samples measured in Southampton with 2 s exposure and 90 nm pixel size, while reconstruction (c) was acquired at the ARTEMIS laser in the Central Laser Facility, Harwell Oxford. Due to less optimized HHG source, 5 s exposure per a position was needed to obtain 140 nm pixel size. Scale bars are 10 μm.

5.8.2 Comparison of EUV ptychography and HXR ptychography reconstructions

The 14-days old sample (Fig. 5.24) was also imaged by a visible light microscope (Fig. 5.27(b)) and by ptychography using synchrotron radiation at 7.3 keV at Diamond Light Source synchrotron at the I13 beamline (Fig. 5.27(c)). The exposure time was 1 s and the probe diameter was roughly 10 μm. Both these parameters are relatively close to our HHG experiment. The image was reconstructed by our difference maps code and refined using the ePIE code. More details about this experiment can be found in Ref. [179]. The hard X-ray ptychography reconstruc-

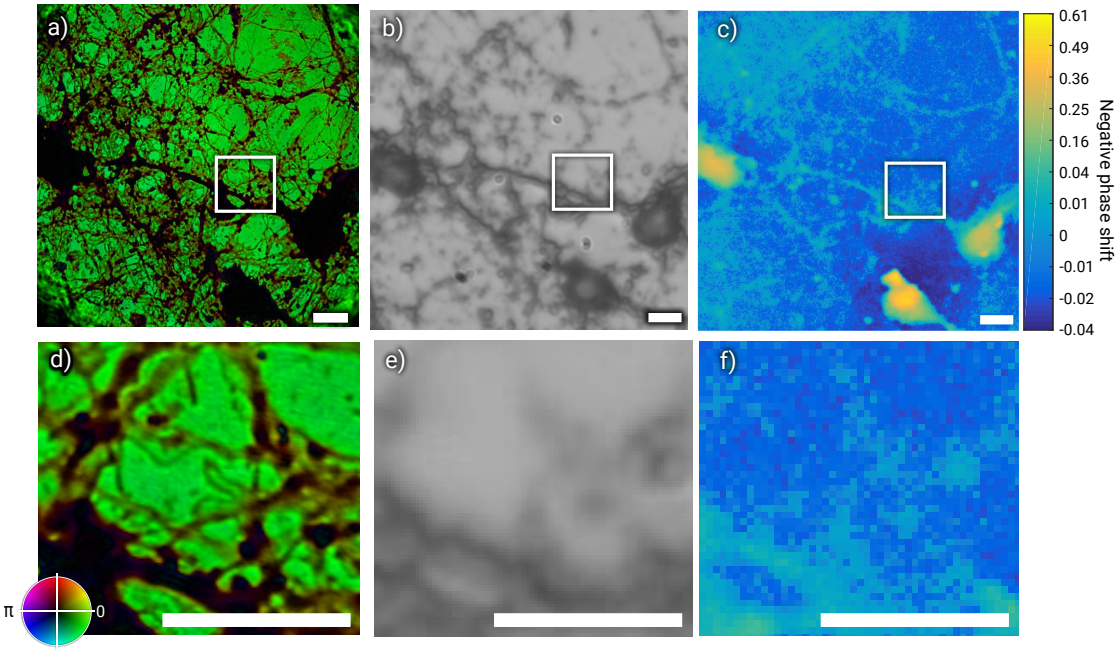

FIGURE 5.27: Comparison of the EUV ptychography reconstruction (a) with visible light microscope (NA=0.6, mag. 40×), and ptychography reconstruction at 7.3 keV at Diamond I13 beamline (c). Note that the phase reconstruction is shown in nonlinear colour scale in order to visualize the weakly phase shifting features. Images (d-f) show a magnified region of the sample. Scale bars are 5 μm.

tion has the advantages that it can easily see through all the thick cell bodies. The HXR phase is shown in nonlinear colour scale, in order to show the thick cell bodies and also weakly interacting neurites. Phase sensitivity in our HXR experiment was roughly 0.01 rad, which is comparable or better than other published reconstructions of weakly scattering specimens [58, 180, 181]. Despite this, neither the resolution nor the quality can be compared with the sharp and detailed EUV reconstruction. Furthermore, the X-ray reconstruction contains large phase

artefacts around high contrast objects such as the cell bodies. It can be seen in the Fig. 5.27(c) as dark blue regions with the phase shift value below zero.

5.8.3 Radiation damage

One of the common issues of the short wavelength microscopy is radiation damage [24, 182]. Most of the publications focused on radiation damage investigated wavelength region around the water-window region or even for higher photon energies. Because EUV range is not commonly used for microscopy of biological samples, there are no studies about radiation damage in this region. The specific property of the EUV is very strong interaction with any material, therefore the deposited dose is relatively high because most of the photons are lost in the first 50 nm of the specimen.

On the other hand, in the hard X-ray regions most of the photon sample interaction is done by elastic scattering that causes only a phase shift. Whereas for our EUV imaging, the δ/β ratio is 2.1 for proteins, the value for 7.1 keV is $\delta/\beta = 400$. Therefore the initial idea of our HXR experiment was that the resolution criterion given in Ref. [24] for bright field SXR microscopy can be avoided because ptychography can get more than 400× more signal from the phase than attenuation based microscopy methods. In order to test this idea, the neuron sample was scanned with a 7.1 keV beam using 10 s exposure time per a position and 2 μm probe diameter. The photon flux on sample was $1.4{\cdot}10^8$ photons/s and the dose deposited into the sample was $3.9{\cdot}10^5$ Gy. If we compare this dose with the dose deposited into the sample during the ptychography scan presented in Sec. 5.8, where the photon flux at 42 eV at the sample was $2.1{\cdot}10^8$ photons/s with 15 s exposure, however due to much stronger absorption the total dose during the ptychography was $4.4{\cdot}10^7$ Gy. Note that the dose was unnecessarily increased by factor of 6 because the neurons were grown on the side of the silicon nitride that was facing the pinhole instead of facing the camera direction. The additional attenuation of 84% by the silicon nitride foil caused this increase. After this correction, the deposited EUV dose should correspond to 40-50 nm resolution according to Ref. [24], which agrees well with our results.

Figure 5.28 shows EUV and HXR reconstruction of the same neuron sample. The EUV ptychographic scan was taken after the HXR scan. Figure 5.28(a) shows

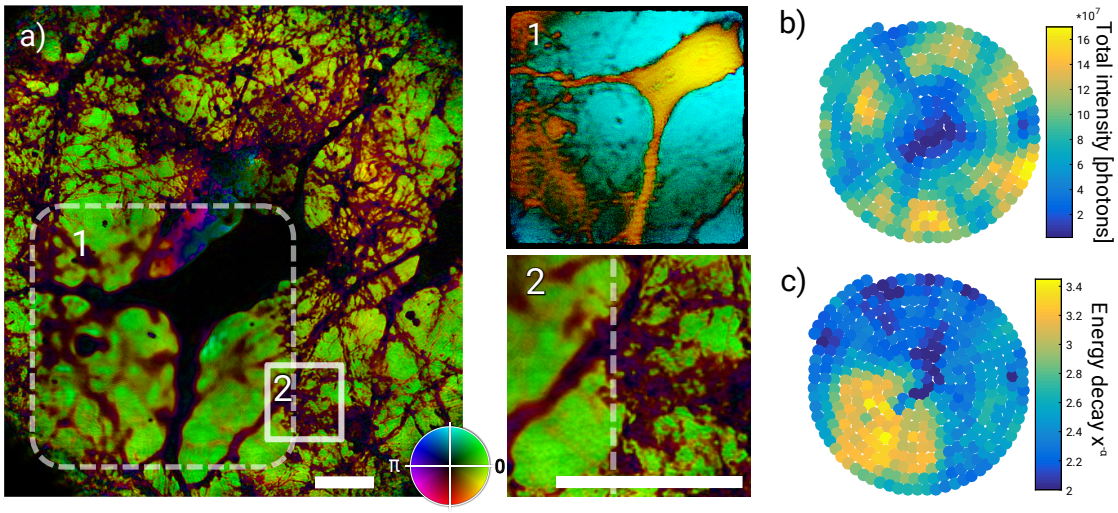

FIGURE 5.28: An example of radiation damage caused by X-ray ptychography at 7.1 keV. (a) EUV ptychography reconstruction after the X-ray ptychography scan. Subregion (1) was scanned by a X-ray beam with 10 s exposure. The X-ray ptychography reconstruction in a contrast enhanced colour scale is shown in (1). The transition region around the damaged region is shown in (2). The image (b) shows total number of photons in each measured diffraction pattern, while (c) shows the decay rate of the scattering energy in dependence on the spatial frequency for each measured diffraction pattern. Scale bar denotes 10 μm.

in the bottom left corner (area 1) the X-ray scan region. The X-ray ptychography reconstruction is shown in Fig. 5.28(1). The detail of the transition area in Fig. 5.28(2) is clearly showing that all fine structures of our sample were lost. In order to get more quantitative estimation of the radiation damage, the radially averaged power spectrum was fitted by $x^{-\alpha}$. Figure 5.28(b) total number of photons collected in each measured diffraction pattern. Figure 5.28(c) shows the decay constant α for each measured position. The energy decay was increased at least twice in all transparent regions (i.e. except dark blue regions in Fig. 5.28(b)). The real decay could have been even higher, because the diffraction pattern starts to be dominated by the probe diffraction in the weakly scattering (damaged) regions.

EUV is ionizing radiation, therefore it is surprising that despite two orders of magnitude higher dose it did not cause any obvious damage, while HXR radiation completely wiped out all finer structures. Our results clearly show that it is not possible to image biological sample with HXR radiation without cryo cooling. Despite the very high δ/β ratio, the radiation damage mechanisms are clearly different to the EUV photons. On the other hand, EUV radiation at 30 nm causes surprisingly low radiation damage. The data presented in Fig. 5.28 were taken

with 2 s exposure, however data presented in Fig. 5.24 needed 15 s exposure per a position and still no obvious damage was observed.

There are no studies about radiation damage in EUV below 123 eV wavelengths, however in the "water-window" soft X-ray region, the photon to sample interaction is done mainly by photoelectric effect. This provides a photoelectron, creates secondary electrons by collisional ionization and leaves a hole in shell of the atom that can be filled by another electron from the outer shells resulting in a ion and another low energy electron due to the Auger effect [41]. Radiation at 42 eV that we use has sufficient energy to remove electrons only from the two most outer shells (11.3 eV, 24.4 eV) and therefore the Auger effect cannot happen. The produced photoelectrons have too low energy to create significant damage as well. On the other hand, the HXR photons interact with the light element more through the Compton scattering and thus they can produce a large number of highly energetic secondary electrons.

In future, a more thorough investigation of the EUV radiation damage is needed, however for our fixed neurons samples it seems that it is not a serious limitation and the dose can be still reduced by 84% if the sample is prepared optimally.

5.9 Effective near-field EUV ptychography

The idea of near-field ptychography is to collect the diffraction pattern in the near-field regime i.e. sufficiently close to the sample so that the Fresnel number satisfies $F \geq 1$. The second option is to use the beam curvature to get into effective near-field regime, where the pixels in the sample plane have different size than in the camera plane (see Section 2.1.3). The optimal propagator on the detector plane is the angular spectrum method (ASM) [36]. The setup (Fig. 5.29) was almost identical to our far-field setup. The main difference was that the pinhole or structured illumination was projected onto the sample using the focusing mirror. The sample was placed 1.5 cm down-stream the focus and the camera was shifted 6.1 cm behind the sample. However, the transformation from far-field to near-field setup could be done simply under vacuum with a sufficiently long translation stage for the camera.

The main gain of the near-field setup is possibility to image a very large field of view without the limitations of the coherence requirement and oversampling of

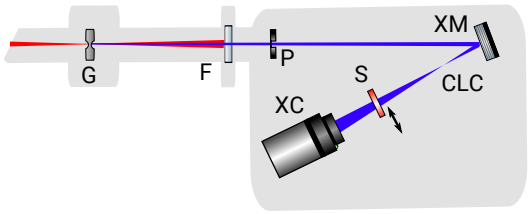

FIGURE 5.29: A schematic of the near-field ptychography setup. Sample S was placed 5-15 mm downstream the focus (CLC) in order to get curved illumination and a large illumination spot on the sample. The pinhole P or other optics creating the structured illumination was placed upstream of the focusing mirror XM and projected onto the sample.

the signal on the detector. As we will show, the coherence is still crucial for high resolution imaging, however it is only limiting the achievable resolution and not the field of view.

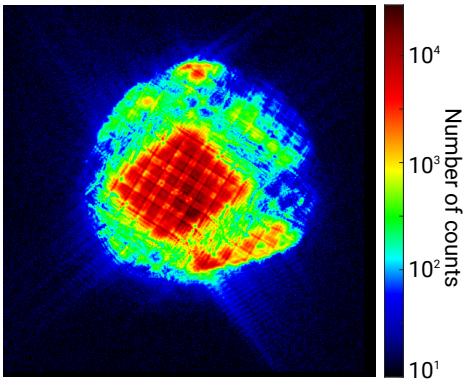

FIGURE 5.30: An example of the measured "near-field" diffraction pattern used for reconstruction in Fig. 5.31(a). The fine grid-like structure was caused by the illumination probe (Fig. 5.31(c)), while the large square in centre is produced by the sample grid. The dynamic range is significantly reduced compared to the far-field pattern.

We have performed several experiments in order to test this imaging method. Firstly, we have attempted to measure the achievable resolution in our near-field imaging setup. Theoretically, the diffraction limited resolution would be 385 nm, however it is not achievable because the flat profile of the collected near-field pattern does not allow to use the full potential of the high dynamic range imaging (see Fig. 5.30). Therefore, only weak signal from the higher spatial frequencies was collected. Another factor limiting the resolution is given by demagnification of the camera pixels on the sample plane (Fig. 2.10(a)). In contrast to the far-field imaging, the near-field approximation results in the same pixel size in the real (sample) plane and Fourier plane in the case of flat illumination wavefront.

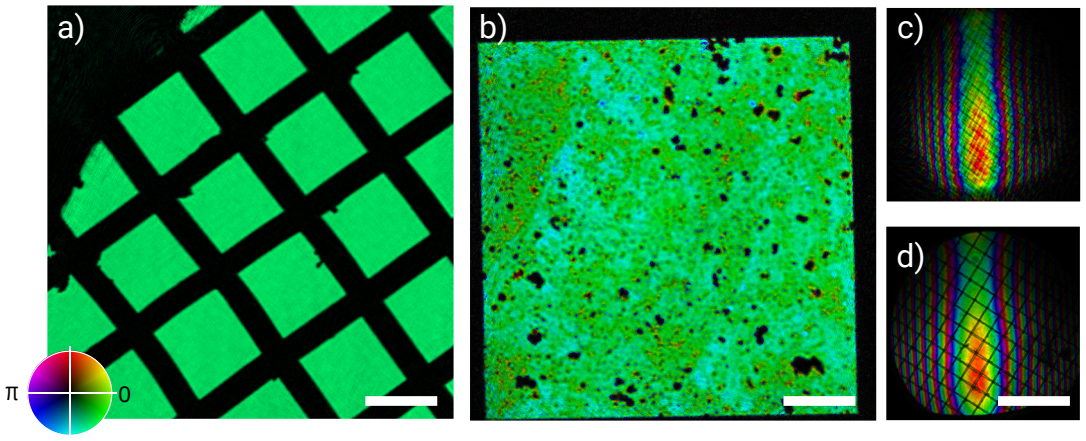

FIGURE 5.31: Reconstruction of a calibration grid (a) and our DNA sample (b). Figure (c) shows an example of the reconstructed probe and (d) is back-propagated probe to image of the metal mesh placed at the filter position P (see Fig. 5.29). Scale bars are 100 μm.

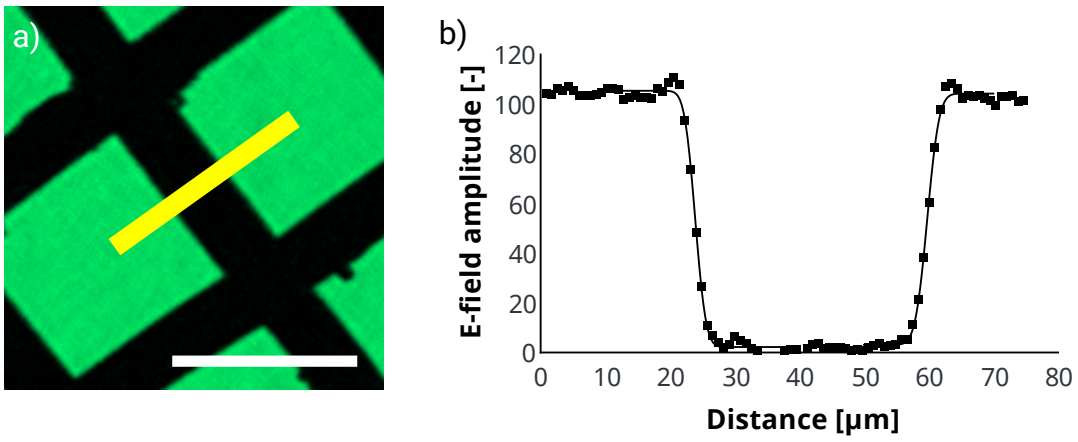

FIGURE 5.32: Results of the knife-edge analysis (b) applied on a sharp edge of the imaged grid (a) for estimation of resolution. Scale bar is 100 μm. The resolution was estimated to 4.1 μm.

In the case of curved wavefront, the pixels are rescaled according to the Fresnel scaling theorem as discussed in Section 2.1.3. The pixel size in the real-plane was estimated to 0.78 μm, therefore the magnification was 17×. This sampling limitation can be theoretically avoided [183] in the case of high coherence and signal. The second limitation is the virtual source size that is closely related to the coherence of the illumination (Fig. 2.10(b)). The virtual source size gives the smallest pixel size on the sample and for very high Fresnel numbers, it directly gives the maximum achievable resolution. However, in our case the Fresnel number was 3 and the coherence is the most limiting factor because the coherence length of our HHG beam is roughly equal to half of the beam diameter.

Two examples of reconstructions are shown in Fig. 5.31. Fig. 5.31(a) shows a

reconstruction of a TEM mesh with 100 μm windows and Fig. 5.31(b) shows reconstruction of a 50 nm thick silicon nitride membrane with 500 μm square window containing a thin poly-D-lysine layer ($\approx$5 nm) and DNA particles (dark spots). The resolution was measured using the knife-edge method using ERF fit of the reconstructed amplitude of the electric field (see Fig. 5.32). The full-period resolution was estimated to 4.1 μm. Note that if the resolution would be measured using the intensity values (square of the modulus), the resolution would be $\approx 3\,\mu$m.

The causes of limited resolution with respect to the diffraction limited pixel size were discussed in Section 2.2.1.

There are several ways to improve the achievable resolution: increase coherence, use shorter wavelength, use shorter propagation distance or use more curved illumination, which has a similar effect as a shorter propagation distance. The illumination NA in our setup was only 0.01, which is not sufficient for any high resolution imaging. Shorter propagation distance would not help due to large pixels of our camera (Andor DX434, 13 μm). On the other hand, high signal intensity that is limited mainly only by the Poisson noise allows the use of much faster readout rate of the EUV sensitive camera and thus much faster imaging. The readout time is usually around 50-95% of the total data acquisition time for our far-field EUV ptychography.

In the second experiment, we have tested the effect of the illumination structure on the reconstruction quality. It was already shown that illumination with a more complex structure helps to achieve better reconstruction quality [38, 45, 124, 184, 185]. Because there are no random phase diffusers available in the EUV range, we have used a nickel mesh placed upstream of the focusing mirror to create a complex pattern on the sample position and compared it with the flat illumination profile. The flat illumination still contains some low contrast features caused by dust on the multilayer mirror and imperfections of the aluminium filter. The reconstruction of the TEM mesh is shown in Fig. 5.33, where Fig. 5.33(a) is a reconstruction with flat illumination (Fig. 5.33(d)) and Fig. 5.33(b) shows reconstruction with structured illumination (Fig. 5.33(f)). Details of the figures 5.33(c,e) clearly show improved quality when the structured illumination was used.

The optimal frequency content of the illumination would need further investigation, however our results indicate that the illumination should contain enough of

the sharp features but without too fine structures that would result in lower visibility of the measured diffraction patterns. Similar conlusions were done in Ref. [184].

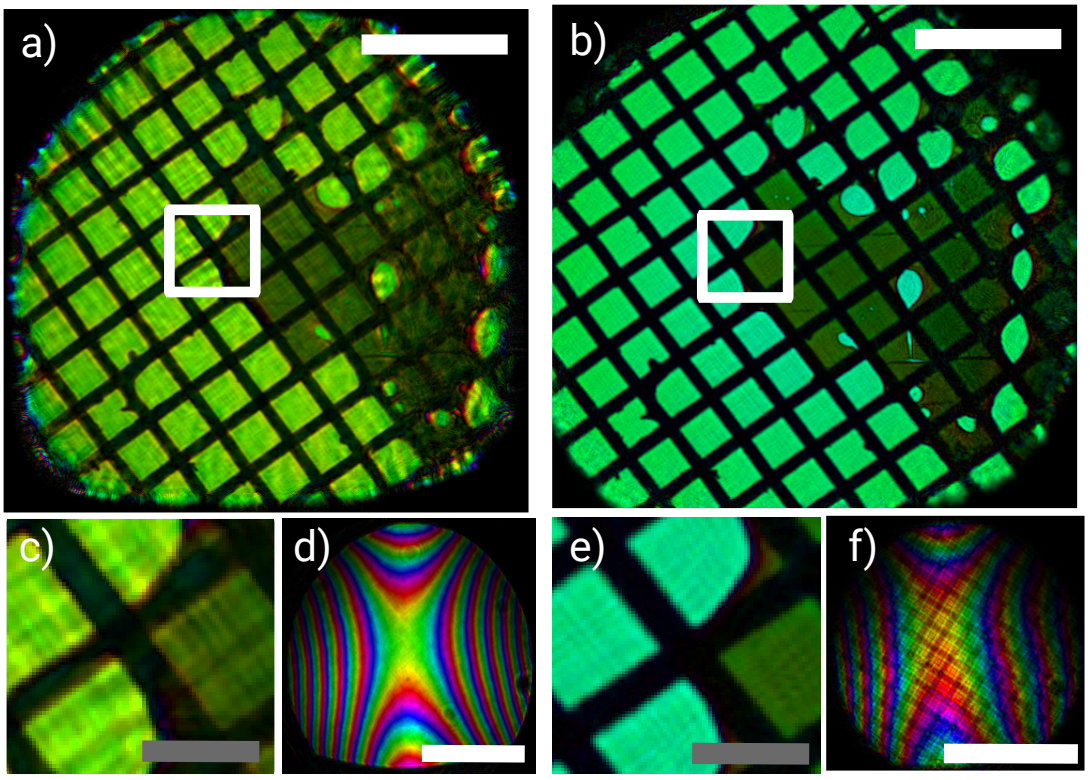

FIGURE 5.33: Near-field ptychography reconstruction of a TEM copper grid with a damaged 80 nm thick formvar foil. Fig. (a) shows reconstruction with mostly flat illumination function (d) while (b) shows reconstruction with more complex illumination probe (f). Comparison of the zoomed regions is shown in (c,e). White scale bars denote 300 μm and the grey ones denote 100 μm.

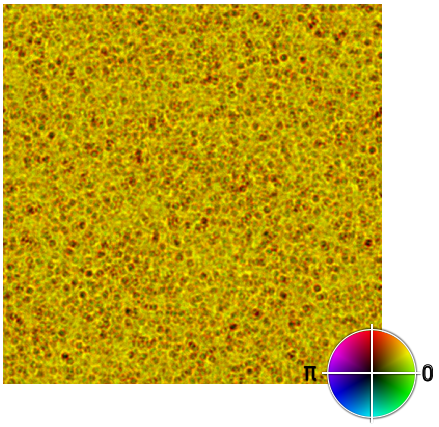

FIGURE 5.34: A simulation of near-field propagation from an array of polyspheres. The propagation distance was 200 μm, wavelength 13 nm and the PMMA spheres diameter was 400 nm. The transmission would be roughly 70%, dependent on the density of the polyspheres.

5.10 Conclusion

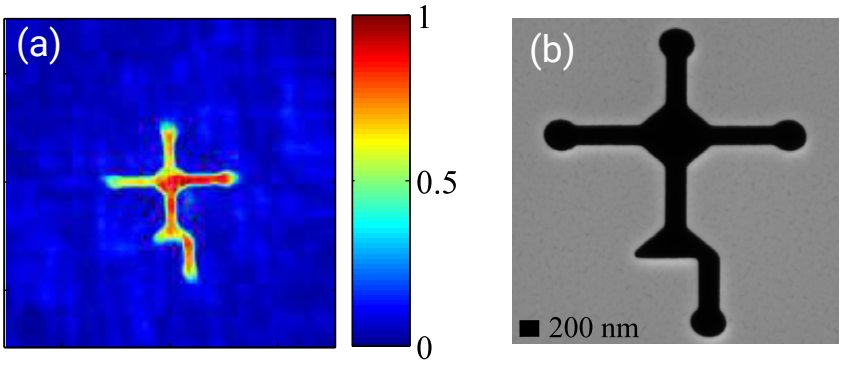

FIGURE 5.35: Reconstruction of a simple binary highly scattering sample (a) published by our group in Ref. [17] and SEM image of the sample (b). Accumulated exposure time was 220 s and the half-period resolution was estimated roughly 100 nm.

Basic properties and physics behind the high harmonic generation were outlined in this chapter. The HHG process can provide highly coherent short wavelength illumination, however the achievable flux, wavelength and level of coherence significantly depend on the experimental parameters. In particular, the optimal phase matching can significantly improve the achievable flux, however the maximum flux often does not lead to optimal parameters for CDI imaging and a trade-off between flux, coherence and stability needs to be found. HHG photons can be generated down to the "water-window" range (2.34-4.4 nm) (see e.g. [161, 162, 186]), however sufficient flux for reasonable CDI microscopy was till now successfully produced only around 30 nm wavelength (42 eV energy).

Based on the HHG EUV source, we have presented a ptychography microscope built in the Southampton university. We have demonstrated that using coherent diffractive imaging at 30 nm wavelength, it is possible to get high quality reconstructions even for relatively weakly scattering real biological samples with resolution down to 50 nm (Sec.5.8).

This is a significant progress since the last best results published by our group shown in Fig. 5.35. The resolution was improved from $\approx$ 100 nm down to 60 nm and the exposure time was reduced from 220 s to 0.2 s per a position. It was done partly by better tuning of the phase-matching process of our HHG source [179], which allowed us to reach higher coherence, flux and focusability. Furthermore, the possibility to accept lower temporal coherence helped to avoid additional losses due to spectral filtering that is inefficient in the EUV range. Moreover, longer

acquisition time was allowed thanks to improved stability by our custom-made laser pointing stabilization system (Appendix A) and better passive stability thanks to other improvements in our lab such as removing sources of vibrations and better sealing of the laser path to avoid air fluctuations. And finally, the major improvements were done in the reconstructed code. Moreover, our OPRP method helped to reduce artefacts in the reconstructions for datasets with drifting beam and significantly changing illumination intensities (see e.g. Sec. 3.2). Our new position correction methods presented in Sec. 2.6.2 reduced the artefacts caused by low precision of the used piezoelectric actuators (see Appendix B).

All these improvements helped to move laboratory EUV microscopy towards a usable tabletop EUV microscope. Despite this, there are still many challenges that need to be tackled before HHG based EUV CDI will become a more standard tool for biology and material science.

In future, it would be possible to increase the generated coherent flux by using a higher power driving laser or a hollow capillary [163]. Another option is to use a thinner silicon nitride support for the biological samples, using a 30 nm foil would increase the transmission roughly twice. Finally, modern EUV sensitive cameras provide 2-3× higher quantum efficiency at the 42 eV energy. All these improvements reduce the exposure time by more than one order of magnitude without need of significant changes in the experiment.

Recently, first examples of ptychography with HHG sources around 90 eV (13 nm) were already demonstrated by Gardner et al. [187] for highly scattering artificial samples. Despite of the fact that the shorter wavelength allows better penetration through sample, significant improvements of sources will be needed to move from simple, practically binary samples towards the real biological imaging. Dependence of the refractive index on the wavelength for polycarbonate (Fig. 5.36) shows that the absorption contrast for an identical biological sample drops 25 times and the phase contrast drops 10 times if the sample would be imaged at 10 nm wavelength compared to our 30 nm wavelength.

The required source radiance for imaging at given resolution and exposure time is proportional to [24]

$$L_{\text{source}} \sim \frac{1}{C^2 \lambda} \quad , \tag{5.11}$$

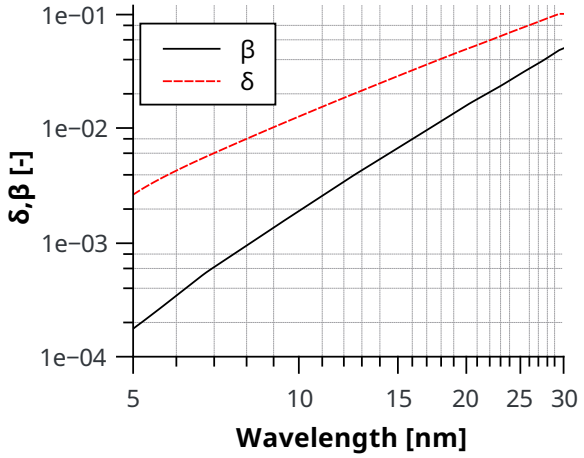

FIGURE 5.36: Dependence of complex refractive index $n = 1 - \delta - \mathrm{i}\beta$ of polycarbonate ($C_{16}H_{14}O_3$) on illumination wavelength [9].

where C denotes contrast (Eq. 2.44) and λ is illumination wavelength. Therefore, imaging of weakly scattering samples using 10 nm wavelength at the same resolution as presented in this chapter would require >300× higher source radiance if the phase contrast would be used. If a similar far-field geometry of the experiment would be used, then the required dynamic range of the collected diffraction patterns would be increased by two orders of magnitude as well.

Chapter 6

Ptychography with plasma based EUV source

In the previous chapter, the high harmonic generation based CDI microscope was presented. As it was mentioned, the advantage of HHG is that the generated light is already highly spatially coherent and the temporal coherence can be easily improved by a reflection from multilayer mirrors. However, HHG is not the only way to generate coherent EUV light in laboratory scale experiments. Another option is to use plasma based incoherent sources, where the plasma is sufficiently heated and compressed in order to generate EUV from a relatively small source. In contrast to HHG, these sources are practically fully incoherent and therefore it is necessary to use large amount of filtering in order to gain coherent light. On the other hand, the advantage is that these sources can generate more than $20\,\mathrm{W}/(2\pi\mathrm{sr})$ [19, 188] of EUV power at 13.5 nm (up to 200 W in the case of high volume manufacturing lithography sources) in contrast to much less than $1\,\mu$W generated by the HHG source. Furthermore, maintenance is much easier compared to the ultrafast lasers which makes these sources more suitable for industrial applications.

EUV plasma sources are powerful and versatile sources commonly used for microscopy in EUV and SXR range [33, 34, 57, 150, 189–192]. We have used two different sources for our ptychography experiments. The first source was a gas discharge plasma (GDP) EUV source from Philips built in cooperation with Fraunhofer Institute for Laser Technology (ILT) in Aachen. GDP sources generate electric field between a hollow cathode and an anode. The subsequent discharge current creates magnetic field that compresses the plasma. This effect is called

pinching. The dense and hot plasma pinch is very unstable with life-time of tens nanoseconds, however with the optimal setting of density and temperature, it can very efficiently generate EUV radiation from a source size ≈200-1000 μm. The discharges can be repeated with up to 3 kHz frequency, however due to limited stability, the frequency was limited to 1 kHz. A photo of the GDP source and schema of the hollow cathode are shown in Fig. 6.1. The plasma source was designed to use xenon gas as the "fuel", however oxygen and nitrogen fuels were successfully tested as well. The xenon fuel provides best output power around 10.9 nm and 13.5 nm wavelengths, while the oxygen plasma results in a strong peak at 17.3 nm wavelength. This kind of discharge source is scalable down to the water-window range (2.3-4.4 nm) [33, 193, 194].

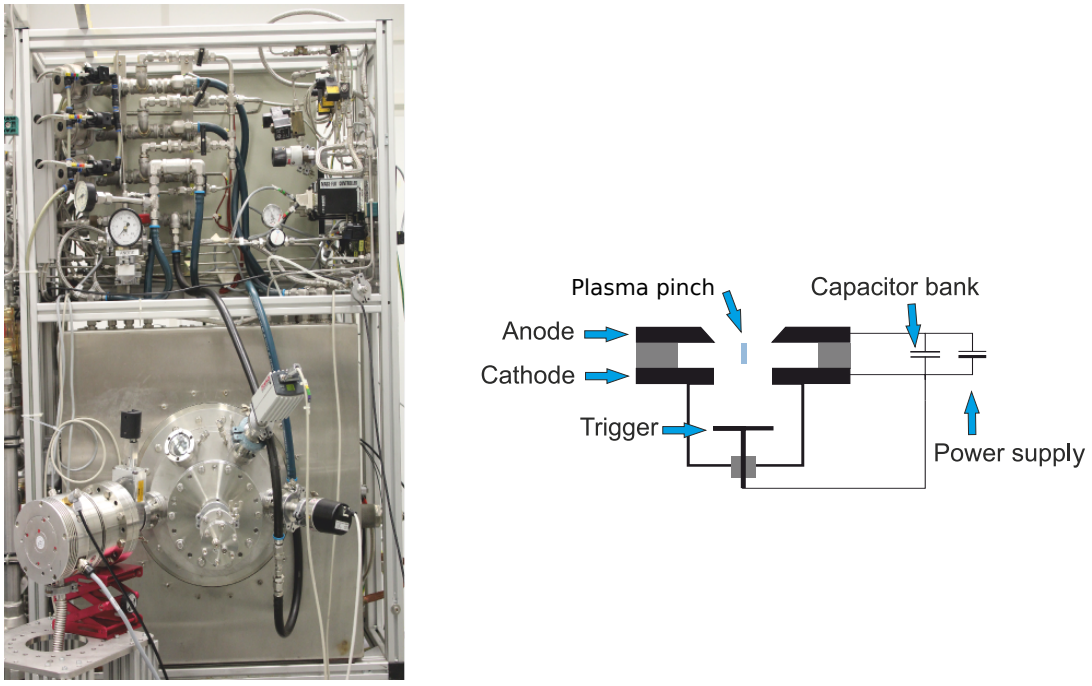

FIGURE 6.1: The left image is showing the used Philips EUV source and the right one shows schema of hollow cathode producing plasma pinches.

For our second ptychography experiment, we have used laser assisted discharge produced plasma (LDP) source [19, 188] from the USHIO company. This LDP source uses an electric plasma discharge between two rotating electrodes. The surface of one of the rotating and tin covered electrodes is first evaporated by a laser pulse and then pinching is done by electric discharge between the electrodes. Compared to laser produced plasma sources (LPP) [34, 57], LDP has the advantage of direct conversion of electric power into photon energy and it can provide higher brilliance. LDP sources are also less limited by erosion of electrodes than GDP sources. On the other hand, LDP sources produce more debris, therefore need

better debris mitigating filters and complicated cleaning of the front optics in order to remove the deposited tin layer. Note that the GDP sources produce lower but still nonnegligible amount of debris. The used LDP source was operated at 4 kHz but it could be operated with up to 12 kHz frequency. The source can produce up to 36 W/(2πsr at the position of the intermediate focus (behind the collector optics) [19].

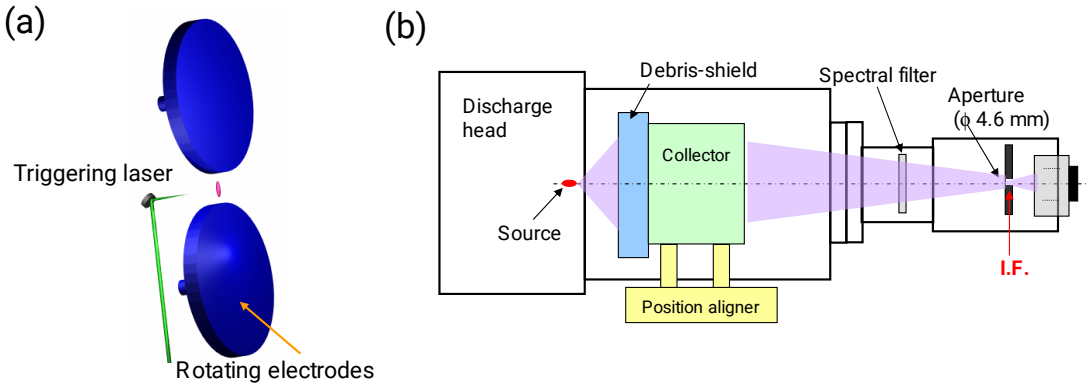

FIGURE 6.2: Simplified schematics of the LDP laser assisted source. (a) shows the rotating electrodes and triggering laser. (b) shows a schematic of the LDP source collector module. Figure adopted from Ref. [18].

EUV and SXR light in plasma is produced by several mechanisms:

1. Bremsstrahlung: Continuous radiation that is emitted by the free electrons decelerated in the Coulomb field of ions. This is an important radiation source in fully ionized plasma i.e. in the fusion experiments.

2. Recombination radiation: Free electrons recombining with ions emit the binding energy as light. This radiation produces a semi-continuous spectrum due to free-to-bound transitions.

3. Spectral line radiation: Electrons gets excited to higher energy states by collisions between ions and electrons or by high energy photons. Photons are radiated during transitions to lower energy levels. Because of discrete energy levels in atoms, the radiation spectrum consists of narrow spectral lines that are dominantly broadened by the Doppler effect.

The maximum power radiated by any material in equilibrium is limited according to Kirchhoff's law of thermal radiation by the power of a black body source at

the same temperature. The maximum spectral radiance is given by the Planck formula

$$L_\lambda(\lambda, T) = \frac{2hc^2}{\lambda^5} \frac{1}{\exp\left(\frac{hc}{k_B T \lambda}\right) - 1} \quad . \tag{6.1}$$

The optimal plasma temperature to observe radiance at 13.5 nm (92 eV) is therefore 18.5 eV. In the case of GDP sources, the right gas fuel needs to be selected. The choice of the working gas is limited by achievable breakdown voltage. The GDP EUV source was designed to produce EUV at 13.5 nm using Xe with pulse energy 500 μJ. However, the produced spectrum is relatively broad. In the case of oxygen fuel, the produced energy is only 30 μJ at 17.3 nm. We have used oxygen plasma surrounded by argon in order to increase the working gas pressure which resulted in reduced breakdown voltage. Oxygen radiates in several narrow lines while radiation of Xe is more broadband (see Fig. 6.3)(a). The spectral width of a single peak is dominated by the Doppler broadening

$$\Delta f_{\mathrm{FWHM}} = \sqrt{\frac{8 \ln 2 kT}{mc^2}} f_0 \quad .$$

This would results in very narrow bandwidth $\approx 10^{-4}$. However, the intense spectral peaks are often doublets or triplets for lighter elements and in the case of heavy ions such as xenon, the density of lines is so high that it is impossible to resolve the individual lines.

hence the bandwidth is limited by distance between the lines. In the case of oxygen plasma at 17.3 nm, the estimated bandwidth is $\approx 10^{-2}$. For heavier elements such as Xe or Sn, the spectral lines of individual peaks are practically indistinguishable and the emission is very broadband as it can be seen in Fig. 6.3(b).

Both EUV CDI setups presented in this chapter were built in cooperation with Jan Bußmann and Raoul Bresenitz from RWTH Aachen university, however I created all the data acquisition scripts and ptychography reconstructions.

A simplified schema of our setup is shown in the Fig 6.4. The light was first spatially filtered by a 500 μm pinhole and then reflected and spectrally filtered by a B_4C/Si multilayer mirror. Moreover, the produced light was filtered by argon gas that is surrounding the oxygen pinch in order to increase oxygen pressure. Spectral filtering of each component is shown in the Fig. 6.5.

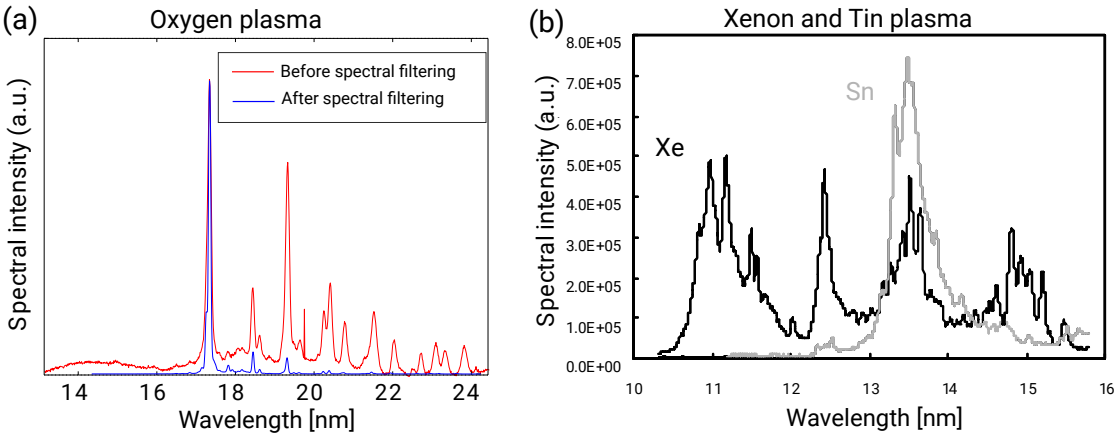

FIGURE 6.3: (a) Measured spectrum of our EUV source with B_4C/Si multilayer mirror and without. Both spectra were normalized to their maximum intensity, (b) an example of the EUV spectrum from Xe and Sn fuel based sources. Figure (b) was adapted from Ref. [19].

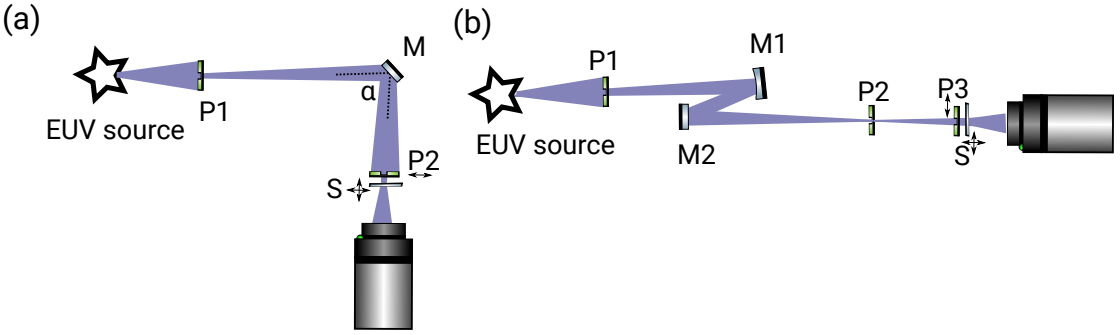

FIGURE 6.4: (a) simplified schema of the CDI setup used with the gas discharge plasma source. P1 is a filtering pinhole with 500 μm diameter, M is filtering multilayer mirror and P2 is pinhole that produces the structured illumination on the sample S. Setup (b) was used with the laser assisted discharge plasma source. Pinholes P1, P2 were used to adjust the coherence and P3 to illuminate the sample S.

6.1 Ptychography with GDP source at 17.3 nm

In contrast to HHG EUV sources, the EUV produced by plasma source is practically fully incoherent. The light is produced during compression of the plasma pinch by the magnetic field. Smaller pinch size will result in higher brightness and thus less filtering is required to get the same coherence length. In the first proof of concept ptychography experiment with oxygen fuel, the pinch diameter, as seen from the sample, was estimated to be between 0.8-1 mm. Plasma produces light at a broad range of wavelengths from IR to EUV. Therefore, the spectrum was filtered by reflection from a B_4C mirror (Figs. 6.4(a), 6.3). The angle of reflection was optimized in order to maximize the reflectance at 17.3 nm. Wavelengths in the

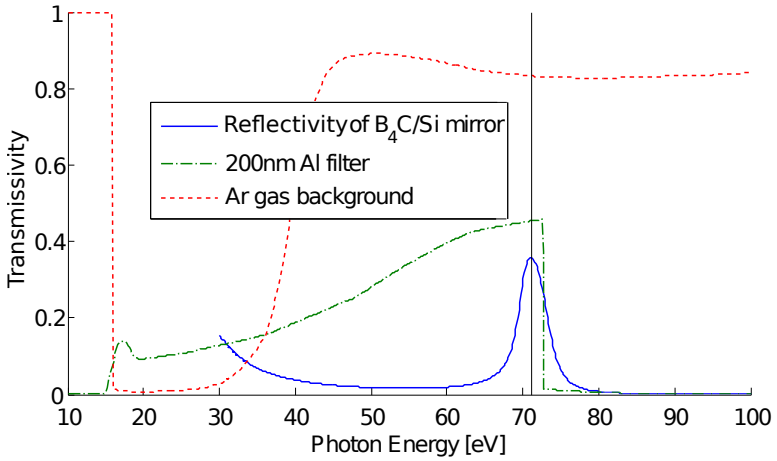

FIGURE 6.5: Transmission of different filtering elements in our setup and reflectivity of B_4C/Si mirror [9]. Vertical line denotes position of 17.3 nm oxygen line.

DUV and UV range were further filtered by the argon gas surrounding the source. Finally, a 200 nm thick aluminium filter could be added to filter the visible light, however it was not used in our imaging experiments in order to get higher EUV photons flux (Fig 6.5). The consequence was a low amount of visible light passing through all the filtering pinholes and the sample as can be seen in the Fig. 6.6.

CDI imaging is basically limited by flux and coherence. However, both these factors are linked, because if the flux is sufficient, then it is possible to get better coherence by spatial filtering of the beam. We have used a pinhole with diameter 500 μm in distance 1.2 m from the sample. There would be no gain in using a larger pinhole in our setup because the spatial coherence would be already dominated by the pinch size and the flux gain would not compensate the loss of coherence.

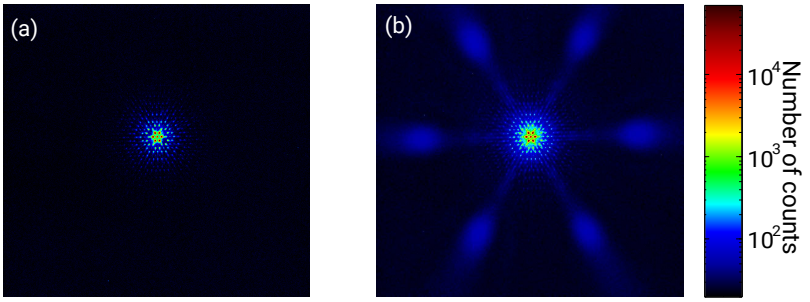

FIGURE 6.6: An example of a diffraction pattern from a hexagonal grid with (a) and without (b) aluminium filter. The image (b) shows visible light radiation in background that is creating the large smooth speckles far from the centre.

Using the transverse coherence length defined in Section 2.2.2, the coherence is $l_{\text{transverse}} = 9\,\mu$m, therefore the maximum sample size for CDI should be below 5 μm

according to [48], however ptychography can accept more incoherence, therefore the illumination size was set to 10 μm. Using the estimated bandwidth of 1%, the longitudinal coherence is $l_{\text{longitudinal}} = 1.7\,\mu$m (Eq. 2.47), which roughly limits the maximum extent of the illuminated probe to ≈200 px (Eq. 2.52).

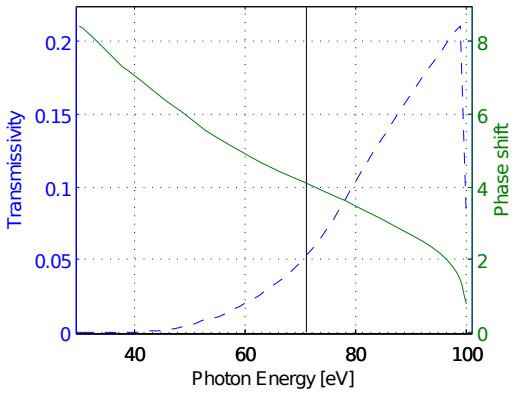

FIGURE 6.7: Phase shift and transmission for a 200 nm thick silicon nitride foil calculated from [9].

The proof of concept ptychography reconstruction was performed using simple relatively high contrast sample. The imaged object was 200 nm thick silicon nitride holey grid from PELCO. The circular holes with 2.5 μm diameter were arranged in a hexagonal pattern. The transmission through the foil was only 5.5% (see Fig. 6.7), however it is still better than transmission at 29 nm wavelength that is mere 0.1%. We have used a regular grid scan with 10×10 positions, 2 μm steps. The random offset between the positions was set to 10% in order to avoid periodic artefacts. The exposure time was 60 s per diffraction pattern. The diffraction pattern was collected by EUV sensitive camera (Andor iKon M). In order to reduce the incoherence effects caused by low spatial coherence and slightly polychromatic spectrum, we have applied the online deconvolution methods mentioned in the Section 2.5 in combination with background subtraction to remove effects of the visible coherent background light [20]. The final reconstruction is show in Fig. 6.8. The pixel size is 226 nm, however the real resolution depends on the transmission and varies from 250 to >500 nm.

The reconstruction (Figs. 6.8,6.9) is clearly showing thickness (or composition) variations in the flat region around the grid. Small imperfections of the silicon nitride foil (top left corner) are visible in both the amplitude and phase but the phase information is less noisy. The presented dataset is available at DOI:10.5258/SOTON/382965.

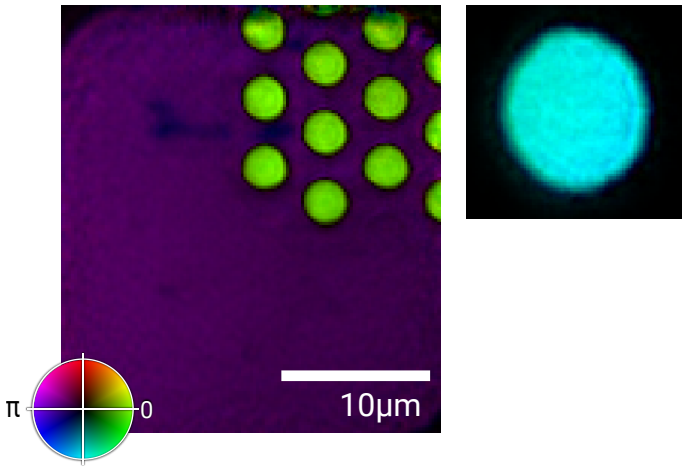

FIGURE 6.8: A ptychography reconstruction of a hexagonal grid at 17.3 nm wavelength shown in the complex colour scale. Results published in Ref. [20].

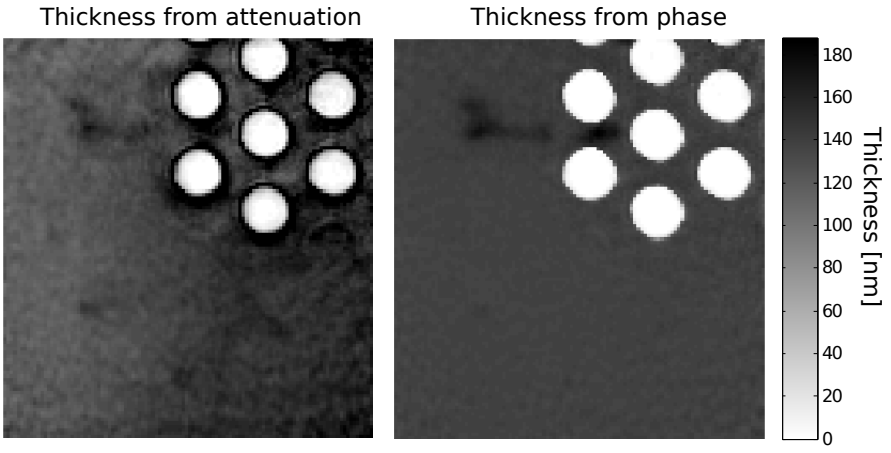

FIGURE 6.9: Comparison of the silicon nitride foil thickness estimated from amplitude (left) and phase (right). Thickness estimated from amplitude reveal ramp up from left to right that is probably caused by slow changes in the EUV source power over the scanning time.

6.2 Ptychography with LDP source at 13.5 nm

The previous experiment was only a first ptychography test with incoherent EUV sources, however despite the fact that it was the first ptychography with a tabletop source at wavelength below 30 nm, the selected wavelength 17.3 nm had relatively low significance. Moreover, the exposure time and resolution were not competitive with HHG sources.

Our second experiment was performed with a tin based laser assisted discharge plasma source. A simplified schematic of the setup is shown in Fig. 6.4(b). The spectral bandwidth was reduced to 2% by the Si/Mo mirrors optimized for 13.5 nm

wavelength. The visible and UV light was filtered by a 200 nm thick Zr filter. The spatial coherence could have been optimized by the pinhole P2, however due to the technical constraints of the light source, the spatial coherence at the sample position was larger than required (i.e. $l_{\mathrm{coh}} \gtrsim 20\,\mu$m), which resulted in less effective use of the available flux. Furthermore, the spectral filtering tool (mirrors M1,M2) was not possible to adjust for our application, which resulted in further significant reduction of the available flux. The transmission through the spectral filtering mirror system was below 2%.

The ptychography spiral scan consisted of 194 positions with scan step 3 μm. The exposure time per a position was 30 s. The diffraction patterns were collected by an EUV sensitive camera (Andor iKon M). The sample was illuminated through a pinhole with 10 μm diameter placed 395 μm from the sample. The reconstructed pixel size was 77 nm, however the real resolution was flux limited to around 100 nm. The total flux at the sample position was $5{\cdot}10^5$ photons/s. This is still 400× lower compared to the HHG flux at 30 nm. The final reconstruction of the polyspheres testing sample is shown in Fig. 6.10. The sample was prepared in identical way as the HHG testing sample in Section 5.7. The low resolution compared to the size of the spheres limits the possible analysis, however the magnified region shows that the polyspheres were partially transparent.

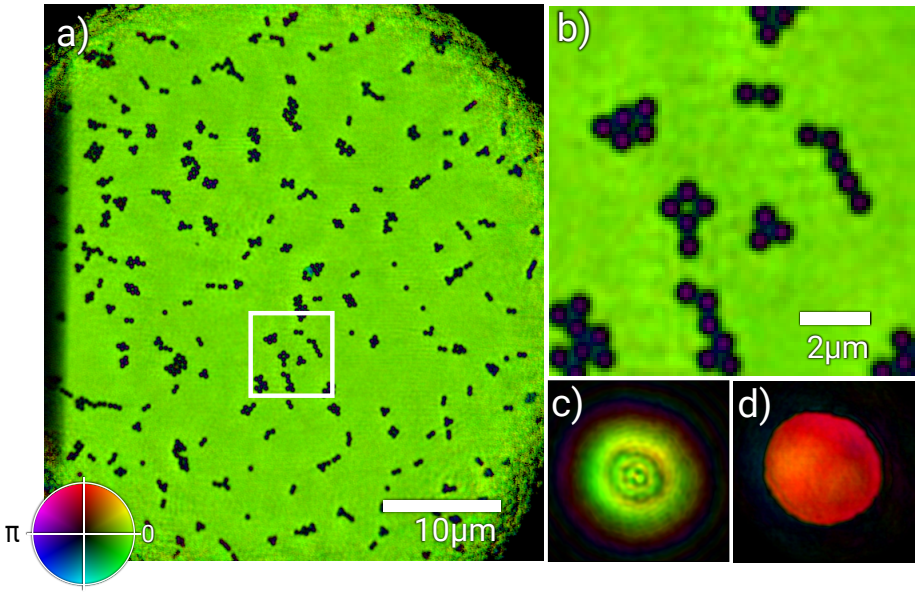

FIGURE 6.10: Ptychography reconstruction of complex transmission through a polyspheres sample. (a) shows the full reconstruction and (b) contains a magnified central region. (c,d) are the illumination probe at the sample position and the probe back-propagated 395 μm at the pinhole plane.

6.3 Conclusion

In this chapter, we have presented the first results of ptychography with laboratory incoherent EUV sources. Both experiments were only proofs of concept without real applications because the available coherent flux is too low for reasonably high resolution microscopy.

One of the main issues is that the light produced by these sources is radiated into all directions, in contrast to the HHG sources, where all the power is confined in one direction and the peak brightness is therefore much higher. Although the incoherent sources can provide tens of watts of EUV power, the resulting coherent flux on the sample is roughly 3 orders of magnitude smaller than for the HHG source that we have tested. On the other hand, the tested HHG source was at 30 nm wavelength and the available HHG flux at 13 nm wavelength is significantly lower [56].

In future, we expect that the coherent flux requirements can be reduced by using high NA illumination that could provide reasonable high resolution with significantly lower coherence requirements [38].

Chapter 7

Conclusions and future work

The main goal of this thesis was to demonstrate that laboratory short wavelength sources can be used for high quality coherent diffractive imaging. We believe that this was well shown in Chapter 5, where we have shown high resolution reconstructions of hippocampal neurons at 30 nm wavelength.

Laboratory sources provide much lower coherent flux than the large facilities like synchrotrons, however their price and availability is much better. Despite this, there are still many things that need to be improved, before it can become a valuable tool for biology and material sciences. Currently the main limitations of our high harmonics generation based CDI microscope are:

- low long term stability: the pulse duration and intensity stability of our laser is not sufficient for very long scans, or for repeatability of a single scan under different conditions. This is necessary if 3D reconstruction methods such as laminography would be used or in the case of pump-probe ptychography imaging of evolving systems.
- low spatial resolution: we have reached ≈50 nm resolution for biological samples, however in order to significantly outclass the visible light super-resolution techniques such as PALM and STORM that can reach also resolution below 50 nm, it is necessary to further increase the numerical aperture and the dynamic range.
- slow scanning time: currently most of the scanning time is spend by readout time from our CCD EUV camera. However, the newer EUV CCD cameras

can provide similar readout noise with faster readout speed or they can also readout several sectors of the camera simultaneously.

- low photon flux: in particular for thicker biological samples, the photon flux is limiting the achievable resolution. In future, it will be possible to use high repetition rate fibre ultrafast lasers with more than 30 kHz repetition rate [178].

- low penetration: penetration depth is one of the common issues for EUV that comes as a price for good natural contrast. The best option is to move the imaging to shorter wavelengths as it was already done with HHG sources [56, 178, 187, 195], where the penetration is improved. On the other hand, due to lower contrast of the samples, the requirements on the photon flux and collected dynamic range get even more serious.

Some of these issues can be improved by better source tuning or some small improvements in the imaging system such as fast movable beamblock. A lot of the stability issues were reduced by our OPRP method introduced in Section 2.7 and by our new gradient based position correction method 2.6.2. However, some of the issues will require further development of high power ultrafast lasers. In future, it may be possible to use the full potential of the HHG sources and realize multicolour ptychography. The first but still very preliminary attempts were already published [151, 196, 197]. Another option is pump-probe ptychography that could make possible imaging with tens of femtosecond time resolution thanks to perfect natural synchronization with the driving IR laser. It is also possible to perform magnetic contrast phase imaging using the circularly polarized high harmonic beam [198, 199] or 3D topographical mapping of surfaces using the laminography method in reflection mode. It would be possible to use HHG based ptychography for polarization sensitive microscopy [200] using a flat 45° multilayer EUV mirror between the sample and camera as a polarization filter. Finally, EUV microscopy is a valuable tool for EUV lithography mask inspection because some kinds of defects cannot be observed at different than the lithography wavelength 13.5 nm. Possible inspection setups are shown in Fig. 7.1.

In Chapters 3 and 4, we have shown novel methods for imaging of sample anisotropy and non-linearity. These methods are not easily applicable in the short wavelength regions, however they can prove to be a valuable tool for phase imaging of non-linear samples. Anisotropy extension of ptychography can be also useful for high

NA imaging of multilayer samples in the EUV range, when the reflectivity of the surface depends on the illumination angle.

Finally, in Chapter 6, we have demonstrated ptychography with incoherent plasma based EUV sources. The achieved resolution and photon flux was not competitive with the HHG sources. This happened because the far-field diffraction geometry is well suitable for the coherent laser like sources, however it is not using the full potential of the weakly coherent sources. One way how to use more of the incoherent power is to do the imaging in the effectively near-field mode as it was tested in Section 5.9 for the HHG source. Near-field mode will alleviate the coherence requirements and allow to use larger amount of the incoherent flux. Moreover, because the measured noise will be dominated by Poisson noise, it is possible to use the fastest readout speed of the camera and reduce the overhead. The resolution of the effectively near-field mode is approximately limited by the geometrical scaling of the illumination beam, therefore in order to get high resolution, it is necessary to use highly diverging illumination without losing flux by using diffractive optics and at the same time guarantee the long working distance that is needed for reflective microscopy. This is difficult to accomplish in the X-ray range, because only grazing incidence mirrors can be used, however in the EUV range, multilayer mirrors that can provide the illumination NA more than 0.2 are available. Moreover, ptychography should not be so sensitive to the quality of these mirrors as long as they are used only for illumination of the sample, therefore the optics can be relatively cheap. The second and more expensive option is to use high quality optics behind the sample and to perform the imaging slightly out of focus. This would also result in the near-field ptychography task with coherence requirements significantly reduced. Examples of such setups are shown in Fig. 7.1.

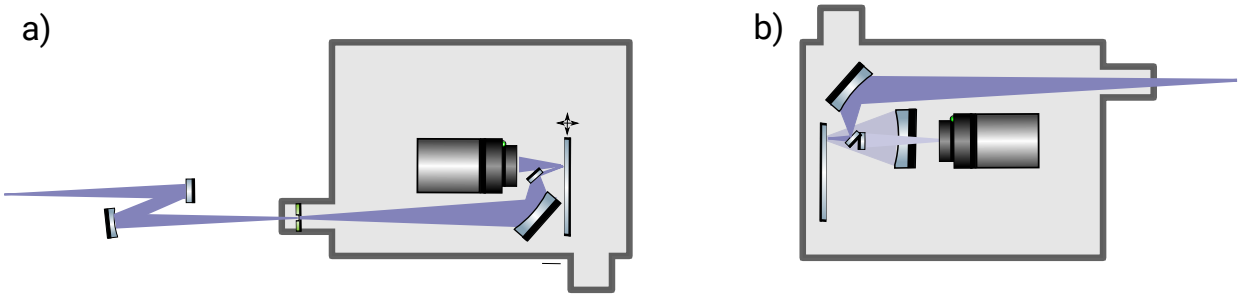

FIGURE 7.1: Examples of the near-field EUV ptychography setups designed for lithography mask inspection in the reflection mode. The first example uses high NA illumination optics that can have relatively low quality. The second example (b) is using optics behind the sample, therefore high quality, high NA optics is required and small pixels on the camera or larger distance of the camera from the optics is needed.

Appendix A

Beam stabilization systems

Although regenerative amplifier based laser systems are generally considered very stable, the CDI applications have high very high demands on the overall beam stability. Good pointing stability allows longer scans, better day to day experiment repeatability and in the case of long exposures it leads to better visibility of the collected diffraction patterns. However, stability of the produced HHG beam is not given only by the pointing stability of the driving laser, but if the beam quality is low, intensity and pulse variations are changing the pointing stability of the HHG beam as well. M^2 of our laser was estimated to 1.8, hence this effect cannot be neglected for our setup.

The beam stability improvements can be separated into two groups: passive and active. Passive improvements consist mainly of better sealing of the beam path in order to avoid unwanted air movements that can disturb the beam direction and improving the air-conditioning system. Air movements are usually very fast and produce instabilities with a frequencies higher than several Hz. On the other hand, it is relatively easy to avoid them almost completely. Further passive stability improvements can be achieved if all sources of vibrations such as heavy prepumps are moved out of the laser laboratory and placed on damping pads.

The active beam position stabilization systems usually use a combination of two actuators and two beam position detectors. In the case of continuous wave lasers, very fast systems based on quadrant photodiodes (QPDs) and piezo-actuated mirrors are available [201]. The most common configuration includes two actuated mirrors and two QPDs that are used to stabilize the beam spatially and angularly. The stabilization is more difficult if a laser with a lower repetition rate (our Spitfire

amplifier produces only 1 kHz) is used. If the stabilization system is not designed to stabilize laser beam with limited repetition rate, the stabilization performance can be poor [202].

We have developed a stabilization system optimized for pulsed lasers with pulse rate around 1 kHz with high pointing precision requirements.

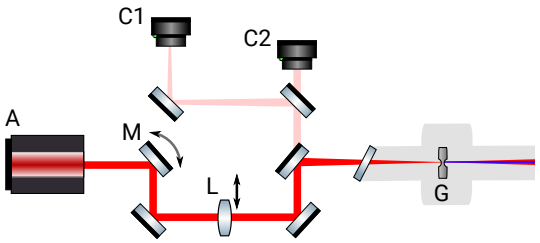

FIGURE A.1: A simplified schema of the stabilization system used in the current HHG imaging setup. The IR laser (A) beam is reflected by an actively stabilizing mirror (M) and passes though an active lens (L). The lens (L) focuses beam onto a gas-cell (G), where EUV is generated. A small fraction of the IR beam passes through the back of a dielectric mirror and is collected by cameras C1, C2 at different positions.

Geometry of the stabilization system

The geometry of our stabilization system is shown in the Fig. A.1. The beam is controlled by one piezo-actuated mirror (M) with angular step 0.2 μrad (Picomotor piezo linear actuators 8302 in a custom made mirror holder) and by an actively stabilizing lens (L) on a XY stage with 0.1 μrad minimum step size also driven by Picomotor actuators. Distance between the active mirror and lens is approximately 5 m in order to more decouple their movements. The lens movements produce mainly lateral shift of the beam at the gas-cell position, while the mirror causes both lateral and angular movements. The main beam is reflected by a multilayer mirror to the gas-cell and only a small fraction ($< 0.1\%$) passes through the back side of the dielectric mirror to the detectors (C1, C2). We have used Thorlabs DCC-1545M CMOS cameras as the detectors, however the systems can work with any sufficiently fast camera. Camera C2 was placed at sufficient distance to capture the whole IR beam on the chip. No additional lenses were used, only neutral density filters in order to avoid oversaturation, reduce background light. The second camera (C1) was placed behind an additional beam splitter (10:90) and positioned close to the focal position of the beam. This allows us to directly estimate the pointing stability at the gas-cell position. This configuration of the

detectors guarantees good decoupling of the beam movements and therefore robust stabilization performance.

Software implementation

The signal from both CMOS cameras was collected and processed by our self-made stabilization code written in Python and running on a Raspberry PI2 microcomputer. Only a limited region (camera C2) and a down-sampled signal (camera C1) were read out from the cameras in order to increase the maximum frame-rate and processing speed. The beam position was calculated as a centre of mass of the measured centroids. Before the calculations, the background signal was subtracted and median filter was applied in the vertical dimension. Median filter helps to remove artefacts caused by the pulsed nature of the laser that causes different number of pulses collected by different lines of the chip. The recommended exposure time should be roughly around 10 pulses (i.e. 10 ms). Further, our system automatically selects the optimal exposure time very close to the multiple of the pulse delay. The required precision of the exposure time setting is better than 0.2%. This is crucial in order to avoid beating of the readout and laser frequency, which would result in fast oscillations of the measured centre of mass of the beam. Another option to avoid this issue, would be to use better cameras with global shutter and external trigger; however the precision of the beam position estimation was not limiting our stabilization system. The analogue gain is automatically adjusted to stretch dynamic range of the collected signal. Also it is important to use fast readout time in order to reduce the acquisition time.

Numerically, the stabilization task is treated as a system of linear equations. Movement of the actuators by distance $\delta\mathbf{X}$ will result in a shift of the detected signal by $\delta\mathbf{Y}$. The system of equations can be written as

$$\delta\mathbf{Y} = \mathbb{A}\delta\mathbf{X} \quad . \tag{A.1}$$

If all relevant parameters of the system are known, i.e. distances of the mirrors, step sizes of the actuators and focal length of the lens, it is possible to calculate matrix $\mathbb{A}$. However, in practice $\mathbb{A}$ is identified during an automatic calibration procedure. Response of detectors on the movement of each of the mirrors in calculated and the matrix $\mathbb{A}$ is solved using the least squares method for multiple measurements. The optimal correction response $\delta\mathbf{X}$ of the actuators can be then

obtained as

$$\delta\mathbf{X} = \mathbb{A}^{-1}\delta\mathbf{Y} \quad . \tag{A.2}$$

The matrix $\mathbb{A}$ needs to be as well conditioned as possible otherwise small errors during the calibration procedure may result in large errors in the inversion of the system matrix and the stabilization will be suboptimal.

The equation A.2 is sensitive to noise in the measured data and therefore usually a digital proportional integral derivative (PID) controller is implemented [202, 203]

$$\mathbf{X}_{\text{next}} = \mathbf{X}_{\text{prev}} + K_p\delta\mathbf{X} + K_i\int \delta\mathbf{X}dt + K_d\frac{\delta\mathbf{X}}{dt} \quad . \tag{A.3}$$

The derivative term K_d is helping to speed up convergence by predicting the future step, integration term K_i includes "memory" of the system and proportional constant $K_p \in (0, 1)$ is just a smoothed estimation of the correction term. Our test of the stabilization performance, however showed that the derivative term is never improving results and the optimal integration constant is usually negligible as well. This is not so unexpected because the stabilization system has no memory and with noiseless ideal data and infinitely small step size of the motors, the position would be fully corrected immediately in the following step. The optimal proportional term depends mainly on the amount of fast vibrations in the signal. The optimal value is usually selected between $K_p = 0.3$ and $K_p = 0.8$ in order to minimize the beam movements on the camera. The stabilization rate of our system is between 5 to 15 Hz dependent on the motors and camera settings. It is roughly equally limited by response time of the piezo-actuator controllers and by the readout time of the cameras. The processing time by Raspberry PI2 is approximately 10% of the step time.

Figure A.2 shows an example of vertical position of the IR beam on the C2 camera, which is equivalent to gas-cell position movements. The longer time stability is significantly improved with the stabilization but the short term "smoothness" is worse. This is mainly caused by imperfections of the mirror mounts, which are not able to slide smoothly with the extremely high precision that is required. The average beam position stability at the gas-cell position is around 0.1 μm. The stabilization could be improved if fast piezoelectric actuators were used instead of cheaper but slower picomotors based positioners.

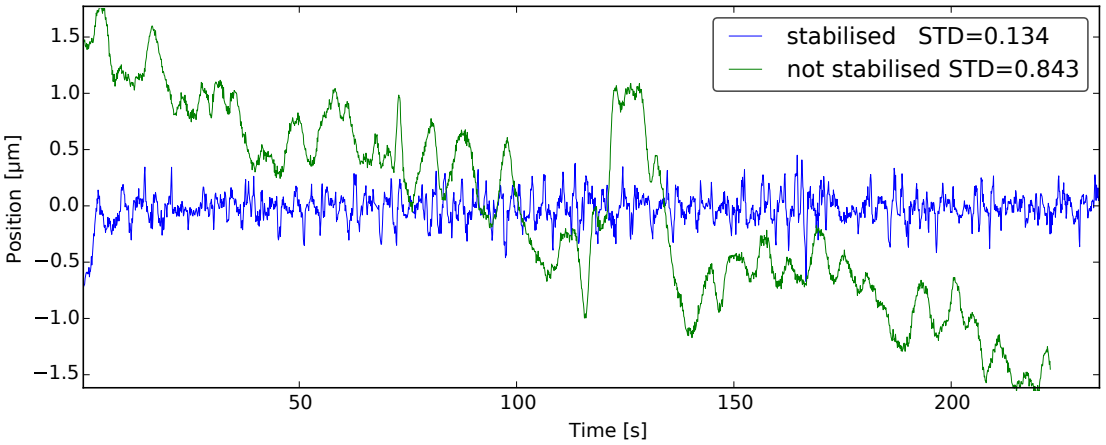

FIGURE A.2: Vertical position of the IR beam on the C2 camera with stabilization (blue) and without (green). The stability over longer time periods is significantly improved.

The pointing stability of the HHG beam is crucial for CDI applications. HHG beam in the first approximation inherits the same pointing stability as the driving IR beam but it has much lower divergence (≈2 mrad compared to ≈18 mrad for IR beam), therefore the ratio of beam diameter versus position stability is significantly worse. Despite this, the pointing stability of the EUV beam, calculated directly from IR stability, should be smaller than 0.01 μm at the pinhole position, which would be more than sufficient for our applications. However, this estimation does not correspond to our measurements of the EUV stability, which are showing drifts up to 2 μm over 4 hours period. We assume that the residual EUV beam drifts are caused by changes in the IR pulse duration and pulse energy. These effects are resulting in the movements of the EUV beam through the phase-matching process.

Appendix B

Nano-precision sample positioning system

One of the major issues in high resolution ptychography is precise knowledge of the sample position with respect to the illumination optics (pinhole) position. This issue is shared by any high resolution microscopy, however ptychography is using the precise knowledge of sample position for the realspace constraint and therefore the reconstruction quality is particularly sensitive. This issue was well known and different numerical solutions were proposed in multiple publications [89–91, 95, 114, 149]. Also our OPRP method and gradient based method (Sec. 2.6.2) are able to correct small positioning errors. Despite this, it is still preferable to reduce the positioning errors, because the position correction algorithms work well only with high contrast sample and high quality datasets - e.g. high coherence and no systematic errors caused by the camera such as over-saturation artifacts.

Here, we present our simple positioning feedback system based on white light that allows to reach 0.2 nm short term accuracy. Long term accuracy is limited by thermal drifts to roughly 2 nm.

Our sample positioning system is using 3D nanopositioning piezoelectric actuators Smaract 1730 that should have sub-nm scanning resolution and absolute accuracy 1 μm/m according to the manufacturer. The position of the stages is automatically measured by proprietary optical encoders built in the actuators.

In order to estimate reliability of the Smaract optical encoders, we have built our own external optical positioning system. A simplified schema and real image are

shown in Fig B.1. A diffuser optics mounted directly on our sample holder was back-illuminated by a white LED. The illumination angle needs to be sufficiently narrow in order to create high contrast sharp features in the imaged pattern. The collected image (Fig. B.2) was collected by a standard 8-bit CMOS camera (DCC1545M). Resolution is diffraction limited to 0.75 μm.

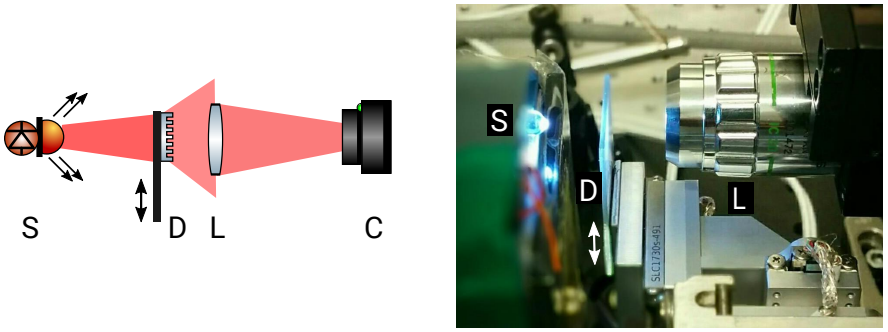

FIGURE B.1: A schema of our simple nano-position feedback system. A light source S (white LED) is illuminating a diffuser D. Scattered light from the diffuser is collected by a microscope lens L (Olympus, NA=0.4, 20×) and imaged on a CMOS camera (DCC1545M).

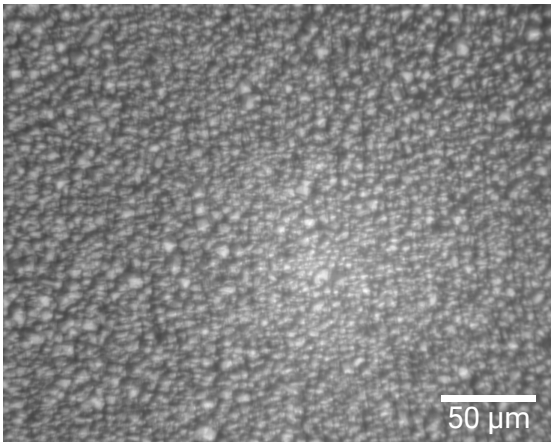

FIGURE B.2: Back-illuminated diffuser imaged by our positioning system.

Relative position of the diffuser was estimated using a sub-pixel precision cross-correlation. Originally, the subpixel cross-correlation proposed in Ref. [115] was tested, however from our tests we found that simple normalized cross-correlation X_{corr} can be used to get faster and more accurate results. Normalized cross-correlation

$$X_{\mathrm{corr}}(x_0, y_0) = \frac{1}{n} \sum_{x,y} \frac{\left(f(x,y) - \overline{f}\right) \left(g^*(x - x_0, y - x_0) - \overline{g^*}\right)}{\sigma_f \sigma_g} \quad , \tag{B.1}$$

between template f and image g, where $\overline{f}$ denotes average of f and σ_f is the standard deviation of f was calculated using FFT. Instead of taking positions of maximum to estimate the shift as proposed in Ref. [115], we have selected a small region (11×11 px) around the maximum. The minimal value in this region was subtracted and the shift between template and image was calculated as center of mass of squared intensity in this small region.

Sensitivity of our system was tested using sinusoidal movements of the actuators with increasing amplitude as shown in Fig. B.3. The first 100 positions were collected with the piezoelectric actuators switched off. After activation of the actuators, the noise is significantly increased, however this is caused by the limited precision of the active Smaract position feedback. The maximal achievable precision of our system can be estimated from the initial passive signal, after subtraction of the long term thermal drift, to 0.17 nm. This precision is more than sufficient to test accuracy of the Smaract nanoprecision actuators used for our ptychography experiments, where only precision around 30 nm is required.

The Smaract actuators exhibit significant jumps during linear movements caused by their stick-slip shift principle each 1.5-1.8 μm as it is shown in Fig B.4. This is causing jumps in the perpendicular dimension to the direction of movement with amplitude ≈50 nm. Resolution of our ptychography experiments was usually around 50 nm, therefore the error between adjacent positions is up to 2 px. However, also in the direction of movements, an error with period of roughly 10 μm and amplitude 100 nm can be observed leading to the worst case error up to 4 px.

Another source of positioning errors is thermal drift. We have identified several main sources of thermal variation that may result into drifts up to hundreds of nanometers

1. Smaract positioning sensors - if the Smaract optical sensors are permanently running, it can increase temperature of the sample plate by up to 3°C. However, our system is designed to keep the position sensors switched off, when the motors are not moving. This can be however still an issue if several large and quick mapping scans were taken before the ptychograhy scan was initialized.

2. Andor camera - our Andor camera is water cooled, however it is also conductively connected to the rest of the CDI microscope. Therefore, Peltier

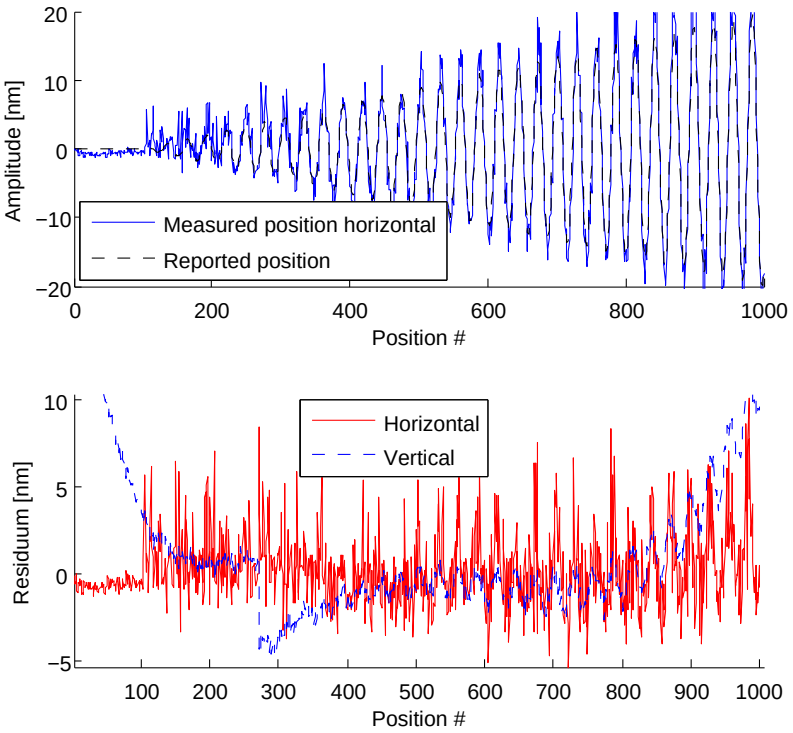

FIGURE B.3: Sensitivity tests of our feedback system based on moving the Smaract actuator in horizontal direction with growing amplitude. The top plot shows position reported by the Smaract system (dashed) and measured position from our system (full line). The actuators were switched off during the first 100 positions and then the position was sinusoidally modulated up to amplitude 20 nm in horizontal direction. The bottom plot shows residuum between reported and measured positions for horizontal and vertical directions. This different misfit effects are shown in Fig. B.4.

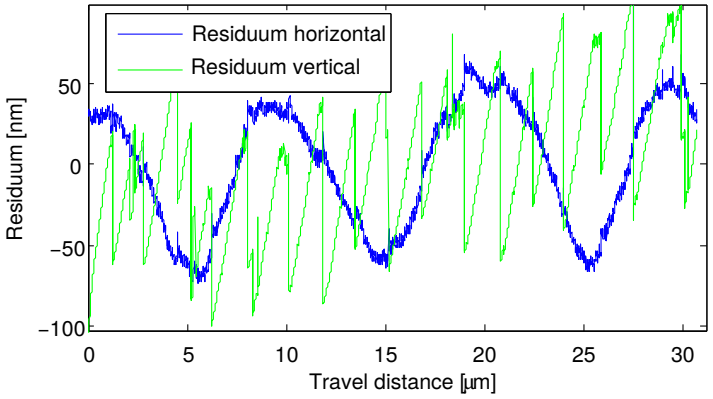

FIGURE B.4: Residuum between the requested position and the measured position during linear movement in the horizontal direction. The perpendicular direction to the requested linear movement (i.e. vertical direction) shows sudden jumps cased by the stick-slip shift principle. These jump are automatically compensated by the included Smaract feedback in the horizontal direction, however still some position imperfections with period of $\approx 10\,\mu$m are can be observed.

cooling of the camera can be a significant heat source, despite larger distance from the sample of the water cooling flow is insufficient.

3. Vacuum pump down - during pumping down the chamber, the remaining gas in vacuum chamber has lower temperature than the rest of the metal construction. This "heat shock" is caused due to large pumped volume and slow pumping speed which result in heat exchange between the remaining gas and the microscopy setup. We have observed that the setup needs roughly 2 hours after pump down to return into thermal equilibrium.

4. Air-conditioning in lab - air temperature fluctuations can be still seen in vacuum by our feedback system (Fig B.5) despite the vacuum chamber weighting over 500 kg and thermal connection between chamber and our setup being minimized.

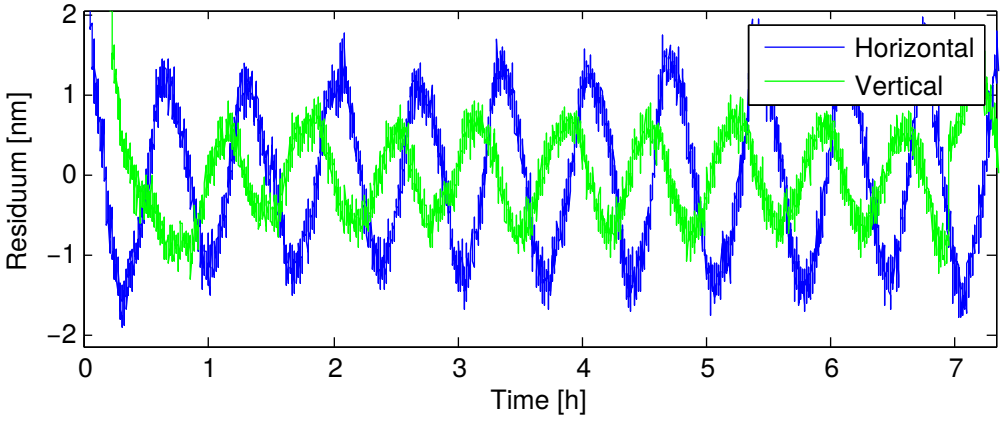

FIGURE B.5: Long-term measurement of sample position is showing periodic oscillations caused by lab air-conditioning system. Note that the feedback system was placed in vacuum inside a steel vacuum chamber weighting roughly 500 kg. These position variations are negligible for our imaging setup, however the subnanometer precision of our position estimation system is clearly visible.

Bibliography

[1] J. Miao, P. Charalambous, J. Kirz, and D. Sayre, "Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," *Nature*, vol. 400, no. 6742, pp. 342–344, 1999.

[2] G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution Fourier ptychographic microscopy," *Nature photonics*, vol. 7, no. 9, pp. 739–745, 2013.

[3] G. Zheng, X. Ou, R. Horstmeyer, J. Chung, and C. Yang, "Fourier ptychographic microscopy: A gigapixel superscope for biomedicine," *Optics and Photonics News*, vol. 25, no. 4, pp. 26–33, 2014.

[4] M. Guizar-Sicairos, I. Johnson, A. Diaz, M. Holler, P. Karvinen, H.-C. Stadler, R. Dinapoli, O. Bunk, and A. Menzel, "High-throughput ptychography using eiger: scanning X-ray nano-imaging of extended regions," *Optics Express*, vol. 22, no. 12, pp. 14 859–14 870, 2014.

[5] X. Ou, G. Zheng, and C. Yang, "Embedded pupil function recovery for Fourier ptychographic microscopy," *Optics Express*, vol. 22, no. 5, pp. 4960–4972, 2014.

[6] M. Odstrcil, P. Baksh, S. Boden, R. Card, J. Chad, J. Frey, and W. Brocklesby, "Ptychographic coherent diffractive imaging with orthogonal probe relaxation," *Optics Express*, vol. 24, no. 8, pp. 8360–8369, 2016.

[7] M. Odstrcil, P. Baksh, C. Gawith, R. Vrcelj, J. Frey, and W. Brocklesby, "Nonlinear ptychographic coherent diffractive imaging," *Optics Express*, vol. 24, no. 18, pp. 20 245–20 252, 2016.

[8] X. He, M. Miranda, J. Schwenke, O. Guilbaud, T. Ruchon, C. Heyl, E. Georgadiou, R. Rakowski, A. Persson, M. Gaarde *et al.*, "Spatial and spectral

properties of the high-order harmonic emission in argon for seeding applications," *Physical Review A*, vol. 79, no. 6, p. 063829, 2009.

[9] B. L. Henke, E. M. Gullikson, and J. C. Davis, "X-ray interactions: Photoabsorption, scattering, transmission, and reflection," *Atomic data and nuclear data tables*, vol. 54, no. 2, pp. 181–342, 1993.

[10] P. D. Baksh, M. Odstrčil, H.-S. Kim, S. A. Boden, J. G. Frey, and W. S. Brocklesby, "Wide-field broadband extreme ultraviolet transmission ptychography using a high-harmonic source," *Optics Letters*, vol. 41, no. 7, pp. 1317–1320, 2016.

[11] M. Odstrčil, P. Baksh, H. Kim, S. Boden, W. Brocklesby, and J. Frey, "Ultra-broadband ptychography with self-consistent coherence estimation from a high harmonic source," in *SPIE Optical Engineering+ Applications*. International Society for Optics and Photonics, 2015, pp. 958 912–958 912.

[12] D. Shapiro, P. Thibault, T. Beetz, V. Elser, M. Howells, C. Jacobsen, J. Kirz, E. Lima, H. Miao, A. M. Neiman *et al.*, "Biological imaging by soft x-ray diffraction microscopy," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 43, pp. 15 343–15 346, 2005.

[13] H. N. Chapman, A. Barty, S. Marchesini, A. Noy, S. P. Hau-Riege, C. Cui, M. R. Howells, R. Rosen, H. He, J. C. Spence *et al.*, "High-resolution ab initio three-dimensional x-ray diffraction microscopy," *JOSA A*, vol. 23, no. 5, pp. 1179–1200, 2006.

[14] M. Van Heel and M. Schatz, "Fourier shell correlation threshold criteria," *Journal of Structural Biology*, vol. 151, pp. 250–262, 2005.

[15] O. Scherzer, "The use of morozov's discrepancy principle for Tikhonov regularization for solving nonlinear ill-posed problems," *Computing*, vol. 51, no. 1, pp. 45–60, 1993.

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

[17] A. Parsons, R. Chapman, P. Baksh, B. Mills, S. Bajt, W. Brocklesby, and J. Frey, "Ultra-broadband support determination for extreme ultraviolet coherent diffractive imaging from a high harmonic source," *Journal of Optics*, vol. 15, no. 9, p. 094009, 2013.

[18] Y. Teramoto, Z. Narihiro, D. Yamatani, T. Yokoyama, K. Bessho, Y. Joshima, T. Shirai, S. Mouri, T. Inoue, H. Mizokoshi, G. Niimi, T. Hosokai, H. Yabuta, K. C. Paul *et al.*, "High-power DPP EUV source development toward HVM," in *International EUVL Symposium*, 2007.

[19] Y. Teramoto, G. Niimi, D. Yamatani, Y. Joshima, K. Bessho, T. Shirai, T. Takemura, T. Yokota, H. Yabuta, K. C. Paul *et al.*, "Development of Xe- and Sn-fueled high-power z-pinch EUV source aiming at HVM," in *SPIE 31st International Symposium on Advanced Lithography*. International Society for Optics and Photonics, 2006, pp. 615 147–615 147.

[20] M. Odstrcil, J. Bussmann, D. Rudolf, R. Bresenitz, J. Miao, W. Brocklesby, and L. Juschkin, "Ptychographic imaging with a compact gas-discharge plasma extreme ultraviolet light source," *Optics Letters*, vol. 40, no. 23, pp. 5574–5577, 2015.

[21] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. CUP Archive, 1999.

[22] H. N. Chapman, "X-ray imaging beyond the limits," *Nature Materials*, vol. 8, no. April, pp. 299–301, 2009.

[23] I. Schlichting and J. Miao, "Emerging opportunities in structural biology with X-ray free-electron lasers." *Current opinion in structural biology*, Aug. 2012.

[24] M. Howells, T. Beetz, H. Chapman, C. Cui, J. Holton, C. Jacobsen, J. Kirz, E. Lima, S. Marchesini, and H. Miao, "An assessment of the resolution limitation due to radiation-damage in X-ray diffraction microscopy," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 170, no. 1-3, pp. 4–12, 2009.

[25] A. M. Maiden, M. J. Humphry, and J. M. Rodenburg, "Ptychographic transmission microscopy in three dimensions using a multi-slice approach." *Journal of the Optical Society of America. A, Optics, image science, and vision*, vol. 29, no. 8, pp. 1606–14, Aug. 2012.

[26] P. Li, D. J. Batey, T. B. Edo, and J. M. Rodenburg, "Separation of three-dimensional scattering effects in tilt-series Fourier ptychography," *Ultramicroscopy*, vol. 158, pp. 1–7, 2015.

[27] L. Tian and L. Waller, "3d intensity and phase imaging from light field measurements in an led array microscope," *Optica*, vol. 2, no. 2, pp. 104–111, 2015.

[28] M. D. Seaberg, B. Zhang, D. F. Gardner, E. R. Shanblatt, M. M. Murnane, H. C. Kapteyn, and D. E. Adams, "Tabletop nanometer extreme ultraviolet imaging in an extended reflection mode using coherent fresnel ptychography," *Optica*, vol. 1, no. 1, pp. 39–44, 2014.

[29] B. Zhang, D. F. Gardner, M. D. Seaberg, E. R. Shanblatt, H. C. Kapteyn, M. M. Murnane, and D. E. Adams, "High contrast 3d imaging of surfaces near the wavelength limit using tabletop EUV ptychography," *Ultramicroscopy*, vol. 158, pp. 98–104, 2015.

[30] A. Greenbaum, W. Luo, T.-W. Su, Z. Göröcs, L. Xue, S. O. Isikman, A. F. Coskun, O. Mudanyali, and A. Ozcan, "Imaging without lenses: achievements and remaining challenges of wide-field on-chip microscopy," *Nature methods*, vol. 9, no. 9, pp. 889–895, 2012.

[31] S. Dong, K. Guo, P. Nanda, R. Shiradkar, and G. Zheng, "Fpscope: a field-portable high-resolution microscope using a cellphone lens," *Biomedical optics Express*, vol. 5, no. 10, pp. 3305–3310, 2014.

[32] M. Berglund, L. Rymell, M. Peuker, T. Wilhein, and H. M. Hertz, "Compact water-window transmission X-ray microscopy," *Journal of microscopy*, vol. 197, no. 3, pp. 268–273, 2000.

[33] M. Benk, K. Bergmann, D. Schäfer, and T. Wilhein, "Compact soft x-ray microscope using a gas-discharge light source," *Optics Letters*, vol. 33, no. 20, pp. 2359–2361, 2008.

[34] P. Wachulak, A. Bartnik, H. Fiedorowicz, P. Rudawski, R. Jarocki, J. Kostecki, and M. Szczurek, "Water window compact, table-top laser plasma soft X-ray sources based on a gas puff target," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 268, no. 10, pp. 1692–1700, 2010.

[35] P. W. Wachulak, A. Bartnik, L. Wegrzynski, J. Kostecki, R. Jarocki, T. Fok, M. Szczurek, and H. Fiedorowicz, "Sub 1-μm resolution water-window microscopy using a compact, laser-plasma sxr source based on a double stream

gas-puff target," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 311, pp. 42–46, 2013.

[36] J. W. Goodman, *Introduction to Fourier optics*. Roberts and Company Publishers, 2005.

[37] D. Paganin, *Coherent X-ray optics*. Oxford University Press, 2006.

[38] M. Stockmar, P. Cloetens, I. Zanette, B. Enders, M. Dierolf, F. Pfeiffer, and P. Thibault, "Near-field ptychography: phase retrieval for inline holography using a structured illumination," *Scientific reports*, vol. 3, 2013.

[39] M. Zürch, J. Rothhardt, S. Hadrich, S. Demmler, M. Krebs, J. Limpert, A. Tünnermann, A. Guggenmos, U. Kleineberg, and C. Spielmann, "Real-time and sub-wavelength ultrafast coherent diffraction imaging in the extreme ultraviolet," *Scientific reports*, vol. 4, 2014.

[40] A. Dutt and V. Rokhlin, "Fast Fourier transforms for nonequispaced data," *SIAM Journal on Scientific computing*, vol. 14, no. 6, pp. 1368–1393, 1993.

[41] D. Attwood, *Soft x-rays and extreme ultraviolet radiation: principles and applications*. Cambridge university press, 1999.

[42] D. F. Gardner, B. Zhang, M. D. Seaberg, L. S. Martin, D. E. Adams, F. Salmassi, E. Gullikson, H. Kapteyn, and M. Murnane, "High numerical aperture reflection mode coherent diffraction microscopy using off-axis apertured illumination." *Optics Express*, vol. 20, no. 17, pp. 19 050–9, Aug. 2012.

[43] T. Sun, Z. Jiang, J. Strzalka, L. Ocola, and J. Wang, "Three-dimensional coherent X-ray surface scattering imaging near total external reflection," *Nature Photonics*, vol. 6, no. 9, pp. 586–590, 2012.

[44] M. Zürch, C. Kern, and C. Spielmann, "Xuv coherent diffraction imaging in reflection geometry with low numerical aperture," *Optics Express*, vol. 21, no. 18, pp. 21 131–21 147, 2013.

[45] A. M. Maiden, M. J. Humphry, F. Zhang, and J. M. Rodenburg, "Superresolution imaging via ptychography," *JOSA A*, vol. 28, no. 4, pp. 604–612, 2011.

[46] P. Thibault and M. Guizar-Sicairos, "Maximum-likelihood refinement for coherent diffractive imaging," *New Journal of Physics*, vol. 14, no. 6, p. 063004, 2012.

[47] R. Hegerl and W. Hoppe, "Influence of electron noise on three-dimensional image reconstruction," *Zeitschrift für Naturforschung A*, vol. 31, no. 12, pp. 1717–1721, 1976.

[48] J. C. H. Spence, U. Weierstall, and M. Howells, "Coherence and sampling requirements for diffractive imaging." *Ultramicroscopy*, vol. 101, no. 2-4, pp. 149–52, Nov. 2004.

[49] J. N. Clark and A. G. Peele, "Simultaneous sample and spatial coherence characterisation using diffractive imaging," *Applied Physics Letters*, vol. 99, no. 15, p. 154103, 2011.

[50] 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. 171104, Apr. 2014.

[51] B. Abbey, L. W. Whitehead, H. M. Quiney, D. J. Vine, G. A. Cadenazzi, C. A. Henderson, K. A. Nugent, E. Balaur, C. T. Putkunz, A. G. Peele, G. J. Williams, and I. McNulty, "Lensless imaging using broadband X-ray sources," *Nature Photonics*, vol. 5, no. 7, pp. 420–424, 2011.

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

[53] K. Giewekemeyer, H. Neubauer, S. Kalbfleisch, S. P. Krüger, and T. Salditt, "Holographic and diffractive X-ray imaging using waveguides as quasi-point sources," *New Journal of Physics*, vol. 12, no. 3, p. 035008, 2010.

[54] J. N. Clark, C. T. Putkunz, E. K. Curwood, D. J. Vine, R. Scholten, I. McNulty, K. A. Nugent, and A. G. Peele, "Dynamic sample imaging in coherent diffractive imaging," *Optics Letters*, vol. 36, no. 11, pp. 1954–1956, 2011.

[55] P. Thibault, V. Elser, C. Jacobsen, D. Shapiro, and D. Sayre, "Reconstruction of a yeast cell from X-ray diffraction data," *Acta Crystallographica Section A: Foundations of Crystallography*, vol. 62, no. 4, pp. 248–261, 2006.

[56] M. D. Seaberg, D. E. Adams, E. L. Townsend, D. a. Raymondson, W. F. Schlotter, Y. Liu, C. S. Menoni, L. Rong, C.-C. Chen, J. Miao, H. C. Kapteyn, and M. M. Murnane, "Ultrahigh 22 nm resolution coherent diffractive imaging using a desktop 13 nm high harmonic source." *Optics Express*, vol. 19, no. 23, pp. 22 470–9, Nov. 2011.

[57] P. W. Wachulak, A. Bartnik, and H. Fiedorowicz, "Sub-70 nm resolution tabletop microscopy at 13.8 nm using a compact laser–plasma EUV source," *Optics Letters*, vol. 35, no. 14, pp. 2337–2339, 2010.

[58] M. Dierolf, P. Thibault, A. Menzel, C. M. Kewish, K. Jefimovs, I. Schlichting, K. Von Koenig, O. Bunk, and F. Pfeiffer, "Ptychographic coherent diffractive imaging of weakly scattering specimens," *New Journal of Physics*, vol. 12, no. 3, p. 035017, 2010.

[59] M. Van Heel, "Similarity measures between images," *Ultramicroscopy*, vol. 21, no. 1, pp. 95–100, 1987.

[60] M. Van Heel and M. Schatz, "Fourier shell correlation threshold criteria," *Journal of structural biology*, vol. 151, no. 3, pp. 250–262, 2005.

[61] J. Cederquist, J. Fienup, J. Marron, and R. Paxman, "Phase retrieval from experimental far-field speckle data," *Optics Letters*, vol. 13, no. 8, pp. 619–621, 1988.

[62] G. Williams, H. Quiney, B. Dhal, C. Tran, K. A. Nugent, A. Peele, D. Paterson, M. De Jonge *et al.*, "Fresnel coherent diffractive imaging," *Physical Review Letters*, vol. 97, no. 2, p. 025506, 2006.

[63] B. Abbey, K. A. Nugent, G. J. Williams, J. N. Clark, A. G. Peele, M. A. Pfeifer, M. De Jonge, and I. McNulty, "Keyhole coherent diffractive imaging," *Nature Physics*, vol. 4, no. 5, pp. 394–398, 2008.

[64] J. Rodenburg, "Ptychography and related diffractive imaging methods," *Advances in Imaging and Electron Physics*, vol. 150, pp. 87–184, 2008.

[65] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Applied Optics*, vol. 21, no. 15, pp. 2758–2769, 1982.

[66] J. Fienup, "Reconstruction of an object from the modulus of its fourier transform," *Optics letters*, vol. 3, no. 1, pp. 27–29, 1978.

[67] U. Schnars and W. Jueptner, *Digital holography*. Springer, 2005.

[68] S. Marchesini, H. He, H. N. Chapman, S. P. Hau-Riege, A. Noy, M. R. Howells, U. Weierstall, and J. C. Spence, "X-ray image reconstruction from a diffraction pattern alone," *Physical Review B*, vol. 68, no. 14, p. 140101, 2003.

[69] D. R. Luke, "Relaxed averaged alternating reflections for diffraction imaging," *Inverse Problems*, vol. 21, no. 1, p. 37, 2005.

[70] O. Raz, B. Leshem, J. Miao, B. Nadler, D. Oron, and N. Dudovich, "Direct phase retrieval in double blind Fourier holography," *Optics Express*, vol. 22, no. 21, pp. 24 935–24 950, 2014.

[71] B. Leshem, R. Xu, Y. Dallal, J. Miao, B. Nadler, D. Oron, N. Dudovich, and O. Raz, "Direct single-shot phase retrieval from the diffraction pattern of separated objects," *Nature Communications*, vol. 7, 2016.

[72] J. N. Clark, X. Huang, R. Harder, and I. K. Robinson, "High-resolution three-dimensional partially coherent diffraction imaging." *Nature Communications*, vol. 3, p. 993, Jan. 2012.

[73] W. Hoppe, "Beugung im inhomogenen primärstrahlwellenfeld. i. prinzip einer phasenmessung von elektronenbeungungsinterferenzen," *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, vol. 25, no. 4, pp. 495–501, 1969.

[74] B. McCallum and J. Rodenburg, "Two-dimensional demonstration of wigner phase-retrieval microscopy in the stem configuration," *Ultramicroscopy*, vol. 45, no. 3-4, pp. 371–380, 1992.

[75] H. N. Chapman, "Phase-retrieval x-ray microscopy by wigner-distribution deconvolution," *Ultramicroscopy*, vol. 66, no. 3-4, pp. 153–172, 1996.

[76] J. M. Rodenburg and H. M. Faulkner, "A phase retrieval algorithm for shifting illumination," *Applied Physics Letters*, vol. 85, no. 20, pp. 4795–4797, 2004.

[77] H. Faulkner and J. Rodenburg, "Movable aperture lensless transmission microscopy: a novel phase retrieval algorithm," *Physical Review Letters*, vol. 93, no. 2, p. 023903, 2004.

[78] J. Rodenburg, A. Hurst, and A. Cullis, "Transmission microscopy without lenses for objects of unlimited size," *Ultramicroscopy*, vol. 107, no. 2, pp. 227–231, 2007.

[79] J. Rodenburg, A. Hurst, A. Cullis, B. Dobson, F. Pfeiffer, O. Bunk, C. David, K. Jefimovs, and I. Johnson, "Hard-x-ray lensless imaging of extended objects," *Physical Review Letters*, vol. 98, no. 3, p. 034801, 2007.

[80] P. Thibault, M. Dierolf, A. Menzel, O. Bunk, C. David, and F. Pfeiffer, "High-resolution scanning x-ray diffraction microscopy," *Science*, vol. 321, no. 5887, pp. 379–382, 2008.

[81] A. M. Maiden and J. M. Rodenburg, "An improved ptychographical phase retrieval algorithm for diffractive imaging," *Ultramicroscopy*, vol. 109, no. 10, pp. 1256–1262, 2009.

[82] P. Thibault, M. Dierolf, O. Bunk, A. Menzel, and F. Pfeiffer, "Probe retrieval in ptychographic coherent diffractive imaging," *Ultramicroscopy*, vol. 109, no. 4, pp. 338–343, 2009.

[83] E. H. Tsai, A. Diaz, A. Menzel, and M. Guizar-Sicairos, "X-ray ptychography using a distant analyzer," *Optics Express*, vol. 24, no. 6, pp. 6441–6450, 2016.

[84] S. Marchesini, A. Schirotzek, C. Yang, H. Wu, and F. Maia, "Augmented projections for ptychographic imaging," *Inverse Problems*, vol. 29, no. 11, p. 115009, 2013.

[85] D. C. Youla and H. Webb, "Image restoration by the method of convex projections: Part 1theory," *IEEE transactions on medical imaging*, vol. 1, no. 2, pp. 81–94, 1982.

[86] 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. 115010, 2012.

[87] P. Thibault, "Algorithmic methods in diffraction microscopy," Ph.D. dissertation, Cornell University, 2007.

[88] R. Dilanian, G. J. Williams, L. W. Whitehead, D. J. Vine, G. Peele, E. Balaur, I. McNulty, H. M. Quiney, and K. Nugent, "Coherent diffractive imaging: a new statistically regularized amplitude constraint," *New Journal of Physics*, vol. 12, no. 9, p. 093042, 2010.

[89] 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.

[90] F. Zhang, I. Peterson, J. Vila-Comamala, A. Diaz, F. Berenguer, R. Bean, B. Chen, A. Menzel, I. K. Robinson, and J. M. Rodenburg, "Translation position determination in ptychographic coherent diffraction imaging," *Optics Express*, vol. 21, no. 11, pp. 13 592–13 606, 2013.

[91] M. Beckers, T. Senkbeil, T. Gorniak, K. Giewekemeyer, T. Salditt, and A. Rosenhahn, "Drift correction in ptychographic diffractive imaging," *Ultramicroscopy*, vol. 126, pp. 44–47, 2013.

[92] 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. 481–7, Apr. 2008.

[93] 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.

[94] A. M. Maiden and J. M. Rodenburg, "An improved ptychographical phase retrieval algorithm for diffractive imaging." *Ultramicroscopy*, vol. 109, no. 10, pp. 1256–62, Sep. 2009.

[95] 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.

[96] V. Elser, I. Rankenburg, and P. Thibault, "Searching with iterated maps," *Proceedings of the National Academy of Sciences*, vol. 104, no. 2, pp. 418–423, 2007.

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

[98] P. M. Pelz, M. Guizar-Sicairos, P. Thibault, I. Johnson, M. Holler, and A. Menzel, "On-the-fly scans for X-ray ptychography," *Applied Physics Letters*, vol. 105, no. 25, p. 251101, 2014.

[99] X. Huang, K. Lauer, J. N. Clark, W. Xu, E. Nazaretski, R. Harder, I. K. Robinson, and Y. S. Chu, "Fly-scan ptychography," *Scientific reports*, vol. 5, 2015.

[100] J. Deng, Y. S. Nashed, S. Chen, N. W. Phillips, T. Peterka, R. Ross, S. Vogt, C. Jacobsen, and D. J. Vine, "Continuous motion scan ptychography: characterization for increased speed in coherent x-ray imaging," *Optics Express*, vol. 23, no. 5, pp. 5438–5451, 2015.

[101] S. Marchesini, H. Krishnan, D. A. Shapiro, T. Perciano, J. A. Sethian, B. J. Daurer, and F. R. Maia, "SHARP: a distributed, GPU-based ptychographic solver," *arXiv preprint arXiv:1602.01448*, 2016.

[102] C. Yang, J. Qian, A. Schirotzek, F. Maia, and S. Marchesini, "Iterative algorithms for ptychographic phase retrieval," *arXiv preprint arXiv:1105.5628*, 2011.

[103] J. Qian, C. Yang, A. Schirotzek, F. Maia, and S. Marchesini, "Efficient algorithms for ptychographic phase retrieval," *Inverse Problems and Applications, Contemp. Math*, vol. 615, pp. 261–280, 2014.

[104] S. Marchesini, Y.-C. Tu, and H. Wu, "Alternating projection, ptychographic imaging and phase synchronization," *Applied and Computational Harmonic Analysis*, 2015.

[105] E. Polak and G. Ribiere, "Note sur la convergence de méthodes de directions conjuguées," *Revue française d'informatique et de recherche opérationnelle, série rouge*, vol. 3, no. 1, pp. 35–43, 1969.

[106] W. H. Press, *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007.

[107] L.-H. Yeh, J. Dong, J. Zhong, L. Tian, M. Chen, G. Tang, M. Soltanolkotabi, and L. Waller, "Experimental robustness of Fourier ptychography phase retrieval algorithms," *Optics Express*, vol. 23, no. 26, pp. 33 214–33 240, 2015.

[108] G. Williams, H. Quiney, A. Peele, and K. Nugent, "Coherent diffractive imaging and partial coherence," *Physical Review B*, vol. 75, no. 10, p. 104102, Mar. 2007.

[109] L. W. Whitehead, G. J. Williams, H. M. Quiney, D. J. Vine, R. a. Dilanian, S. Flewett, K. a. Nugent, and I. McNulty, "Diffractive Imaging Using Partially Coherent X Rays," *Physical Review Letters*, vol. 103, no. 24, p. 243902, Dec. 2009.

[110] A. Tikhonov, "Solution of incorrectly formulated problems and the regularization method," *Soviet Math. Dokl.*, vol. 5, pp. 1035–1038, 1963.

[111] S. T. Thurman and J. R. Fienup, "Phase retrieval with signal bias," *JOSA A*, vol. 26, no. 4, pp. 1008–1014, 2009.

[112] G. H. Golub, P. C. Hansen, and D. P. O'Leary, "Tikhonov regularization and total least squares," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 1, pp. 185–194, 1999.

[113] P. C. Hansen and D. P. O'Leary, "The use of the l-curve in the regularization of discrete ill-posed problems," *SIAM Journal on Scientific Computing*, vol. 14, no. 6, pp. 1487–1503, 1993.

[114] A. Tripathi, I. McNulty, and O. G. Shpyrko, "Ptychographic overlap constraint errors and the limits of their numerical recovery using conjugate gradient descent methods," *Optics Express*, vol. 22, no. 2, pp. 1452–1466, 2014.

[115] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," *Optics Letters*, vol. 33, no. 2, pp. 156–158, 2008.

[116] J. Rodenburg and R. Bates, "The theory of super-resolution electron microscopy via wigner-distribution deconvolution," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 339, no. 1655, pp. 521–553, 1992.

[117] N. O'Brien, M. Mavrogordato, R. Boardman, I. Sinclair, S. Hawker, and T. Blumensath, "Comparing cone beam laminographic system trajectories for composite ndt," *Case Studies in Nondestructive Testing and Evaluation*, 2016.

[118] K. S. Raines, S. Salha, R. L. Sandberg, H. Jiang, J. A. Rodriguez, B. P. Fahimian, H. C. Kapteyn, J. Du, and J. Miao, "Three-dimensional structure

determination from a single view," *Nature*, vol. 463, no. 7278, pp. 214–217, 2010.

[119] N. Halko, P.-G. Martinsson, Y. Shkolnisky, and M. Tygert, "An algorithm for the principal component analysis of large data sets," *SIAM Journal on Scientific computing*, vol. 33, no. 5, pp. 2580–2594, 2011.

[120] N. Halko, P.-G. Martinsson, and J. A. Tropp, "Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions," *SIAM review*, vol. 53, no. 2, pp. 217–288, 2011.

[121] N. Hurley and S. Rickard, "Comparing measures of sparsity," *Information Theory, IEEE Transactions on*, vol. 55, no. 10, pp. 4723–4741, 2009.

[122] T. Godden, R. Suman, M. Humphry, J. Rodenburg, and A. Maiden, "Ptychographic microscope for three-dimensional imaging," *Optics Express*, vol. 22, no. 10, pp. 12 513–12 523, 2014.

[123] Z. Wang, L. Millet, M. Mir, H. Ding, S. Unarunotai, J. Rogers, M. U. Gillette, and G. Popescu, "Spatial light interference microscopy (SLIM)," *Optics Express*, vol. 19, no. 2, pp. 1016–1026, 2011.

[124] M. Guizar-Sicairos, M. Holler, A. Diaz, J. Vila-Comamala, O. Bunk, and A. Menzel, "Role of the illumination spatial-frequency spectrum for ptychography," *Physical review B*, vol. 86, no. 10, p. 100103, 2012.

[125] R. T. Bailey, G. Bourhill, F. R. Cruickshank, D. Pugh, J. N. Sherwood, and G. S. Simpson, "The linear and nonlinear optical properties of the organic nonlinear material 4-nitro-4-methylbenzylidene aniline," *Journal of Applied Physics*, vol. 73, no. 4, p. 1591, 1993.

[126] C. M. Bishop, "Pattern recognition," *Machine Learning*, vol. 128, pp. 1–58, 2006.

[127] E. Shaffer, C. Moratal, P. Magistretti, P. Marquet, and C. Depeursinge, "Label-free second-harmonic phase imaging of biological specimen by digital holographic microscopy," *Optics Letters*, vol. 35, no. 24, pp. 4102–4104, 2010.

[128] E. Shaffer, N. Pavillon, J. Kühn, and C. Depeursinge, "Digital holographic microscopy investigation of second harmonic generated at a glass/air interface," *Optics Letters*, vol. 34, no. 16, pp. 2450–2452, 2009.

[129] X. Chen, O. Nadiarynkh, S. Plotnikov, and P. J. Campagnola, "Second harmonic generation microscopy for quantitative analysis of collagen fibrillar structure," *Nature protocols*, vol. 7, no. 4, pp. 654–669, 2012.

[130] H. Yokota, J. Kaneshiro, and Y. Uesu, "Optical second harmonic generation microscopy as a tool of material diagnosis," *Physics Research International*, vol. 2012, 2012.

[131] K. Shi, P. S. Edwards, J. Hu, Q. Xu, Y. Wang, D. Psaltis, and Z. Liu, "Holographic coherent anti-stokes raman scattering bio-imaging," *Biomedical optics Express*, vol. 3, no. 7, pp. 1744–1749, 2012.

[132] G. Boyd, R. C. Miller, K. Nassau, W. Bond, and A. Savage, "Linbo3: an efficient phase matchable nonlinear optical material," *Applied Physics Letters*, vol. 5, no. 11, pp. 234–236, 1964.

[133] S. J. Luo, W. F. Du, and H. Z. Wang, "A first-principles study of linear and nonlinear optical properties of 4-nitro-4'-methylbenzylidene aniline." *The Journal of chemical physics*, vol. 129, no. 9, p. 094705, Sep. 2008.

[134] A. Yariv, *Introduction to optical electronics*. Holt, Rinehart and Winston, Inc., New York, NY, Jan 1976.

[135] A. Yariv and P. Yeh, *Optical waves in crystals*. Wiley, New York, 1984, vol. 10.

[136] D. E. Zelmon, D. L. Small, and D. Jundt, "Infrared corrected sellmeier coefficients for congruently grown lithium niobate and 5 mol.% magnesium oxide–doped lithium niobate," *JOSA B*, vol. 14, no. 12, pp. 3319–3322, 1997.

[137] D. W. Berreman, "Optics in stratified and anisotropic media: 4× 4-matrix formulation," *Josa*, vol. 62, no. 4, pp. 502–510, 1972.

[138] P. Franken, A. Hill, C. Peters, and G. Weinreich, "Generation of optical harmonics," *Physical Review Letters*, vol. 7, no. 4, pp. 118–119, 1961.

[139] G. Boyd and D. Kleinman, "Parametric interaction of focused Gaussian light beams," *Journal of Applied Physics*, vol. 39, no. 8, pp. 3597–3639, 1968.

[140] D. F. Gardner, B. Zhang, M. D. Seaberg, L. S. Martin, D. E. Adams, F. Salmassi, E. Gullikson, H. Kapteyn, and M. Murnane, "High numerical aperture reflection mode coherent diffraction microscopy using off-axis

apertured illumination," *Optics Express*, vol. 20, no. 17, pp. 19 050–19 059, 2012.

[141] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio science*, vol. 23, no. 4, pp. 713–720, 1988.

[142] G. Miller, R. Batchko, W. Tulloch, D. Weise, M. Fejer, and R. Byer, "42%-efficient single-pass cw second-harmonic generation in periodically poled lithium niobate," *Optics Letters*, vol. 22, no. 24, pp. 1834–1836, 1997.

[143] O. Gayer, Z. Sacks, E. Galun, and A. Arie, "Temperature and wavelength dependent refractive index equations for MgO-doped congruent and stoichiometric LiNbO3," *Applied Physics B*, vol. 91, no. 2, pp. 343–348, 2008.

[144] K. A. Nugent, "Coherent methods in the X-ray sciences," *Advances in Physics*, vol. 59, no. 1, pp. 1–99, 2010.

[145] J. Miao, R. L. Sandberg, and C. Song, "Coherent X-ray diffraction imaging," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 18, no. 1, pp. 399–410, 2012.

[146] U. Neuhäusler, G. Schneider, W. Ludwig, M. Meyer, E. Zschech, and D. Hambach, "X-ray microscopy in zernike phase contrast mode at 4 keV photon energy with 60 nm resolution," *Journal of Physics D: Applied Physics*, vol. 36, no. 10A, p. A79, 2003.

[147] G. A. van Riessen, M. Junker, N. W. Phillips, and A. G. Peele, "A soft X-ray beamline for quantitative nanotomography using ptychography," pp. 885 117–885 117, 2013.

[148] N. Burdet, G. R. Morrison, X. Huang, X. Shi, J. N. Clark, F. Zhang, M. Civita, R. Harder, and I. K. Robinson, "Observations of artefacts in the x-ray ptychography method," *Optics Express*, vol. 22, no. 9, pp. 10 294–10 303, 2014.

[149] P. Hessing, B. Pfau, E. Guehrs, M. Schneider, L. Shemilt, J. Geilhufe, and S. Eisebitt, "Holography-guided ptychography with soft X-rays," *Optics Express*, vol. 24, no. 2, pp. 1840–1851, 2016.

[150] L. Juschkin, L. Loetgering, D. Rudolf, R. Xu, S. Brose, S. Danylyuk, and J. Miao, "Tabletop coherent diffraction imaging with a discharge plasma EUV source," pp. 88 490Y–88 490Y, 2013.

[151] D. F. Gardner, B. Zhang, M. D. Seaberg, E. R. Shanblatt, C. L. Porter, R. Karl, C. Mancuso, H. C. Kapteyn, M. M. Murnane, and D. E. Adams, "Lensless hyperspectral spectromicroscopy with a tabletop extreme-ultraviolet source," in *SPIE Advanced Lithography*. International Society for Optics and Photonics, 2016, pp. 977 808–977 808.

[152] D. F. Gardner, C. L. Porter, E. R. Shanblatt, G. F. Mancini, R. Karl, M. Tanksalvala, C. Bevis, H. C. Kapteyn, M. M. Murnane, and D. E. Adams, "Spectroscopic imaging of buried layers in 2+ 1D via tabletop ptychography with high-harmonic EUV illumination," pp. 97 780J–97 780J, 2016.

[153] M. Ammosov, N. B. Delone, V. P. Krainov *et al.*, "Tunnel ionization of complex atoms and of atomic ions in an alternating electromagnetic field," *Sov. Phys. JETP*, vol. 64, no. 6, pp. 1191–1194, 1986.

[154] P. B. Corkum, "Plasma perspective on strong field multiphoton ionization," *Physical Review Letters*, vol. 71, no. 13, p. 1994, 1993.

[155] T. Brabec and F. Krausz, "Intense few-cycle laser fields: Frontiers of nonlinear optics," *Reviews of Modern Physics*, vol. 72, no. 2, p. 545, 2000.

[156] M. Lewenstein, P. Balcou, M. Y. Ivanov, A. Lhuillier, and P. B. Corkum, "Theory of high-harmonic generation by low-frequency laser fields," *Physical Review A*, vol. 49, no. 3, p. 2117, 1994.

[157] A. Shiner, C. Trallero-Herrero, N. Kajumba, H.-C. Bandulet, D. Comtois, F. Légaré, M. Giguère, J. Kieffer, P. Corkum, and D. Villeneuve, "Wavelength scaling of high harmonic generation efficiency," *Physical Review Letters*, vol. 103, no. 7, p. 073902, 2009.

[158] M. Schnürer, Z. Cheng, M. Hentschel, G. Tempea, P. Kálmán, T. Brabec, and F. Krausz, "Absorption-limited generation of coherent ultrashort soft-x-ray pulses," *Physical review letters*, vol. 83, no. 4, p. 722, 1999.

[159] P. Balcou, P. Salieres, A. L'Huillier, and M. Lewenstein, "Generalized phase-matching conditions for high harmonics: The role of field-gradient forces," *Physical Review A*, vol. 55, no. 4, p. 3204, 1997.

[160] F. Lindner, G. G. Paulus, H. Walther, A. Baltuška, E. Goulielmakis, M. Lezius, and F. Krausz, "Gouy phase shift for few-cycle laser pulses," *Physical review letters*, vol. 92, no. 11, p. 113001, 2004.

[161] E. A. Gibson, A. Paul, N. Wagner, D. Gaudiosi, S. Backus, I. P. Christov, A. Aquila, E. M. Gullikson, D. T. Attwood, M. M. Murnane *et al.*, "Coherent soft x-ray generation in the water window with quasi-phase matching," *Science*, vol. 302, no. 5642, pp. 95–98, 2003.

[162] M.-C. Chen, P. Arpin, T. Popmintchev, M. Gerrity, B. Zhang, M. Seaberg, D. Popmintchev, M. M. Murnane, and H. C. Kapteyn, "Bright, Coherent, Ultrafast Soft X-Ray Harmonics Spanning the Water Window from a Tabletop Light Source," *Physical Review Letters*, vol. 105, no. 17, p. 173901, Oct. 2010.

[163] R. T. Chapman, T. J. Butcher, P. Horak, F. Poletti, J. G. Frey, and W. S. Brocklesby, "Modal effects on pump-pulse propagation in an Ar-filled capillary," *Optics Express*, vol. 18, no. 12, pp. 13 279–13 284, 2010.

[164] O. Kfir, P. Grychtol, E. Turgut, R. Knut, D. Zusin, D. Popmintchev, T. Popmintchev, H. Nembach, J. M. Shaw, A. Fleischer *et al.*, "Generation of bright phase-matched circularly-polarized extreme ultraviolet high harmonics," *Nature Photonics*, 2014.

[165] A. Rundquist, C. G. Durfee, Z. Chang, C. Herne, S. Backus, M. M. Murnane, and H. C. Kapteyn, "Phase-matched generation of coherent soft x-rays," *Science*, vol. 280, no. 5368, pp. 1412–1415, 1998.

[166] C. G. Durfee III, A. R. Rundquist, S. Backus, C. Herne, M. M. Murnane, and H. C. Kapteyn, "Phase matching of high-order harmonics in hollow waveguides," *Physical Review Letters*, vol. 83, no. 11, p. 2187, 1999.

[167] J. Wesson and D. J. Campbell, *Tokamaks*. Oxford University Press, 2011, vol. 149.

[168] T. Popmintchev, M.-C. Chen, A. Bahabad, M. Gerrity, P. Sidorenko, O. Cohen, I. P. Christov, M. M. Murnane, and H. C. Kapteyn, "Phase matching of high harmonic generation in the soft and hard X-ray regions of the spectrum," *Proceedings of the National Academy of Sciences*, vol. 106, no. 26, pp. 10 516–10 521, 2009.

[169] E. Constant, D. Garzella, P. Breger, E. Mével, C. Dorrer, C. Le Blanc, F. Salin, and P. Agostini, "Optimizing high harmonic generation in absorbing gases: Model and experiment," *Physical Review Letters*, vol. 82, no. 8, p. 1668, 1999.

[170] B. Mills, "Focussing and diffraction using a high harmonic source," Ph.D. dissertation, University of Southampton, 2009.

[171] A. D. Rakic, A. B. Djurišic, J. M. Elazar, and M. L. Majewski, "Optical properties of metallic films for vertical-cavity optoelectronic devices," *Applied Optics*, vol. 37, no. 22, pp. 5271–5283, 1998.

[172] R. Bartels, A. Paul, M. Murnane, H. Kapteyn, S. Backus, Y. Liu, and D. Attwood, "Absolute determination of the wavelength and spectrum of an extreme-ultraviolet beam by a Youngs double-slit measurement," *Optics Letters*, vol. 27, no. 9, pp. 707–709, 2002.

[173] D. Rudolf, J. Bußmann, M. Odstrčil, M. Dong, K. Bergmann, S. Danylyuk, and L. Juschkin, "Interferometric broadband Fourier spectroscopy with a partially coherent gas-discharge extreme ultraviolet light source," *Optics Letters*, vol. 40, no. 12, pp. 2818–2821, 2015.

[174] R. A. Dilanian, B. Chen, S. Teichmann, L. V. Dao, H. M. Quiney, and K. A. Nugent, "High-harmonic-generation spectrum reconstruction from Young's double-slits interference pattern using the maximum entropy method," *Optics Letters*, vol. 33, no. 20, pp. 2341–2343, 2008.

[175] S. Teichmann, B. Chen, R. A. Dilanian, P. Hannaford, and L. V. Dao, "Spectral characteristics across the spatial profile of a high-harmonic beam," *Journal of Physics D: Applied Physics*, vol. 42, p. 5108, 2009.

[176] S. Zayko, E. Mönnich, M. Sivis, D.-D. Mai, T. Salditt, S. Schäfer, and C. Ropers, "Coherent diffractive imaging beyond the projection approximation: waveguiding at extreme ultraviolet wavelengths," *Optics Express*, vol. 23, no. 15, pp. 19 911–19 921, 2015.

[177] H. V. Le, K. B. Dinh, P. Hannaford, and L. Van Dao, "High resolution coherent diffractive imaging with a table-top extreme ultraviolet source," *Journal of Applied Physics*, vol. 116, no. 17, p. 173104, 2014.

[178] G. Tadesse, R. Klas, S. Demmler, S. Hadrich, I. Wahyutama, M. Steinert, C. Spielmann, M. Zürch, A. Tünnermann, J. Limpert *et al.*, "High speed and high resolution table-top nanoscale imaging," *arXiv preprint arXiv:1605.02909*, 2016.

[179] P. Baksh, "Ptychography imaging of real biological samples using a high harmonic and synchrotron source," Ph.D. dissertation, University of Southampton, 2016.

[180] Y. Nishino, Y. Takahashi, N. Imamoto, T. Ishikawa, and K. Maeshima, "Three-dimensional visualization of a human chromosome using coherent x-ray diffraction," *Physical Review Letters*, vol. 102, no. 1, p. 018101, 2009.

[181] L. Shemilt, E. Verbanis, J. Schwenke, A. K. Estandarte, G. Xiong, R. Harder, N. Parmar, M. Yusuf, F. Zhang, and I. K. Robinson, "Karyotyping human chromosomes by optical and x-ray ptychography methods," *Biophysical Journal*, vol. 108, no. 3, pp. 706–713, 2015.

[182] D. Adjei, A. Wiechec, P. Wachulak, M. G. Ayele, J. Lekki, W. M. Kwiatek, A. Bartnik, M. Davidkovi, L. Vysin, L. Juha *et al.*, "DNA strand breaks induced by soft X-ray pulses from a compact laser plasma source," *Radiation Physics and Chemistry*, vol. 120, pp. 17–25, 2016.

[183] D. Claus and J. M. Rodenburg, "Pixel size adjustment in coherent diffractive imaging within the rayleigh–sommerfeld regime," *Applied Optics*, vol. 54, no. 8, pp. 1936–1944, 2015.

[184] R. M. Clare, M. Stockmar, M. Dierolf, I. Zanette, and F. Pfeiffer, "Characterization of near-field ptychography," *Optics Express*, vol. 23, no. 15, pp. 19 728–19 742, 2015.

[185] 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. 054008, 2016.

[186] E. J. Takahashi, T. Kanai, K. L. Ishikawa, Y. Nabekawa, and K. Midorikawa, "Coherent water window x ray by phase-matched high-order harmonic generation in neutral media," *Physical Review Letters*, vol. 101, no. 25, p. 253901, 2008.

[187] D. F. Gardner, G. F. Mancini, M. Tanksalvala, E. Shanblatt, X. Zhang, B. Galloway, C. Porter, R. Karl, C. Bevis, M. Murnane *et al.*, "Ptychographic imaging with 17.5 nm spatial resolution employing high harmonic light at 13.5 nm," in *CLEO: Applications and Technology*. Optical Society of America, 2016, pp. ATu4J–5.

[188] Y. Teramoto, Z. Narihiro, D. Yamatani, T. Yokoyama, K. Bessho, Y. Joshima, T. Shirai, S. Mouri, T. Inoue, H. Mizokoshi *et al.*, "Development of Sn-fueled high-power DPP EUV source for enabling HVM," in *Advanced Lithography*. International Society for Optics and Photonics, 2007, pp. 65 173R–65 173R.

[189] F. Brizuela, G. Vaschenko, C. Brewer, M. Grisham, C. Menoni, M. Marconi, J. Rocca, W. Chao, J. Liddle, E. Anderson *et al.*, "Reflection mode imaging with nanoscale resolution using a compact extreme ultraviolet laser," *Optics Express*, vol. 13, no. 11, pp. 3983–3988, 2005.

[190] G. Vaschenko, C. Brewer, F. Brizuela, Y. Wang, M. Larotonda, B. Luther, M. Marconi, J. Rocca, C. Menoni, E. Anderson *et al.*, "Sub-38 nm resolution tabletop microscopy with 13 nm wavelength laser light," *Optics Letters*, vol. 31, no. 9, pp. 1214–1216, 2006.

[191] K. W. Kim, Y. Kwon, K.-Y. Nam, J.-H. Lim, K.-G. Kim, K. S. Chon, B. H. Kim, D. E. Kim, J. Kim, B. N. Ahn *et al.*, "Compact soft x-ray transmission microscopy with sub-50 nm spatial resolution," *Physics in medicine and biology*, vol. 51, no. 6, p. N99, 2006.

[192] L. Juschkin, R. Freiberger, and K. Bergmann, "EUV microscopy for defect inspection by dark-field mapping and zone plate zooming," in *Journal of Physics: Conference Series*, vol. 186, no. 1. IOP Publishing, 2009, p. 012030.

[193] K. Bergmann, F. Küpper, and M. Benk, "Soft x-ray emission from a pulsed gas discharge in a pseudosparklike electrode geometry," *Journal of Applied Physics*, vol. 103, no. 12, p. 123304, 2008.

[194] M. Benk and K. Bergmann, "Brilliance scaling of discharge sources for extreme-ultraviolet and soft x-ray radiation for metrology applications," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 11, no. 2, pp. 021 106–1, 2012.

[195] K. B. Dinh, H. V. Le, P. Hannaford, and L. Van Dao, "Coherent diffractive imaging microscope with a high-order harmonic source," *Applied Optics*, vol. 54, no. 17, pp. 5303–5308, 2015.

[196] S. Witte, V. T. Tenner, D. W. Noom, and K. S. Eikema, "Lensless diffractive imaging with ultra-broadband table-top sources: from infrared to extreme-ultraviolet wavelengths," *Light: Science & Applications*, vol. 3, no. 3, p. e163, 2014.

[197] G. Jansen, D. Rudolf, L. Freisem, K. Eikema, and S. Witte, "Spatially resolved fourier transform spectroscopy in the extreme ultraviolet," *arXiv preprint arXiv:1607.02386*, 2016.

[198] A. Fleischer, O. Kfir, T. Diskin, P. Sidorenko, and O. Cohen, "Spin angular momentum and tunable polarization in high-harmonic generation," *Nature Photonics*, vol. 8, no. 7, pp. 543–549, 2014.

[199] T. Fan, P. Grychtol, R. Knut, C. Hernandez-Garcia, D. D. Hickstein, D. Zusin, C. Gentry, F. J. Dollar, C. Mancuso, C. Hogle *et al.*, "Bright soft X-ray high harmonic generation with circular polarization for X-ray magnetic circular dichroism," in *Compact EUV & X-ray Light Sources*. Optical Society of America, 2016, pp. ET5A–5.

[200] S. Zayko, M. Sivis, S. Schäfer, and C. Ropers, "Polarization contrast of nanoscale waveguides in high harmonic imaging," *Optica*, vol. 3, no. 3, pp. 239–242, 2016.

[201] S. Grafström, U. Harbarth, J. Kowalski, R. Neumann, and S. Noehte, "Fast laser beam position control with submicroradian precision," *Optics Communications*, vol. 65, no. 2, pp. 121–126, 1988.

[202] A. Cordes and A. Davidson, "CMOS cameras allow robust active stabilization of laser beams," *Laser Focus World*, vol. 47, no. 8, pp. 73–75, 2011.

[203] R. A. Hardin, Y. Liu, C. Long, A. Aleksandrov, and W. Blokland, "Active beam position stabilization of pulsed lasers for long-distance ion profile diagnostics at the Spallation Neutron Source (SNS)," *Optics Express*, vol. 19, no. 4, pp. 2874–2885, 2011.