Research Article | Vol. 29, No. 14 / 5 July 2021 / Optics Express 22345
Optics EXPRESS

Flexible ptychography platform to expand the potential of imaging at free electron lasers

KONSTANTIN KHARITONOV,[1] MASOUD MEHRJOO,[1,*] MABEL RUIZ-LOPEZ,[1] BARBARA KEITEL,[1] SVEA KREIS,[1] MARTIN SEYRICH,[1] MIHAI POP,[2] AND ELKE PLÖNJES[1]

[1]*Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, 22607 Hamburg, Germany*
[2]*MAX-IV, Lund University, 22100 Lund, Sweden*
[*]*masoud.mehrjoo@desy.de*

Abstract: Ptychography, a scanning coherent diffraction imaging method, can produce a high-resolution reconstruction of a sample and, at the same time, of the illuminating beam. The emergence of vacuum ultraviolet and X-ray free electron lasers (FELs) has brought sources with unprecedented characteristics that enable X-ray ptychography with highly intense and ultra-fast short-wavelength pulses. However, the shot-to-shot pulse fluctuations typical for FEL pulses and particularly the partial spatial coherence of self-amplified spontaneous emission (SASE) FELs lead to numerical complexities in the ptychographic algorithms and ultimately restrict the application of ptychography at FELs. We present a general adaptive forward model for ptychography based on automatic differentiation, which is able to perform reconstructions even under these conditions. We applied this model to the first ptychography experiment at FLASH, the Free electron LASer in Hamburg, and obtained a high-resolution reconstruction of the sample as well as the complex wavefronts of individual FLASH pulses together with their coherence properties. This is not possible with more common ptychography algorithms.

1. Introduction

Ptychography is a scanning coherent diffraction imaging (CDI) technique that allows the simultaneous imaging of a sample and an unknown illuminating beam (hereafter called probe) [1,2]. It is capable of simultaneously reconstructing probe and sample functions and can achieve diffraction limited resolution. Generally, scanning of a sample with overlap between adjacent scan positions builds the essence of a ptychography experiment by providing the required redundancy in the measured data [3]. This redundancy prevents the ambiguities common to conventional CDI methods and leads to robust and stable reconstructions. Over the past decade, significant advances in ptychography have been achieved, such as use of partially coherent light [3], correction of errors in scan positions [4–6] and slicing images of extended objects [7].

Despite of all of these advances, a shot-to-shot spatially stable and ideally fully spatially coherent probe is still a prerequisite for successful ptychographic reconstruction. Both conditions can be realized for synchrotron radiation at the cost of significant beam aperturing [8]. Successful application of ptychography at synchrotron facilities has drawn the attention towards implementation of ptychography at free electron lasers (FELs). These sources possess higher levels of spatial coherence and photon flux which, together with their ultra-short pulses, makes them ideal facilities for imaging [9]. On the other hand, FELs themselves benefit from ptychography as a beam characterization tool due to its ability to be applied to high-resolution reconstruction of unknown probe beams.

However, FLASH and other SASE FELs are chaotic in nature and generate pulses with varying shot-to-shot parameters [10–12]. The stochastic nature of the SASE process, where initially

#426931 https://doi.org/10.1364/OE.426931
Journal © 2021 Received 6 Apr 2021; revised 14 Jun 2021; accepted 20 Jun 2021; published 30 Jun 2021

Optics EXPRESS

spontaneous undulator radiation is amplified and gives rise to the narrow-band FEL light pulses, leads to shot-to-shot variations of the FEL radiation [13–15].

Despite that, several ptychographic experiments were successfully carried out at FELs. Schropp et al. [16] has utilized compound refractive lenses (CRLs) for focusing the beam at the hard X-ray FEL LCLS. The small input aperture of the focusing optics, which however leads to clipping of the beam, can overcome the difficulties associated with partially coherent and spatially fluctuating illumination. The reconstruction algorithm used in this work treats the reconstructed sample and probe assuming an average and shot-to-shot constant illumination common for all positions. In a second step, the algorithm refines the probe for each individual position without varying the sample any further. This approach was able to produce meaningful high-resolution reconstructions of the probe and the sample while ignoring the only partial spatial coherence and non-perfect pointing stability of the FEL beam. However, this approximation cannot be applied without significant aperturing of the soft X-ray FEL beam, since both assumptions would otherwise be incorrect [17–19].

Another approach was demonstrated by Sala et al. [20]. Here, the problem of spatial fluctuations was solved by applying an orthogonal probe relaxation (OPR) method [21]. OPR assumes that for only slightly different probes a single-mode reconstruction is possible by linking the probes through a lower dimension representation and thus regularizing the problem. This representation can be found by applying a truncated singular value decomposition (TSVD) to the matrices of all the probes reconstructed at the previous iteration of the reconstruction loop. Consequently, individual probes for each scan position may be seen as a weighted coherent sum of these basis vectors. In this representation, the assumption of full spatial coherence is still necessary, which makes this method inapplicable for characterization of the coherence properties of FEL radiation [17–19,22–25]. Recently, Daurer et al. [26] has applied mixed-state ptychography to characterize wavefronts of FELs for single-particle imaging. Within that work, the effect of partial coherence and pulse instabilities were studied.

At present, most phase retrieval ptychography algorithms rely on either iterative projections [27] or gradient [28] methods such as the ptychographic iterative engine, PIE. The latter permit more flexibility and allow incorporation of experimental conditions as a prerequisite in order to improve the reconstruction quality, to obtain modal expression of the probe, or to refine the scan coordinates. However, the need for an explicit closed-form expression of the error gradients makes any modifications to the forward model undesirable since the gradients need to be re-derived for every slight change in the experimental model. An elegant way to solve this problem is to use various numerical methods that allow automatic differentiation (AD) [29,30] of the error metric with respect to optimizable parameters. AD is able to calculate numerically gradients of a function with respect to its parameters.

Using AD as a promising tool of phase retrieval was first proposed by Jurling and Fienup [31]. The main advantage of this approach is that it divides the ptychographic reconstruction into three separate components: forward model, error metrics, and optimization routine. The forward model describes the experimental setup and approximates the measured intensity patterns based on optimizable input parameters. The error metrics evaluates the similarity between approximated and measured intensities. The optimization routine can be freely selected from a variety of the gradient-based optimization techniques to ensure the best reconstruction results in the particular experiment. Due to this separation, AD ensures additional flexibility in the ptychographic routine. Each of its components can be changed independently. This allows easy changes in the experimental setup and in the forward physical model. Also, additional optimizable parameters can be included and optimization and regularization techniques switched without the need for tedious re-deriving of gradients by hand.

Several recent publications [32–34] demonstrated successful use of AD as a basis for a ptychography framework. However, for a successful ptychographic reconstruction at FELs the

Optics EXPRESS

chaotic nature of the SASE process needs to be taken in account in the forward model. In fact, the partial spatial coherence of pulses as well as position instabilities should be incorporated into the model as additional optimizable parameters. None of the currently proposed AD-based approaches addresses these issues. Also, to our knowledge, nobody has tried to incorporate a scan position refinement into the AD-based ptychographic algorithm.

In this work we demonstrate an improved adaptive automatic differentiation (A-AD) based model for FEL ptychography. The built-in AD-based position correction routine helps to overcome the limited position stability and reproducibility of the sample scanning and limited pointing stability of the probe. The flexibility of the AD-based phase retrieval permits a quick adaptation of the forward propagation model to change the reconstruction algorithm. For instance, an imprecise knowledge of the sample-to-detector distance may lead to a non-ideal convergence. This problem can be refined by adding an angular spectrum propagation to the AD-based phase retrieval with an optimizable distance. This addition, through the AD approach, demands no further analytical recalculation of the entire reconstruction procedure. In addition, fluctuations of the probe features, such as pointing instabilities, partial coherence and probe variations, are encapsulated into the numerical differentiable forward model.

We have developed an improved adaptive automatic differentiation (A-AD) based model for FEL ptychography. As an experimental proof of concept, we have demonstrated ptychographic imaging at the FLASH2 variable micro-focus beamlime FL24 [12]. Our model has allowed us to perform high-resolution reconstruction of the sample as well as a detailed shot-to-shot FEL pulse characterization.

The paper is structured as follows. In Section 2 we introduce the adaptive automatic differentiation for ptychography and discuss its theoretical implementation. Section 3 describes the experimental realization of a ptychography experiment at the FALSH2 facility, at the beamline FL24, and presents the reconstruction results based on the AD and A-AD approaches. Finally, the results are summarized and discussed in Sections 4 and 5.

2. Methods

2.1. Ptychography as an optimization problem

A ptychographic model considers a number P probe scans across the sample O with the scanning positions described by a set of translation vectors, $\mathbf{r}_j = (r_{x_j}, r_{y_j})\, j \in 1..N$, where N is the number of scanning points. Within the paraxial and projection approximations [35], the exit wave, $\Psi_j = P \cdot O_{\mathbf{r}_j}$, propagates to the detector and the corresponding intensity is given as:

$$I_j = \left|\mathcal{L}\{\Psi_j\}\right|^2 + \epsilon, \tag{1}$$

where $\mathcal{L}$ indicates the propagators defined by the experimental geometry and the wavelength through the Fresnel number [35], and ϵ represents the additive noise.

As a heuristic approach, we consider a precise knowledge of the scan position and full spatial coherence of the probe. The consequences of deviating from this ideal case will be discussed for both aspects in the next sections. The optimization task behind basic ptychographic reconstruction can thus be formulated as follows: finding a set of P, O which minimizes a loss function $\mathcal{E}(I_j, \tilde{I}_j)$, where I_j is the measured and $\tilde{I}_j$ is the reconstructed intensity at the j-th scan position ($\mathbf{r_j}, j \in 1..N$). Since Poisson noise is not significant in the measurements, we use a loss function for a Gaussian noise approximation as a likelihood error metrics [36] given by the following:

$$\mathcal{E}(I, \tilde{I}) = \sum_{j=1}^{N} \left|\sqrt{I_j} - \sqrt{\tilde{I}_j}\right|^2. \tag{2}$$

A representation of the Eq. (2) for the low-count measurements, where the Poisson noise is predominant, has been derived in [36].

Optics EXPRESS

Classical ptychography algorithms implemented in [3,28,37,38] require deriving the gradients of the loss function with respect to all optimizable parameters. This is a time-consuming task which has to be repeated after each modification of the forward model of Eq. (1). Therefore, any changes of the forward model are undesirable and this hinders ptychographic reconstruction under conditions typical for SASE FELs. Utilizing AD for phase retrieval [31–34] allows deriving the gradients automatically and significantly mitigates the problem. It should be noted that by analytical gradient calculations combined with relaxation methods, the classical ptychography algorithms can aim at performing ptychography at FELs.

2.2. Brief description of automatic differentiation

To illustrate the idea of automatic differentiation (AD) and its use for phase retrieval reconstruction, consider the forward model description of the phase retrieval experiment as follows:

$$F(p_0 \cdots p_n) = f(\phi_k(\cdots \phi_1(\phi_0(p_0 \cdots p_n, I)))) : \ \mathbb{R}^n : \mathbb{R}, \tag{3}$$

where $f\,(p_0 \cdots p_n)$ denotes the loss function $\mathbb{R}^n : \mathbb{R}$ and all sub-functions are fully differentiable: $\forall\phi\ \exists\frac{\partial\phi}{\partial x_n}$, where x_n is a set of input parameters for the intermediate function ϕ. Now, if we want to perform an optimization for a set of input parameters p_n of F with respect to f, we may use gradient-based optimization methods. This can be done if we are able to compute $\frac{\partial F}{\partial p}$, $\forall p \in p_n$ at the point $p_n = c_n$ where we are interested in the local gradient value. To do so, we may use the chain rule and the existing gradients of $\phi_{0..k}$ functions in the following manner:

$$\left.\frac{\partial F}{\partial p_n}\right|_{p_n=c_n} = \left.\frac{\partial f}{\partial p_n}\right|_{p_n=\phi_k(\cdots\phi_0(c_n))} \cdot \left.\frac{\partial \phi_k}{\partial p_n}\right|_{p_n=\phi_{k-1}(\cdots\phi_0(c_n))} \cdot \cdots \cdot \left.\frac{\partial \phi_1}{\partial p_n}\right|_{p_n=\phi_0(c_n)} \cdot \left.\frac{\partial \phi_0}{\partial p_n}\right|_{p_n=c_n}. \tag{4}$$

The routine tracks the forward computation $F(p_0 \cdots p_n)$ and stores the $\phi_0(c_n)\ \cdots\ \phi_k(\cdots\phi_0(c_n))$ values for their further use in the gradient computation. This allows gradient calculations of arbitrarily complex functions as long as they are composed of fully differentiable sub-functions [39]. To finish the optimization step, we need to pass the obtained gradient $\frac{\partial F}{\partial p_n}$ to the gradient-based optimizer which changes the optimizable parameters p_n accordingly.

AD builds a phase retrieval engine which can be subdivided into three parts: forward differentiable model, loss function and gradient based optimizer. Each part can be modified independently of the others. This flexibility allows the forward model to be adapted to any change in the experimental geometry. Previously, AD was proposed to perform ptychographic reconstructions in different geometries including near-, far-field and Bragg conditions [32,34]. However, many critical points for successful ptychography at FELs such as the ability to treat shot-to-shot variations of the probe and scan position refinement have not yet been considered.

In the following sections, we will elaborate on the scan position refinement and spatial coherence treatment within an A-AD based reconstruction. Furthermore, modifying the loss function using orthonormalization regularization and denoising will be discussed.

2.3. Adaptive automatic differentiation ptychography

2.3.1. Scan position refinement by adaptive automatic differentiation

Ptychographic reconstruction algorithms rely on a mutual overlap of neighbouring probes and thus are very sensitive to sample positioning errors. Ideally, the positions should be well-known with the same or better precision as the desired resolution of the reconstruction. Our implementation of a scan position refinement relies on a gradient-based optimization technique with respect to the loss function. Similar analytical approaches were proposed in [6,40,41] that aim at optimizing the ptychography models by including the position refinement. It has been shown that the nonlinear optimization of the scan positions robustly constrains the reconstruction problem and leads to

Optics EXPRESS

faster and reliable convergence. To our knowledge, our algorithm is the first fully AD-based ptychographic scan position refinement algorithm.

To facilitate an integration of scan position refinement into the AD-based reconstruction algorithm, we use an affine transformation for the parameterized grid generation and a differentiable grid sampler, following a method proposed by Jaderberg et al. [42]. In our model, each individual exit wave is produced by multiplication of the probe's complex wave field and the fixed central region of the shifted object. The shifted object is placed in reference coordinates which are fixed to the detector coordinates. The central part corresponds to the part of the sample that was illuminated as shown in Fig. 1(b) (orange central grid). Thus, the exit wave may be calculated as follows:

$$\Psi = P_{k,l} \cdot O^*_{i,j} \;:\; k,l \in 1 \cdots N,\; i \in \frac{H-N}{2} \cdots \frac{H+N}{2},\; j \in \frac{W-N}{2} \cdots \frac{W+N}{2}, \tag{5}$$

where lower indices k, l, i, j refer to the numerical grid coordinates in the object plane, $N \times N$ is the size of the probe function in pixels, O^* is the $H \times W$ pixel object in the reference coordinate system, and $O^*_{i,j}$ is the illuminated part of $N \times N$ pixels of the shifted object producing the measured intensities. In general, the optimizer tries to find the best set of θ_j parameters of the affine transformation $\mathcal{A}_\theta$ for each individual scan position to minimize the difference between measured and reconstructed intensities. Thus, the best translation vector $\boldsymbol{r} = (r_x, r_y)$ is estimated for each of the scan positions (see Supplement 1, Section S1 for further details).

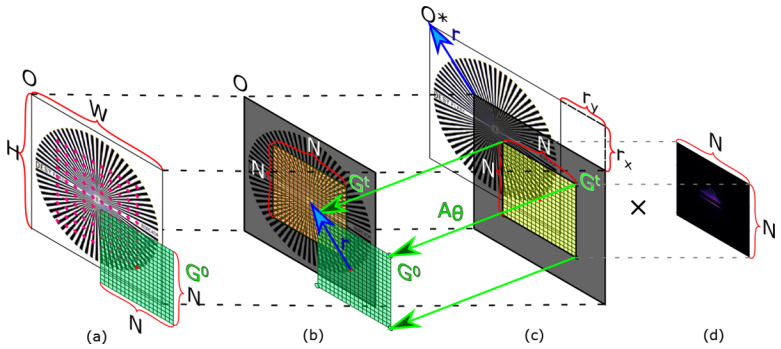


Fig. 1. Application of the affine transformation for the scan position refinement in the A-AD. (a) Sample with estimated scan positions. The red dot is the centre of the G^0 coordinates with the $N \times N$ coordinate grid (green). (b) The centered reference coordinate system G^t (orange grid) has a shift of $\boldsymbol{r}$ (blue arrow) relative to the grid G^0. (c) The sample O^* is shifted by $\boldsymbol{r} = (r_x, r_y)$ to align the part of the sample centered around the current scan point. (d) Exit wave calculated as the multiplication of the probe function with the central part of the computational array. Reference coordinate system and object coordinate system are connected by the affine transformation $\mathcal{A}_\Theta$. Its parameters are iteratively optimized to get the best estimate of $\boldsymbol{r}$ for each individual scan position.

2.3.2. Partial coherence treatment within the adaptive automatic differentiation

At FLASH2 pulses are highly but not fully spatially coherent [19]. Alongside the partial coherence, shot-to-shot fluctuations of the pulses are also considered in the forward ptychography model. The partial spatial coherence can be treated as an expansion of the wave field in terms of mutually incoherent modes in an orthogonal basis. This treatment has been proposed by Thibault et al. [3] as a state mixture reconstruction. Under this formalism a probe P is expressed as a sum of several mutually incoherent orthonormal modes M_i with corresponding modal weights α_i.

Optics EXPRESS

The resulting intensity is calculated as follows:

$$I = \sum_i \alpha_i \left| \mathcal{L}\{M_i \times O\} \right|^2 , \tag{6}$$

where the summation is wrapped over the dummy sub-indices i. However, this technique still works under the assumption of constant illumination at all scan points and should be adapted to take into account the shot-to-shot fluctuations of the beam.

In turn, a solution for ptychography with a fluctuating, but spatially coherent probe is OPR [20,21]. This method assumes that despite spatial fluctuations, probes at different scan positions are still sufficiently similar. So, the problem of reconstructing a unique probe for each individual scan position can be simplified by linking different probes to each other through a lower dimension representation. This lower dimension representation is the span of the principal component vectors which can be found by the truncated singular value decomposition (TSVD) to the matrix of all probes. This method assumes all probes to be fully coherent and should be further adapted to partially coherent illumination. Within the OPR, the measured intensity may be calculated as follows:

$$I = \left| \mathcal{L}\{\Psi\} \right|^2 = \sum_i \left| \mathcal{L}\{\alpha_i M_i\} \right|^2 , \tag{7}$$

where M_i are vectors forming the orthonormal basis of the lower dimensional space, α_i are corresponding weights, and the summation is wrapped around i.

We propose a model which simultaneously takes into account partial spatial coherence of probe and spatial fluctuations of the probe's intensity. Our model considers each FEL pulse as partially coherent. However, the number of unique modes is limited. This appropriateness of this limitation has been experimentally proven by coherence measurements performed at FLASH2 [19]. Under this assumption, each measured intensity pattern is produced by an incoherent sum as:

$$I_j = \sum_{i=1}^{K} \left| \mathcal{L}\left\{ \alpha_{ij} M_i \cdot O_{\mathbf{r_j}} \right\} \right|^2 = \sum_{i=1}^{K} \alpha_{ij}^2 \left| \mathcal{L}\left\{ M_i \cdot O_{\mathbf{r_j}} \right\} \right|^2 \tag{8}$$

where α_{ij} is a set of $K \times N$ modal weights, describing the shot-to-shot fluctuations, unique for each individual mode $i \in 1 \cdots K$ and photon pulse $j \in 1 \cdots N$. A similar representation of single-shot diffraction patterns was given in [26]. The assumption of partial spatial coherence is a key difference between our model and the OPR-based one proposed in [20]. Under this assumption, all fluctuations of the beam may be explained either as a change in corresponding modal weights or as a shift of the probe as a whole by changing the corresponding translational vector $\mathbf{r}_j$, which is found by the position refinement algorithm.

We note that the OPR approach relies on coherent summation and captures the variations in the phase and amplitude of the probe. Our approach is based on incoherent summation and considers partial coherence of the probe and spatial fluctuations of its intensity.

2.3.3. Loss function modifications: mode orthonormalization and denoising

The flexibility of the AD-based methodology permits varying the loss function independently of other parts of the reconstruction to further regularize the optimization problem. We modified the standard likelihood-based loss function (Eq. (2)) to orthonormalize the reconstructed modes and enable simultaneous denoising during the reconstruction.

Radiation generated by a partially spatially coherent source can be expressed as a sum of mutually orthogonal modes [43]. As has been shown in [44], restricting modes to an orthogonal basis dissolves the ambiguities and provides a unique reconstruction of modes.

For non-AD ptychography the mode expansion and orthogonalization could be carried out analytically within the algorithm or numerically after each iteration by TSVD. It can be seen

Optics EXPRESS

from Eq. (8) that any choice of basis vectors may lead to reconstruct the intensity. Therefore, the orthogonalization can be integrated into the phase retrieval that aids interpretation of the reconstructed modes. Otherwise, the orthogonalization can be implemented as a final step after the reconstruction is finished.

Both approaches enhance the analytical and computational complexities without leaving any flexibility. However, in the case of an A-AD reconstruction, this constraint can be integrated directly into the loss function by adding the following regularization term:

$$R(M) = \sqrt{\sum_{i,j}^{K} \left| M_b^{ia} M_{ja}^b - \mathbb{I}_K \right|^2}, \tag{9}$$

where $M_b^{ia} M_{ja}^b = G_j^i,\ j, i \in 1 \cdots K$ is the Einstein notation for a $K \times K$ matrix of all possible mutual dot products of individual modes, and $\mathbb{I}_K$ is the identity matrix of rank K. This regularization forces the reconstructed modes to be orthonormal. Therefore, orthonormalization of the reconstructed modes can be integrated into the loss function and does not dramatically increase the reconstruction time.

Additionally, for increasing the robustness of the reconstruction to noise in the measured intensity patterns, we perform total variation denoising (TVD) [45] by adding a discrete version of a regularizer to the loss function:

$$T(O) \sim |\nabla O| = \sum_{i,j}^{N-1} \sqrt{\left| O_{i+1,j} - O_{i,j} \right|^2 + \left| O_{i,j+1} - O_{i,j} \right|^2}. \tag{10}$$

TVD is an approach for noise reduction developed to preserve sharp edges in the underlying signal. Unlike a conventional low-pass filter, TVD is defined in terms of an optimization problem and can be added to the existing loss function. Additionally, it avoids discontinuities in the reconstructed sample and probe modes.

After adding all aforementioned regularization terms the resulting loss function $L(I, \tilde{I}, O, M)$ used in the optimization process wraps as follows:

$$L(I, \tilde{I}, O, M) = \mathcal{E}(I, \tilde{I}) + \gamma_1 T(O) + \gamma_2 R(M), \tag{11}$$

where $\mathcal{E}(I, \tilde{I})$ is the likelihood term, $T(O)$ is the TVD regularization term, $R(M)$ is the regularization term responsible for the modal orthonormalization, and γ_1, γ_2 are coefficients that determine the force of the regularizers. The γ_1, γ_2 values may be changed dynamically during the reconstruction process to enable or disable certain regularization terms. Currently, there are multiple gradient-based optimizers available out-of-the-box for most computational frameworks. Performance of these methods for the purpose of phase retrieval have never been compared and need to be evaluated. In our work, we utilized the state-of-the-art adaptive moment estimation (Adam) algorithm [46] (see Supplement 1, Section S2 for further details).

2.4. Adaptive automatic differentiation ptychographic model

We used a forward model which utilizes the orthogonal probe decomposition and the scan position refinement algorithms described in sections 2.3.2, 2.3.3, and 2.3.1 for the reconstruction. The forward pass of the model for calculation of the reconstructed intensity at the j-th scan position is expressed as

$$I_j = \sum_i \left| \mathcal{L} \left\{ \alpha_{ij} M_i \cdot S \cdot A_{\theta_j} \{O\} \right\} \right|^2, \tag{12}$$

where A_{θ_j} represents an affine transform driven by its parameter matrix θ_j, S is the support of the probe function, $\mathcal{L}$ is the propagator determined by the experiment configuration. Due to the

Optics EXPRESS

experimental geometry we used a Fresnel propagator. The computation graph of reconstruction is given in Fig. 2.

The algorithm of the A-AD ptychographic reconstruction is illustrated below.

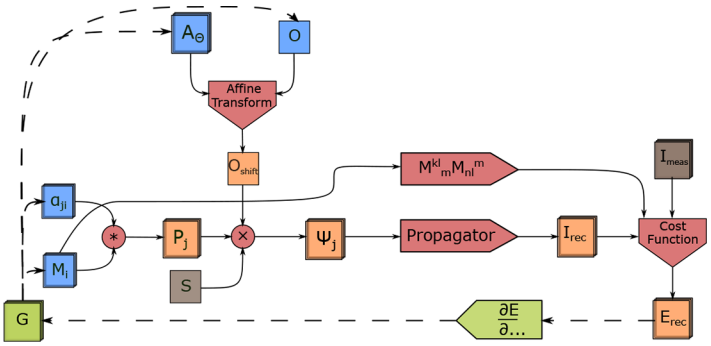


Fig. 2. Computational graph of the A-AD model. Input data (measured intensities - I_{meas}) in brown, operators in red, optimizable variables (modes - M_i , modal weights - α_{ji}, affine transformation matrix - $\mathcal{A}_\theta$, and object function - O) in blue, intermediate variables in orange, backward propagation of error represented with dotted line and green color.

Algorithm 1: A-AD ptychographic reconstruction

Data: Measured intensities I, scan positions $\boldsymbol{r}_j$, initial guesses of the probe modes M_i, modal weights α_i, sample O, support S learning rates lr for each optimizable parameter, number of the iterations N, number of batches B.

for $n \in 0...N$ **do**

  for $b \in 0...B$ **do**

    get scan numbers J for this batch number b

    for $j \in J$ **do**

      $I_j \longleftarrow \sum_i \left|\mathcal{L}\left\{\alpha_{ij} M_i \cdot S \cdot A_{\theta_j}\{O\}\right\}\right|^2$

      $L(I, \tilde{I}, O, M) = \mathcal{E}(I, \tilde{I}) + \gamma_1 T(O) + \gamma_2 R(M)$

      Calculate gradients of the $\mathcal{E}$ w.r.t all optimizable parameters

      Update all of the parameters by performing step of the ADAM algorithm

Result: Reconstructed sample O, probe modes M_i, modal weights α_{ij}, scan positions $\boldsymbol{r}_j$

Currently, there are two main Python frameworks supporting AD which are natively scalable across multiple GPUs - PyTorch [47] and TensorFlow [48]. Both of them were developed mostly to be used in neural network development and provide out-of-the-box state-of-the-art optimization algorithms. PyTorch was selected as the numerical framework for our model.

Overall, ptychography at FELs can be seen as an optimization problem with respect to the sample, scan positions, probe modes and modal weights. Our A-AD ptychography framework was designed to solve these type of problems and provide high-resolution sample reconstruction together with single-shot reconstruction of the probe. This reconstruction of single-shot, varying, partially spatially coherent probes was achieved with the aid of the A-AD engine.

2.4.1. Experimental setup

In the first ptychography experiment at FLASH the described model was applied and verified. As has been shown in [19], the FLASH2 photon pulses are not fully spatially coherent, with the global degree of spatial coherence ζ estimated to be approximately 0.56 (depending on wavelength and accelerator setup). Therefore, the sample was not illuminated fully coherently. Also, in the same work, a degree of beam position instabilities was observed during the data collection

Optics EXPRESS

that significantly influenced the data analysis and interpretation. Consequently, performing ptychography requires an approach which tackles robustly those challenges and responds to the fluctuations.

The experiment was carried out at the beamline FL24 at FLASH2 at a wavelength of $\lambda = 13.5\ nm$ corresponding to $E_{ph} = 91.8\ eV$ with an average photon pulse energy $E_{pulse} = 90\ \mu J$.

The experimental setup is shown in Fig. 3. The FLASH2 beam was focused using bendable Kirkpatrick-Baez (KB) mirrors to focus approximately 2.6 m after the second mirror [12,49]. At this position the KB optics produces a focus of $\approx 10\times10\ \mu m^2$ (FWHM), which was tuned and characterized using a wavefront sensor [50]. The sample was placed approximately 12 cm behind the focus leading to a beam size of approximately $300 \times 300\ \mu m^2$. An aperture of 700 μm diameter was placed 1 cm before the sample to filter out those photon pulses whose pointing was extremely far from the current scanning position. A gas attenuator and metal foil filters were used to attenuate the photon beam and prevent sample damage, as well as suppress higher order FEL radiation [51]. Attenuation was also necessary to avoid detector saturation, since no central beamstop was used.

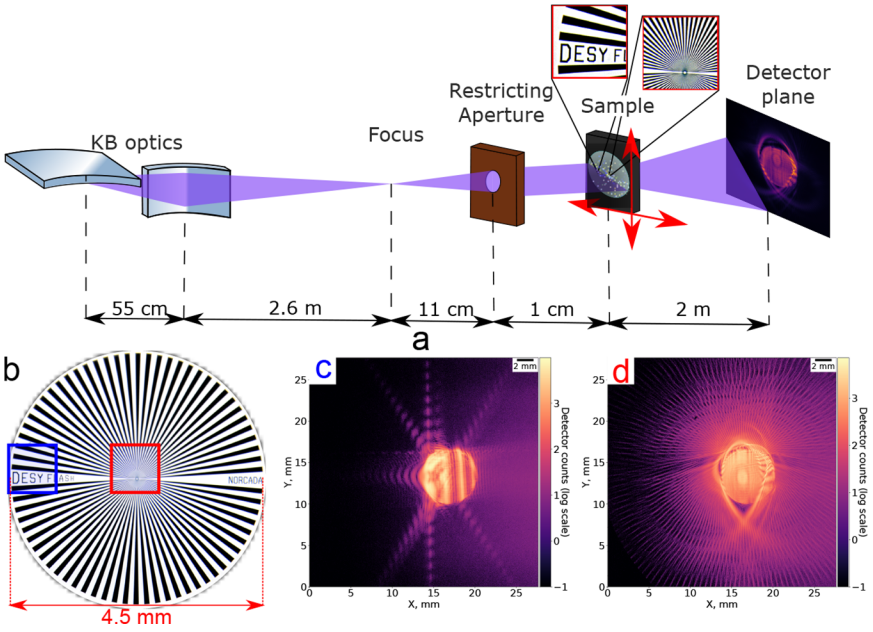


Fig. 3. (a) Schematic outline of the experimental setup. The FEL beam was focused using KB mirrors, thus creating a divergent probe geometry. A restricting aperture several times larger than the beam cross-section was used to limit spatial fluctuations. The sample was placed approximately 12 cm after the focus position. An ANDOR iKon-L SO CCD was placed in the intermediate field 2 m downstream of the sample. Scanned areas of the sample shown in the insets. (b) displays the sample and (c-d) are the diffraction patterns measured by scanning the blue and red boxes.

A $4.5 \times 4.5\ mm^2$ Siemens star pattern produced by Norcada Inc. on a 150 nm thick SiN membrane was used as the sample. The opaque part of the Siemens star was made of 190 nm thick Au and 10 nm thick Cr layers and is thin enough to fulfil the projection approximation for $\lambda = 13.5nm$. At this wavelength, it may be considered a binary black-white object. The sample has features of 0.5 μm and 300 μm size in the central(stripes) and border(letters) regions respectively as shown in the inset in Fig. 3.

Optics EXPRESS

The sample was mounted on a 2-axis translation stage with scanning axes perpendicular to the optical axis. It was scanned in two regions, the borders of the sample (letters *DESY FL*) and the central stripes regions. We used 11×11 scan positions on a rectangular grid, with a $20\mu m$ step size which is equivalent to 90% overlap. This high degree of overlap was used to ensure convergence of the reconstruction. The scanned areas (letters and stripes) are shown in the insets in Fig. 3. Ten measurements per position were taken and the most intense diffraction pattern selected for the analysis, corresponding to the FLASH pulse with the maximum photon flux. The restricting aperture reduced the influence of not ideal pointing stability on the beam positioning. However, we observed no effect of it in the beam, which means that the pointing stability was significantly better than the size of the aperture.

An ANDOR iKon-L SO CCD camera with a pixel size of 13.5 μm and 2048×2048 pixels was placed 2 m behind the sample to measure the diffraction patterns. This results an oversampling factor $\lambda . z_{sd}/p_{det}. D = 2.5$. Here, λ is the wavelength, z_{sd} is the sample-to-detector distance, the detector pixel size is p_{det} and D is the diameter of the beam or the aperture confining the beam [37,52].

Based on geometry and wavelength, the paraxial Fresnel propagator $\mathcal{L}$ is given as [35]

$$\mathcal{L}_{z,f}\left\{\Psi_{z_0,\rho_0}\right\} = \Psi_{z,\rho'} = -\frac{i}{\lambda z}\exp(i\frac{k}{2z}\rho'^2)\cdot\mathcal{F}\left\{\exp(i\frac{k}{2z}\rho_0^2)\cdot\exp(-i\frac{k}{2f}\rho_0^2)\cdot\Psi_{z_0,\rho_0}\right\}, \tag{13}$$

where $\Psi_{z,\rho}$ is the complex wave field at a distance z in transverse coordinates $\rho = (x, y)$, ρ_0 and ρ' denote the coordinates at the initial and resulting plane respectively. $\mathcal{F}\{\cdot\}$ is the Fourier transform, z is the propagation distance, f is the distance to the focus position, and k is the wave number.

2.5. Computational parameters

Although the proposed model simplifies the reconstruction process, it cannot be efficiently applied for a ptychographic reconstruction without choosing a suitable optimizer and optimal set of learning rates. They can be adjusted using different techniques such as a grid search or considered a globally constant value. In this work, we followed [53] to finetune the learning rates (see Supplement 1, Section S2 for further details). This procedure was repeated for each optimization parameter and a final set of the learning rates l_r was obtained as reported in Table 1.

Correctly selected learning rates guarantee the fastest convergence to a unique and stable solution during the reconstruction. Too low values of the learning rates result in very small updates to the optimized parameters and therefore too slow convergence, while too high values lead to rapid changes of the corresponding parameters and instability of the algorithm. However, small deviations from the optimal learning rates still allow the algorithm to reach a solution at the cost of longer computational time.

Table 1. Optimizable parameters and corresponding learning rates.

	Symbol	Learning rate
Object	O	5×10^{-2}
Probe modes	M_i	1×10^{-3}
Probe weights	α_{ji}	1×10^{3}
Affine parameters	$\mathcal{A}_\theta$	5×10^{-3}

During this work we always used the maximum possible batch size to speed-up the computations. The optimization routine was distributed across four Quadro M6000 Nvidia GPUs with 12 GB memory on each. This distribution allowed us to set the batch size to 120 for a single mode (averaged coherent pulse) and 40 scan points for multi-modal reconstructions with 9 modes.

Optics EXPRESS

The central 1024 × 1024 pixels of the diffraction data were selected for reconstruction to ease the computational costs. Background noise was subtracted by using an average of 100 dark frames with a subsequent zeroing of the negative values. The reconstruction was carried out using the model described in Section 2.4. The initial probe assumption was obtained by backpropagating of the square root of the averaged intensities measured without sample in the beam to the sample plane.

$$P = \mathcal{L}^{-}\left\{\sqrt{\bar{I}_j}\right\} \tag{14}$$

The sample was initiated with random module values ranging between 1 and 0 and a planar phase. During the reconstructions, we have used several different starting guesses to initiate the reconstructions including a fully randomized object transparency ranging between [0,1] and phase in [0,2π]. Also, a fully transparent object and an object with randomized transmission varying in the range of [0,1] with a planar phase were used. Overall, for all given scenarios, the reconstructions converged reliably to a similar final error.

The loss function in form of Eq. (11) was optimized during the reconstruction process. Additionally, the quality of the reconstruction was estimated by the average error per pixel as follows:

$$E_{rec} = \frac{\sum_{j=1}^{N} \left|I_j - \tilde{I}_j\right|^2}{N \sum_{j=1}^{N} I_j^2}, \tag{15}$$

where I_j is the reconstructed intensity, $\tilde{I}_j$ is the measured intensity, and N is the batch size.

3. Results

3.1. Preliminary reconstruction using automatic differentiation

As a first step, we performed a reconstruction with an AD-based ptychography model without any adaptations. This approach not only gives an insight into the pure AD-based model, but also yields a result similar to the common and well-known ptychographic engines [28]. The same fully coherent illumination was assumed for all scan positions and position refinement was not performed. The reconstruction procedure consisted of 200 iterations of an object-only optimization followed by 350 iterations of joint optimization of object and probe. The reconstruction of the stripes region did not converge to any meaningful result and the reconstructed letters-area was significantly distorted. The reconstructed probe and sample functions are shown in Fig. 4. The results illustrate that AD-based models may not perform satisfactorily in ptychography experiments at FLASH2 (a SASE FEL) since they cannot represent the spatial fluctuations of the beam and its partial coherence correctly. Therefore, we incorporated additional parameters into the AD-model to form an A-AD variant.

3.2. Reconstruction with the adaptive automatic differentiation based model

3.2.1. Effect of the automatic differentiation based position correction

First, a reconstruction using a fully coherent probe and a simultaneous scan position refinement was conducted to evaluate the performance of the position correction procedure and to determine if a reconstruction under this assumption is appropriate for SASE FELs.

The optimization procedure consisted of a total of 1700 iterations. During the first 200 iterations only the sample function and the scan positions were optimized. Then, the probe function was included for 1000 iterations with a TVD regularization. The TVD regularization preserves the structural details while removing noise and discontinuities. So, over- or under regularization with the TVD term may degrade the overall resolution by removing small sharp features.

Optics EXPRESS

Additionally, during the first 700 iterations the numerical support was applied. This step is not critical for the algorithm convergence, but, as shown in [28], it can speed up the reconstruction by centring the probe.

The reconstructed sample functions for the stripes and for the letter-regions obtained after this step are shown in Fig. 5(a) and (b), respectively. The scan position refinement retrieved some features missing in the first reconstruction such as the upper lines in Fig. 5(b), and removed striped artefacts present in the probe and the sample reconstructions using raw positions (red arrows in Figs. 4(b), 4(a)). In this step, the stripes-part of the sample (Fig. 5(a)) converged to a meaningful result. However, it was reconstructed only in the very centre and was accompanied with numerical artefacts. The reconstructed probe function is shown in Fig. 5(c). It can be seen as an averaged coherent beam illuminating the sample during the ptychographic scanning.

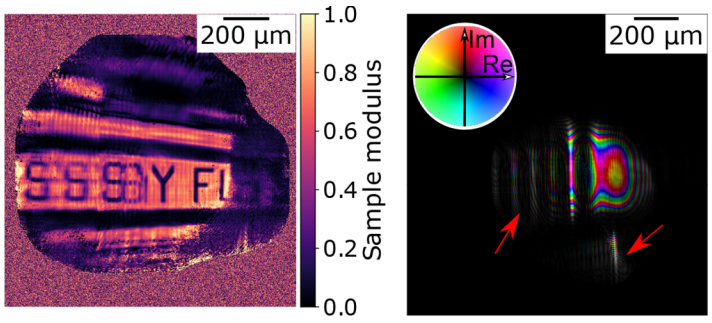


Fig. 4. Results of the ptychographic reconstruction with the AD-based ptychography model using the approximation of a fully coherent and unchanging probe for all of the scan positions without position refinement. (a) reconstructed sample amplitude from the letters area. (b) reconstructed probe function, complex values represented by domain coloring with saturation representing the modulus and hue representing the phase according to the colorwheel. Reconstruction artifacts resulting from imperfectly known positions shown by red arrows.

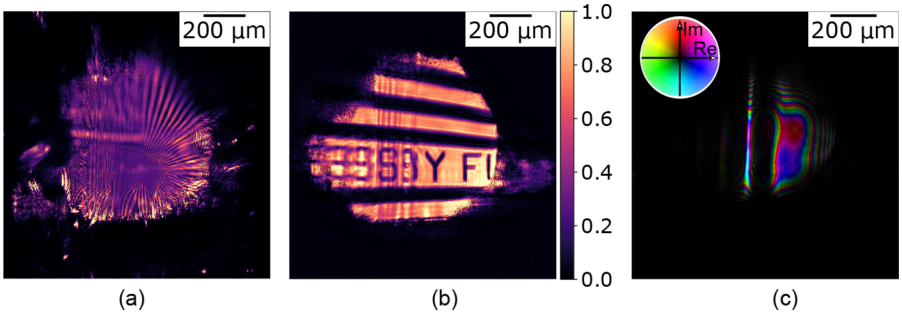


Fig. 5. Results for the reconstruction using an assumption of a fully coherent average beam with a simultaneous position correction. (a) Reconstructed modulus of the stripes-area of the sample. (b) Reconstructed modulus of the letter-area of the sample. (c) Reconstructed probe assuming full coherence and the same illumination for all of the scan points; represented with domain colouring according to the colour wheel.

To evaluate the stability of the refined scan positions we performed several independent reconstructions. The obtained scan position maps are shown at the Fig. 6 together with the initially assumed scan positions. As the scan position refinement algorithm is stable enough to

Optics EXPRESS

reconstruct most of the scan positions within sub-pixel distance from their previous values during the trial-to-trial examination, it can be concluded that it works reliably.

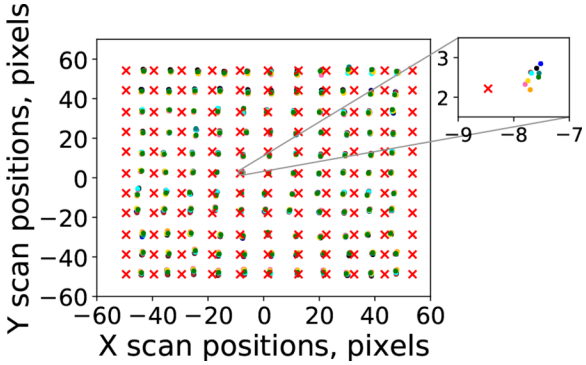
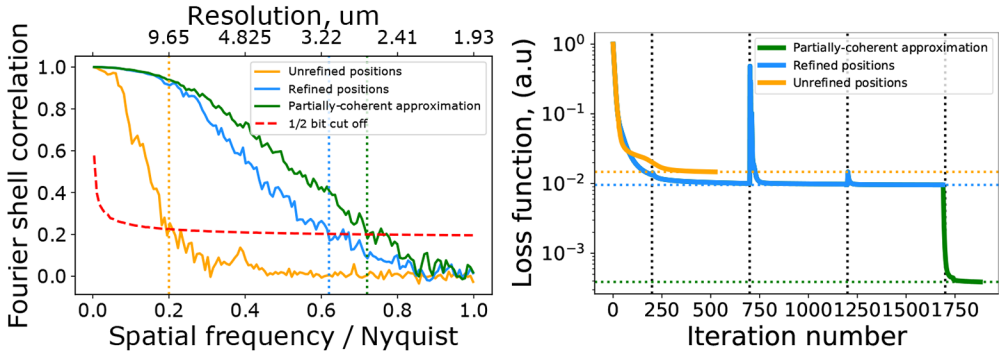


Fig. 6. scan position maps (red crosses) and those obtained from the AD-based position refinement algorithm (circles, different colors correspond to different independent reconstructions). Most of the retrieved scan positions show a sub-pixel stability.

After this step (1700 iterations) of the reconstruction, the value of the loss function (blue curve Fig. 7(b)) was reduced down to 50% of the minimum loss function value achieved during the AD-based reconstruction without position refinement (orange curve Fig. 7(b)) with an average error per pixel $E_{rec} = 3\%$. The resolution of the reconstruction was estimated by the Fourier ring correlation (FRC) method [54] using two independent reconstructions of the letter region. The resolution at 0.5 bit cut-off was found to be approximately $3.11\mu m$ (blue curve Fig. 7(a)).



(a) (b)

Fig. 7. Comparison of the FRC curves (a) and loss function values (b) of different reconstruction regimes. For both (a) and (b): AD-based reconstruction - orange curves; Average beam with position refinement - blue curves; Multi-modal - green curves. (a) Resolution estimated using FRC. The red dashed line represents the 0.5 bit threshold. (b) Evolution of the loss functions during the reconstruction process. Dashed vertical lines mark: the iterations at which the probe optimization was started – 200, support was removed – 700, TV denoising of the sample was started – 1200, and the reconstruction in the partially coherent approximation was started - 1700.

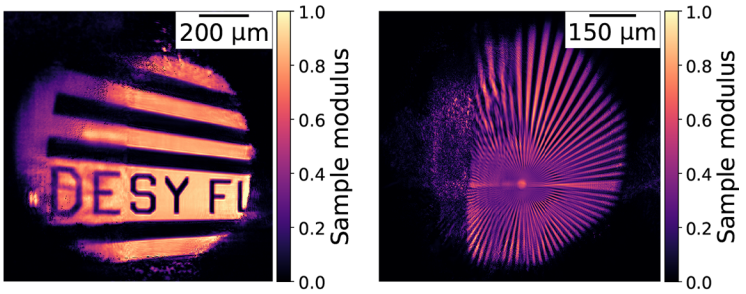
3.2.2. Reconstruction with a partially coherent and spatially fluctuating beam

At the next stage of the reconstruction, the obtained sample and the average probe function were used as the initial data for the multi-modal reconstruction. It lasts for 200 more iterations during

Optics EXPRESS

which the sample, the main modes and the unique modal weights for each of the scan positions were reconstructed. The reconstructed main modes and their average weights are shown in Fig. 9. The average error per pixel was estimated as $E_{rec} = 0.1\%$. The loss function (green curve in Fig. 7(b)) has decreased more than one order of magnitude in comparison to the reconstruction using an average beam and position refinement (blue curve in Fig. 7(b)).

The reconstructed letter-part of the sample is shown in Fig. 8(a). The left part of the sample which was distorted when using the fully coherent reconstruction, (see Fig. 5(b)), can now be clearly distinguished using the multi-modal reconstruction. Additionally, the upper left part of the sample, which was not reconstructed previously due to a dim illumination, is now visible. The resolution resulting from this reconstruction was estimated as 2.68 μm (green curve in Fig. 7(a)). Finally, the multi-modal reconstruction procedure with position correction was repeated for the central region of the sample with the modes obtained for the letter-region reconstruction used as the initial assumption for the probe. The reconstruction of the stripes-region is shown in Fig. 8(b). In contrast to the fully coherent reconstruction, the multi-modal reconstruction was able to retrieve a much larger part of the sample with an average error per pixel of $E_{rec} = 0.5\%$. The image obtained is partially blurred as the size of the stripes in the center is less than a micrometer.



(a) (b)

Fig. 8. Sample modulus in the letters (a) and stripes (b) areas reconstructed in the multi-modal approximation.

3.3. Single-shot beam characterization

The unique properties of each SASE FEL pulse prompt the use of ptychography models capable of reconstructing not only average but also single-shot probe properties. Our A-AD based model can reconstruct the main coherent modes of the FEL radiation together with unique modal weights for each individual shot. The modal expansion can be used to reconstruct and analyze each FEL shot individually. We used the reconstructed coherent modes and modal weights to characterize the global degree of spatial coherence of the FEL radiation.

The reconstructed coherent modes M_i describing the single-shot probe variations are shown in Fig. 9 with their average weights α_i. It is possible to determine an average global degree of transverse coherence of the FLASH radiation considering the retrieved modes and their average weights α_i from Fig. 9. If all the modes originate from the partial coherence of the FLASH2 source we can estimate its degree of transverse coherence from the mode weights [55] as

$$\zeta = \frac{\sum \beta_i{}^2}{(\sum \beta_i)^2} = 54\% \; ; \; \beta_i = \alpha_i^2. \tag{16}$$

This result is in agreement with coherence measurements performed at FLASH2 [19] and FLASH1 [25].

Optics EXPRESS

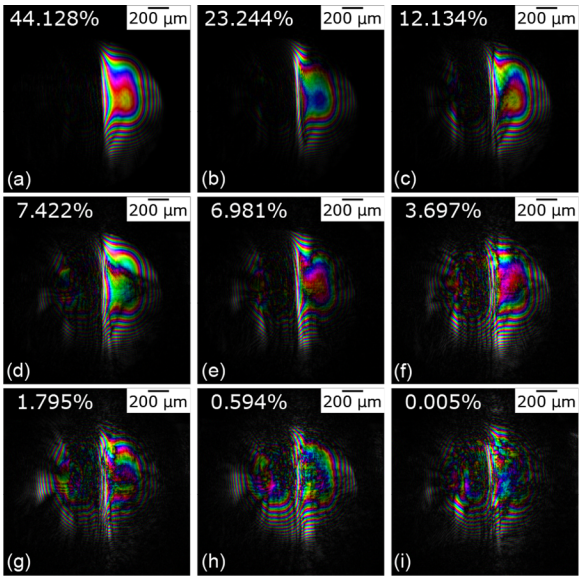


Fig. 9. Probe modes at the object plane reconstructed under partial coherence approximation and their average weights. Amplitude mapped in brightness and phase in hue. Average modal weights shown in upper-left corners.

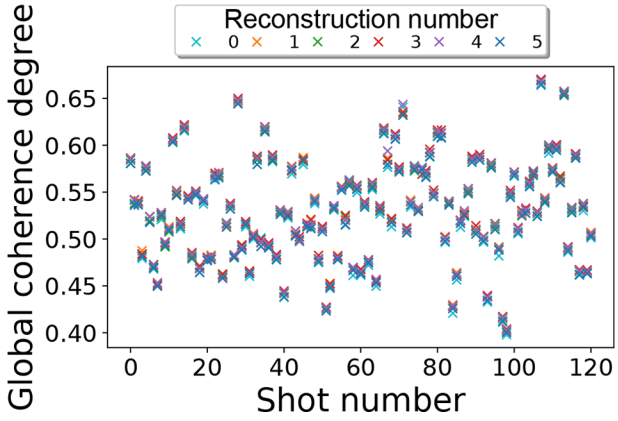


Fig. 10. Global degree of spatial coherence calculated for each reconstructed single-shot probe. Results obtained from several independent reconstructions are plotted on top of each other and distinguished by color. The overlap of the obtained values indicates a high stability of the reconstruction.

Optics EXPRESS

It is also possible to use the retrieved modal weights α_{ij} to estimate the shot-to-shot fluctuations of the FLASH pulses and their spatial coherence, since our reconstruction algorithm provides unique modal weights α_{ij} for each individual shot. To ensure the stability of the reconstruction with respect to the coherence properties, we performed several independent multi-modal reconstructions from the same diffraction patterns measured in the letter-area. The single-shot global degree of spatial coherence was determined from each of the reconstructions and for each scan position according to Eq. (16). The obtained values are presented in Fig. 10 and show that our algorithm steadily converges to the same values of single-shot global coherence degree for independent reconstructions.

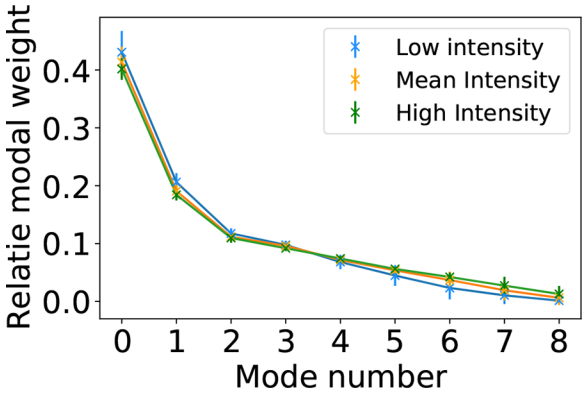
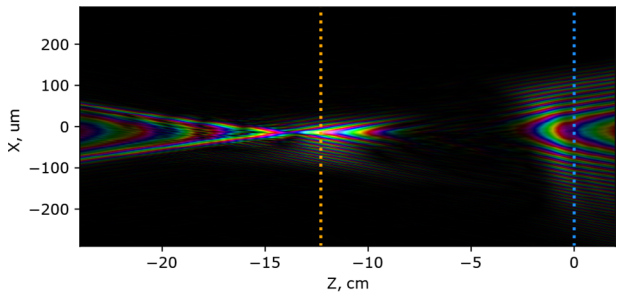


Fig. 11. Distribution of the reconstructed modal weights for three groups of pulses binned by their intensity. The higher relative weight of the first mode corresponds to more coherent shots.

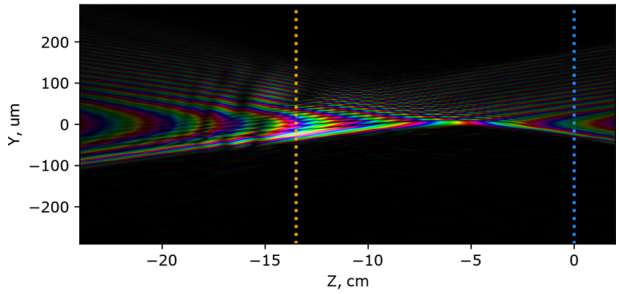
Further analysis of the modal weights distribution for individual shots may give additional statistical insights about the SASE FELs performance. The modal weights distribution for pulses grouped by their intensity into three different equal sized bins is shown in Fig. 11. We did not find a significant dependence between the intensity of the pulses and their coherence degree.

The access to the single-shot modal weights gives us the ability to spatially characterize each of the reconstructed shots individually (see Supplement 1, Section S3), while average characteristics may be estimated from the investigation of the main mode (with the highest average weight). We numerically propagated the main reconstructed mode into the vicinity of the focal plane to investigate the focus of the KB mirrors. The sections of the propagated main mode are shown in Figs. 12(a),12(b). The blue and orange dashed lines indicate the sample and focus positions respectively. The x and y focal planes were estimated by the maximum sharpness of the propagated beam [56] at 12.3 *cm* and 13.5 *cm* upstream of the sample. The main mode propagated to the plane of the second KB mirror is shown in Fig. 12(c) The reconstructed modes and their single-shot weights may be used for the reconstruction of the individual FLASH pulses.

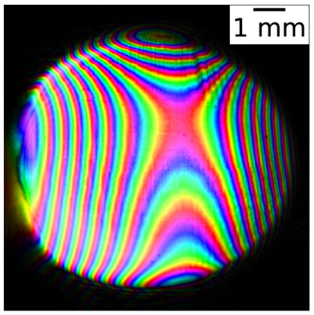
Optics EXPRESS



(a)



(b)



(c)

Fig. 12. Horizontal (a) and vertical (b) caustics of the propagation of the reconstructed main mode. Sample and focal planes shown with blue and orange lines, respectively. The horizontal and vertical focal planes are different due to the astigmatism that the KB system had in this special setup. (c) Main mode backpropagated to the second KB mirror plane with tilts and defocus removed. Complex values shown with domain coloring, with modulus mapped to saturation and phase mapped to hue.

Optics EXPRESS

4. Discussion

In this work, A-AD ptychography was introduced as a practical scheme for performing ptychographic imaging at FELs. In the proposed implementation, we considered the effects of fluctuations of the probe, such as pointing instabilities, intensity fluctuations, partial spatial coherence and wavefront variations, and limited precision of the sample scanning. A proper forward model to optimize pulse and sample recovery was developed. A more generalized implementation can involve complex noise and detector response models [57], multi-slice sample recovery [7] and multi-color forward models [58].

As a proof of concept, we have performed ptychographic imaging at the FLASH2 variable micro-focus beamlime FL24 [12]. A successful reconstruction of a sample and a partially coherent probe was performed. The single-shot probe fluctuations were estimated using the reconstructed main modes and modal weights. We performed the reconstruction at an intermediate Fresnel distance [35], without cutting the beam or selecting the most coherent part. This allowed us to utilize the full flux of the FEL source, and, advantageously, perform a detailed shot-to-shot FEL pulses characterization.

The set of retrieved pulses was represented in the basis of mutually incoherent modes with the varying modal weights. The existence of such a common basis for the FEL pulses measured at the FLASH2 allows the application of the mixed-state ptychography at FELs. Indeed, a series of several-shot long-exposure measurements can be performed for each scan position. Averaging through the intensity measured for each scan position increases the photon flux and the signal-to-noise ratio [26]. Since the basis is known, the average intensity can be expressed in vectors of it. Therefore, the application of mixed-state ptychography [3] is extended to FELs with a chance of obtaining a better reconstruction.

A 90%-overlap scanning pattern was combined with a restricting aperture to decrease the expected significant spatial fluctuations of the FLASH pulses. Since the restricting aperture was several times larger than the beam size, we never observed that it plays a significant role in the analysis. Additionally, the scan position correction algorithm converged steadily for all positions. Therefore, it would be possible to decrease the overlap safely to 60% preserving the sampling requirements [52] and without disturbing the algorithm convergence.

We described the influence of partial coherence in the model by expressing the probe in terms of orthogonal modes. The modal expansion continued until the last mode's weight occupied below 1% of the total intensity distribution. A lower rank representation may prohibit a full description of the probe. On the other hand, a higher rank of representation will increase the computational load. A preliminary understanding of the modes can be obtained with various techniques such as [17,25,59]. We also note that multi-modal reconstruction can improve the retrieval quality for fully coherent pulses since less significant modes (appearing after the dominant) can absorb the noise of the diffraction patterns. Generally, the multi-modal analysis always results in a better reconstruction for SASE FELs with partial coherence. Notably, analysis of the modal weight could also be employed to improve the performance of the beamline optics by optimizing the focusing of the most significant contributing modes.

The experimental geometry can be further optimized to improve the achievable resolution. Decreasing the sample to detector distance together with use of the shortest possible focusing distance for the KB mirrors would allow for sub μm resolution at $\lambda = 13.5nm$ at beamline FL24 at FLASH. Another way of improving the resolution would be the use of a detector with higher dynamic range and possibly a beamstop.

We used a bilinear kernel for differentiable sampling in our scan position correction routine. In general, any formalism providing gradients of the error metrics with respect to the scan positions and the sample function can be used. We chose a bilinear kernel due to its low computational load and artefact-free results. We note that most of the similarly developed algorithms [5,6]

Optics EXPRESS

cannot be integrated within the forward model which significantly reduces the possibility of parallel computation and increases the computational load.

Within this work, we used an Adam optimizer. Other gradient-based optimization techniques can be straightforwardly substituted. However, the analysis of the performance is out of the scope of this work.

The global degree of coherence estimated from the reconstructed data in Fig. 10 varies from 40% to 60%. The fluctuations are mostly produced by changes of the intensity of the dominant FEL modes. However, a further study is required to estimate the precision of the proposed method and determine the minimum values of the global degree of coherence which will still allow successful ptychographic reconstruction. Additionally, a thorough study on the changes of the estimated global degree of coherence for different FLASH settings is required.

5. Conclusion

We have developed an adaptive AD-based forward model including a fully AD-based position refinement routine which is capable of analyzing ptychography experiments at SASE FELs. We applied it to the first ptychography measurements at the FLASH2 facility. The A-AD approach enabled us to successfully retrieve a test sample and characterize single-shot pulses of a SASE FEL without suppressing inherent fluctuations or selecting the most coherent part of the photon pulses.

The AD nature of the presented work removes the need for cumbersome analytical calculations and allows significant methodological advancements for hard and soft X-ray CDI. Applications of our A-AD routine can be extended to the recently growing field of electron ptychography, as well as 3D ptycho-tomography at FELS and synchrotrons. In addition, all techniques discussed in this work can be applied directly to the field of visible and Fourier ptychography as well.

Funding. Deutsches Elektronen-Synchrotron; Horizon 2020 Framework Programme (730872).

Acknowledgments. The authors would like to thank the Maxwell and FLASH teams for support during this research.

Disclosures. The authors declare that they have no conflict of interest. M.M. and E.P. conceived the idea and supervised this work. K.K. and M.M. contributed to the development of the theory and the algorithms. S.K. M.R-L, B.K., M.M., and K.K. designed experiments. M.R-L, E.P., B.K., M.M., K.K., M.P. and M.S. performed the experiments as well as contributing to the data acquisition. K.K. and M.M. analyzed the data. All authors participated in the writing and the revisions of the manuscript.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

References

1. F. Pfeiffer, "X-ray ptychography," Nat. Photonics **12**(1), 9–17 (2018).
2. P. Thibault, M. Dierolf, A. Menzel, O. Bunk, C. David, and F. Pfeiffer, "High-Resolution Scanning X-ray Diffraction Microscopy," Science **321**(5887), 379–382 (2008).
3. P. Thibault and A. Menzel, "Reconstructing state mixtures from diffraction measurements," Nature **494**(7435), 68–71 (2013).
4. A. Maiden, M. Humphry, M. Sarahan, B. Kraus, and J. Rodenburg, "An annealing algorithm to correct positioning errors in ptychography," Ultramicroscopy **120**, 64–72 (2012).
5. 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," Opt. Express **21**(11), 13592 (2013).
6. P. Dwivedi, A. Konijnenberg, S. Pereira, and H. Urbach, "Lateral position correction in ptychography using the gradient of intensity patterns," Ultramicroscopy **192**, 29–36 (2018).
7. A. Suzuki, S. Furutaku, K. Shimomura, K. Yamauchi, Y. Kohmura, T. Ishikawa, and Y. Takahashi, "High-Resolution Multislice X-Ray Ptychography of Extended Thick Objects," Phys. Rev. Lett. **112**(5), 053903 (2014).
8. P. Thibault, M. Guizar-Sicairos, and A. Menzel, "Coherent imaging at the diffraction limit," J. Synchrotron Radiat. **21**(5), 1011–1018 (2014).

Optics EXPRESS

9. A. He, L. Yang, and L. Yu, "High-gain free-electron laser theory, introduction," *Synchrotron Light Sources and Free-Electron Lasers: Accelerator Physics, Instrumentation and Science Applications* pp. 149–190 (2020).
10. K. Tiedtke, A. Azima, N. von Bargen, L. Bittner, S. Bonfigt, S. Düsterer, B. Faatz, U. Frühling, M. Gensch, C. Gerth, N. Guerassimova, U. Hahn, T. Hans, M. Hesse, K. Honkavaar, U. Jastrow, P. Juranic, S. Kapitzki, B. Keitel, T. Kracht, M. Kuhlmann, W. B. Li, M. Martins, T. Núñez, E. Plönjes, H. Redlin, E. L. Saldin, E. A. Schneidmiller, J. R. Schneider, S. Schreiber, N. Stojanovic, F. Tavella, S. Toleikis, R. Treusch, H. Weigelt, M. Wellhöfer, H. Wabnitz, M. V. Yurkov, and J. Feldhauss, "The soft x-ray free-electron laser flash at desy: beamlines, diagnostics and end-stations," New J. Phys. **11**(2), 023029 (2009).
11. F. Mayet, R. Aßmann, S. Schreiber, and M. Vogt, "Implementation of a diagnostic pulse for beam optics stability measurements at flash," MOPHA031, These Proceedings, IPAC **15**, 850–853 (2015).
12. E. Plönjes, B. Faatz, M. Kuhlmann, and R. Treusch, "FLASH2: Operation, beamlines, and photon diagnostics," in *AIP Conference Proceedings*, vol. 1741 (2016), p. 020008.
13. E. L. Saldin, E. A. Schneidmiller, and M. Yurkov, "Statistical properties of radiation from vuv and x-ray free electron laser," Opt. Commun. **148**(4-6), 383–403 (1998).
14. E. Saldin, E. Schneidmiller, and M. Yurkov, "Study of a noise degradation of amplification process in a multistage hghg fel," Opt. Commun. **202**(1-3), 169–187 (2002).
15. S. Schreiber and B. Faatz, "The free-electron laser flash," High Power Laser Sci. Eng. **3**, e20 (2015).
16. A. Schropp, R. Hoppe, V. Meier, J. Patommel, F. Seiboth, H. J. Lee, B. Nagler, E. C. Galtier, B. Arnold, U. Zastrau, J. B. Hastings, D. Nilsson, F. Uhlén, U. Vogt, H. M. Hertz, and C. G. Schroer, "Full spatial characterization of a nanofocused x-ray free-electron laser beam by ptychographic imaging," Sci. Rep. **3**(1), 1633 (2013).
17. A. Singer, F. Sorgenfrei, A. P. Mancuso, N. Gerasimova, O. M. Yefanov, J. Gulden, T. Gorniak, T. Senkbeil, A. Sakdinawat, Y. Liu, D. Attwood, S. Dziarzhytski, D. D. Mai, R. Treusch, E. Weckert, T. Salditt, A. Rosenhahn, W. Wurth, and I. A. Vartanyants, "Spatial and temporal coherence properties of single free-electron laser pulses," Opt. Express **20**(16), 17480–17495 (2012).
18. I. A. Vartanyants and A. Singer, "Coherence properties of third-generation synchrotron sources and free-electron lasers," *Synchrotron Light Sources and Free-Electron Lasers: Accelerator Physics, Instrumentation and Science Applications* pp. 987–1029 (2020).
19. T. Wodzinski, M. Mehrjoo, M. Ruiz-Lopez, B. Keitel, M. Kuhlmann, M. Brachmanski, S. Künzel, M. Fajardo, and E. Plönjes, "Single-shot transverse coherence measurements with Young's double pinholes at FLASH2," J. Phys. Commun. **4**(7), 075014 (2020).
20. S. Sala, B. J. Daurer, M. Odstrcil, F. Capotondi, E. Pedersoli, M. F. Hantke, M. Manfredda, N. D. Loh, P. Thibault, and F. R. N. C. Maia, "Pulse-to-pulse wavefront sensing at free-electron lasers using ptychography," J. Appl. Crystallogr. **53**(4), 949–956 (2020).
21. M. Odstrcil, P. Baksh, S. A. Boden, R. Card, J. E. Chad, J. G. Frey, and W. S. Brocklesby, "Ptychographic coherent diffractive imaging with orthogonal probe relaxation," Opt. Express **24**(8), 8360 (2016).
22. I. A. Vartanyants and A. Singer, "Coherence properties of hard x-ray synchrotron sources and x-ray free-electron lasers," New J. Phys. **12**(3), 035004 (2010).
23. S. Lee, W. Roseker, C. Gutt, B. Fischer, H. Conrad, F. Lehmkühler, I. Steinke, D. Zhu, H. Lemke, M. Cammarata, D. M. Fritz, P. Wochner, M. Castro-Colin, S. O. Hruszkewycz, P. H. Fuoss, G. B. Stephenson, G. Grübel, and A. Robert, "Single shot speckle and coherence analysis of the hard x-ray free electron laser lcls," Opt. Express **21**(21), 24647–24664 (2013).
24. F. Lehmkühler, C. Gutt, B. Fischer, M. A. Schroer, M. Sikorski, S. Song, W. Roseker, J. Glownia, M. Chollet, S. Nelson, K. Tono, T. Katayama, M. Yabashi, T. Ishikawa, A. Robert, and G. Grübel, "Single shot coherence properties of the free-electron laser sacla in the hard x-ray regime," Sci. Rep. **4**(1), 5234 (2014).
25. A. Singer, I. A. Vartanyants, M. Kuhlmann, S. Duesterer, R. Treusch, and J. Feldhaus, "Transverse-Coherence Properties of the Free-Electron-Laser FLASH at DESY," Phys. Rev. Lett. **101**(25), 254801 (2008).
26. B. J. Daurer, S. Sala, M. F. Hantke, H. K. N. Reddy, J. Bielecki, Z. Shen, C. Nettelblad, M. Svenda, T. Ekeberg, G. A. Carini, P. Hart, T. Osipov, A. Aquila, N. D. Loh, F. R. N. C. Maia, and P. Thibault, "Ptychographic wavefront characterization for single-particle imaging at x-ray lasers," Optica **8**(4), 551–562 (2021).
27. S. Marchesini, "Invited Article: A unified evaluation of iterative projection algorithms for phase retrieval," Rev. Sci. Instrum. **78**(1), 011301 (2007).
28. A. Maiden, D. Johnson, and P. Li, "Further improvements to the ptychographical iterative engine," Optica **4**(7), 736–742 (2017).
29. T. J. Barth, D. E. Keyes, and D. Roose, *Automatic Differentiation: Applications, Theory, and Implementations*, vol. 50 of *Lecture Notes in Computational Science and Engineering* (Springer, Berlin, Heidelberg, 2006).
30. C. C. Margossian, "A review of automatic differentiation and its efficient implementation," WIREs Data Min. Knowl. Discov. **9**(4), 1–19 (2019).
31. A. S. Jurling and J. R. Fienup, "Applications of algorithmic differentiation to phase retrieval algorithms," J. Opt. Soc. Am. A **31**(7), 1348 (2014).
32. Y. S. Nashed, T. Peterka, J. Deng, and C. Jacobsen, "Distributed Automatic Differentiation for Ptychography," Procedia Comput. Sci. **108**, 404–414 (2017).
33. S. Kandel, S. Maddali, M. Allain, S. O. Hruszkewycz, C. Jacobsen, and Y. S. G. Nashed, "Using automatic differentiation as a general framework for ptychographic reconstruction," Opt. Express **27**(13), 18653 (2019).

Optics EXPRESS

34. S. Ghosh, Y. S. G. Nashed, O. Cossairt, and A. Katsaggelos, "ADP: Automatic differentiation ptychography," in *2018 IEEE International Conference on Computational Photography (ICCP)*, (IEEE, 2018), pp. 1–10.
35. M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light* (Elsevier, 2013).
36. P. Thibault and M. Guizar-Sicairos, "Maximum-likelihood refinement for coherent diffractive imaging," New J. Phys. **14**(6), 063004 (2012).
37. T. B. Edo, D. J. Batey, A. M. Maiden, C. Rau, U. Wagner, T. A. Waigh, and J. M. Rodenburg, "Sampling in x-ray ptychography," Phys. Rev. A **87**(5), 053850 (2013).
38. S. Marchesini, A. Schirotzek, C. Yang, H. tieng Wu, and F. Maia, "Augmented projections for ptychographic imaging," Inverse Probl. **29**(11), 115009 (2013).
39. A. Griewank and A. Walther, *Evaluating Derivatives* (Society for Industrial and Applied Mathematics, 2008).
40. M. Guizar-Sicairos and J. R. Fienup, "Phase retrieval with transverse translation diversity: a nonlinear optimization approach," Opt. Express **16**(10), 7264–7278 (2008).
41. M. Odstrčil, A. Menzel, and M. Guizar-Sicairos, "Iterative least-squares solver for generalized maximum-likelihood ptychography," Opt. Express **26**(3), 3108–3123 (2018).
42. M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu, "Spatial Transformer Networks," *Advances in Neural Information Processing Systems* pp. 2017–2025 (2015).
43. E. Wolf, "New Theory of Partial Coherence in the Space-Frequency Domain - 1. Spectra and Cross Spectra of Steady-State Sources," J. Opt. Soc. Am. **72**(3), 343–351 (1982).
44. P. Li, T. Edo, D. Batey, J. Rodenburg, and A. Maiden, "Breaking ambiguities in mixed state ptychography," Opt. Express **24**(8), 9038 (2016).
45. A. Chambolle, "An Algorithm for Total Variation Minimization and Applications," J. Math. Imaging Vis. **20**(1/2), 163–177 (2004).
46. D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings* pp. 1–15 (2014).
47. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, *PyTorch: An Imperative Style, High-Performance Deep Learning Library* (Curran Associates, Inc., 2019).
48. M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," ARXIV (2016).
49. M. Ruiz-Lopez, "Status of flash2 (conference presentation)," Proc. SPIE **11038**, 110380F (2019).
50. B. Keitel, E. Plönjes, S. Kreis, M. Kuhlmann, K. Tiedtke, T. Mey, B. Schäfer, and K. Mann, "Hartmann wavefront sensors and their application at flash," J. Synchrotron Radiat. **23**(1), 43–49 (2016).
51. S. Schreiber, B. Faatz, J. Feldhaus, K. Honkavaara, R. Treusch, and M. Vogt, "Status of the flash facility," *MOPD01, these proceedings* (2012).
52. J. C. da Silva and A. Menzel, "Elementary signals in ptychography," Opt. Express **23**(26), 33812–33821 (2015).
53. L. N. Smith, "Cyclical Learning Rates for Training Neural Networks," in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, (IEEE, 2017), pp. 464–472.
54. N. Banterle, K. H. Bui, E. A. Lemke, and M. Beck, "Fourier ring correlation as a resolution criterion for super-resolution microscopy," J. Struct. Biol. **183**(3), 363–367 (2013).
55. I. A. Vartanyants and A. Singer, "Coherence Properties of Third-Generation Synchrotron Sources and Free-Electron Lasers," in *Synchrotron Light Sources and Free-Electron Lasers*, (Springer International Publishing, Cham, 2017), pp. 1–38.
56. J. Vila-Comamala, A. Diaz, M. Guizar-Sicairos, A. Mantion, C. M. Kewish, A. Menzel, O. Bunk, and C. David, "Characterization of high-resolution diffractive X-ray optics by ptychographic coherent diffractive imaging," Opt. Express **19**(22), 21333 (2011).
57. H. Chang, P. Enfedaque, J. Zhang, J. Reinhardt, B. Enders, Y.-S. Yu, D. Shapiro, C. G. Schroer, T. Zeng, and S. Marchesini, "Advanced denoising for x-ray ptychography," Opt. Express **27**(8), 10395–10418 (2019).
58. D. J. Batey, D. Claus, and J. M. Rodenburg, "Information multiplexing in ptychography," Ultramicroscopy **138**, 13–21 (2014).
59. E. L. Saldin, E. A. Schneidmiller, and M. V. Yurkov, "Coherence properties of the radiation from x-ray free electron laser," Opt. Commun. **281**(5), 1179–1188 (2008).