Automatic Differentiation Near-Field Ptychography

Zhang Yiwei , DESY CFEL

September 4, 2024

HELMHOLTZ

DESY.

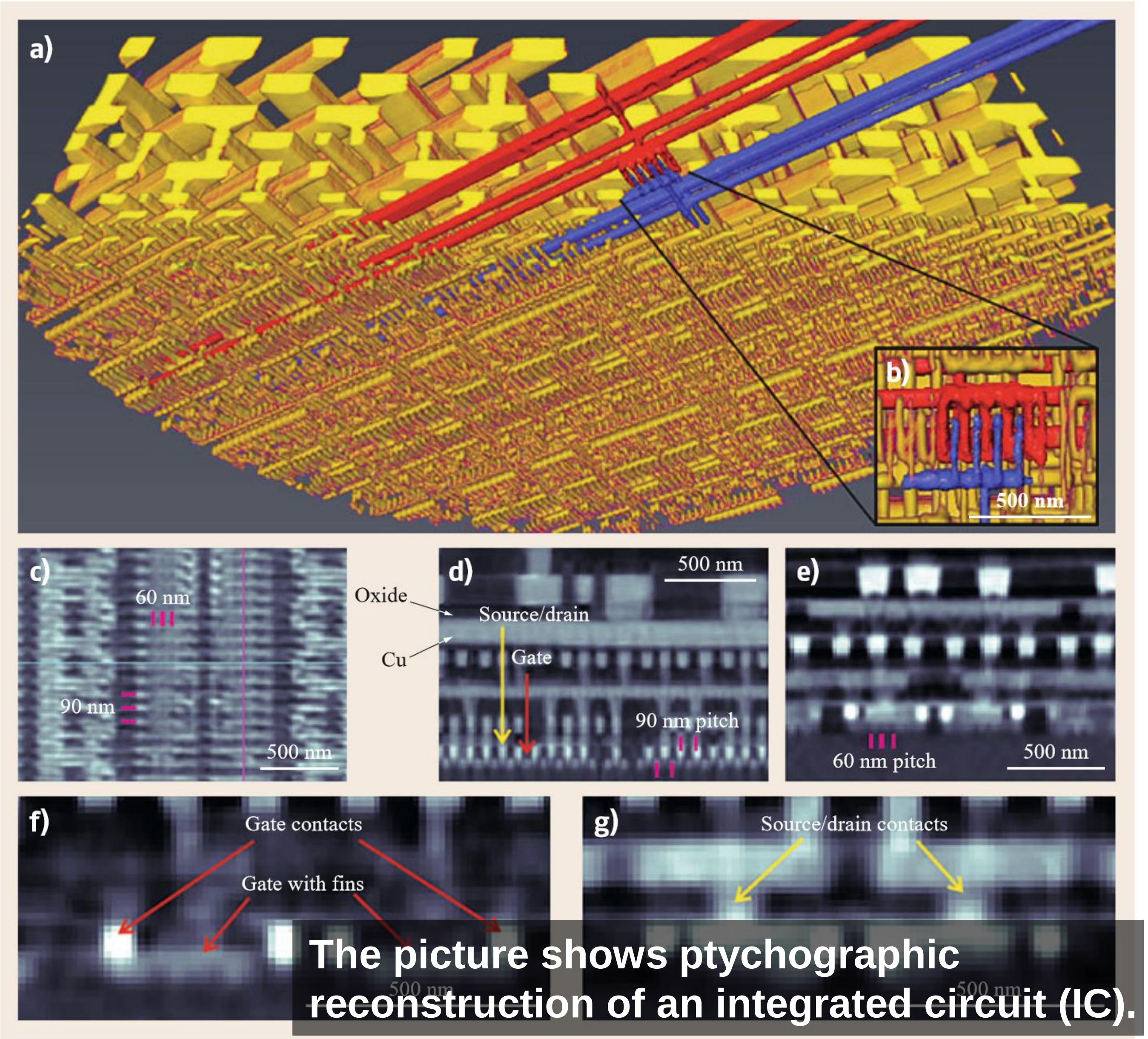
Table of Contents

Background

A high-resolution imaging method is needed

- Materials Science
- Biological Imaging
- Semiconductor Industry
- ……

Advancement of third generation synchrotrons like PETRA III excites a big interest on X-ray ptychography.



The picture shows ptychographic reconstruction of an integrated circuit (IC).

M. Holler, M. Guizar-Sicairos, E.H. Tsai, R. Dinapoli, E. Müller, O. Bunk, J. Raabe, G. Aeppli: High-resolution non-destructive three-dimensional imaging of integrated circuits, Nature 543(7645), 402–406 (2017)

Introduction — Challenges of traditional X-ray

X-ray: a high-resolution imaging method

- diffraction limit: $\frac{1.22\lambda}{D}$

- **higher resolution** than visible light imaging ($\lambda_{X-ray} \ll \lambda_{400-760nm}$)

Why is it challenging for X-ray to be imaged by lens?

→ **Ptychography**

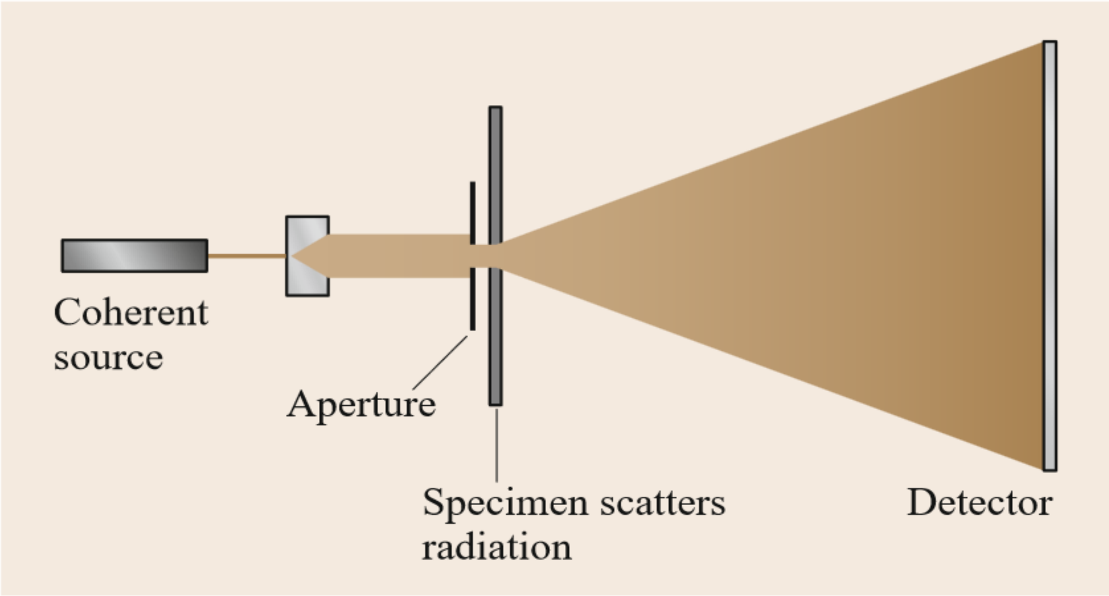
- high quality and high NA of lens is challenging to make

Ptychography

Greek word ‘ptycho’ means ‘fold’

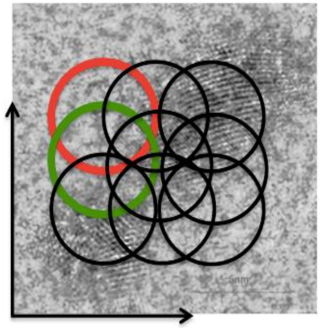
Principle :

1. Illumination probe scans over the specimen with sufficient overlap
2. Intensity of diffraction pattern is recorded
3. Use algorithm to reconstruct the sample and probe



Springer Handbook of Microscopy. (Springer International Publishing, Cham, 2019).

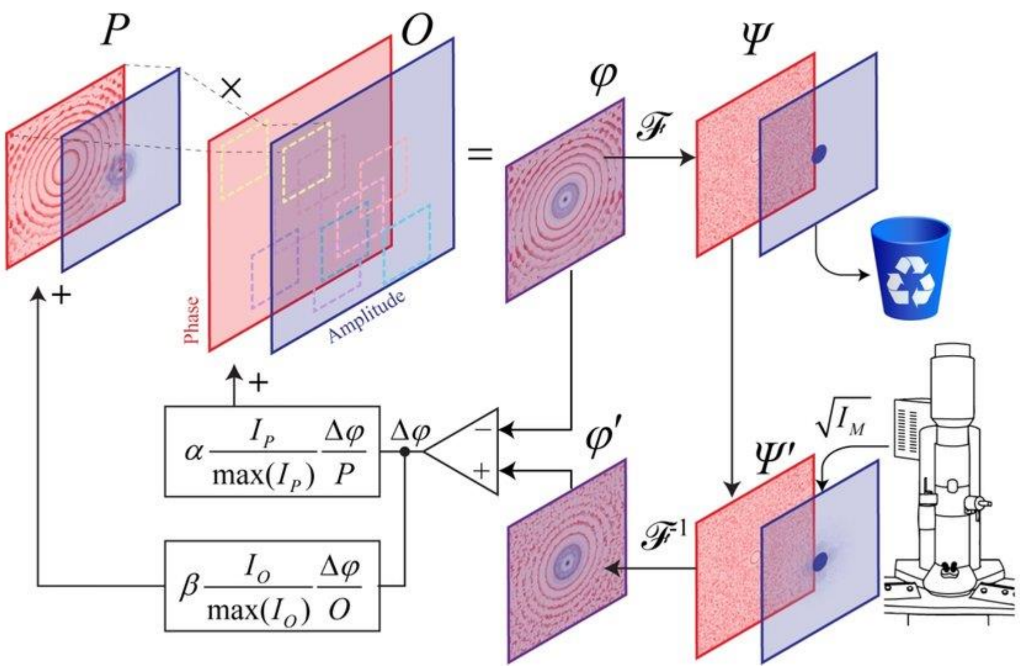
- **Sufficient Overlap in illumination between adjacent positions to recover the lost phase information**



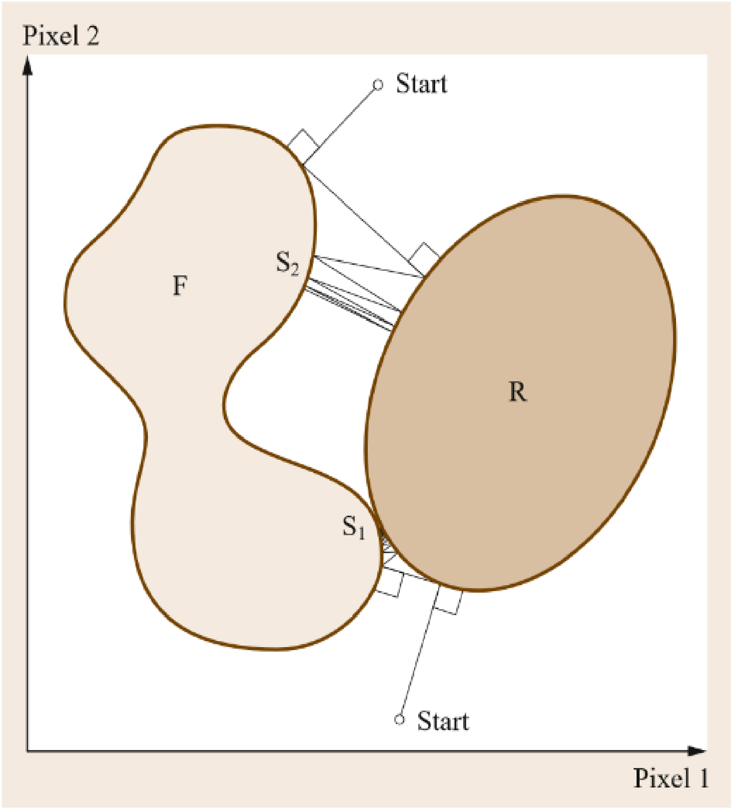
https://en.wikipedia.org/wiki/Ptychography

Conventional Algorithms

1、PIE family：mPIE, ePIE……



2、Alternating projections scheme: DM, RAAR



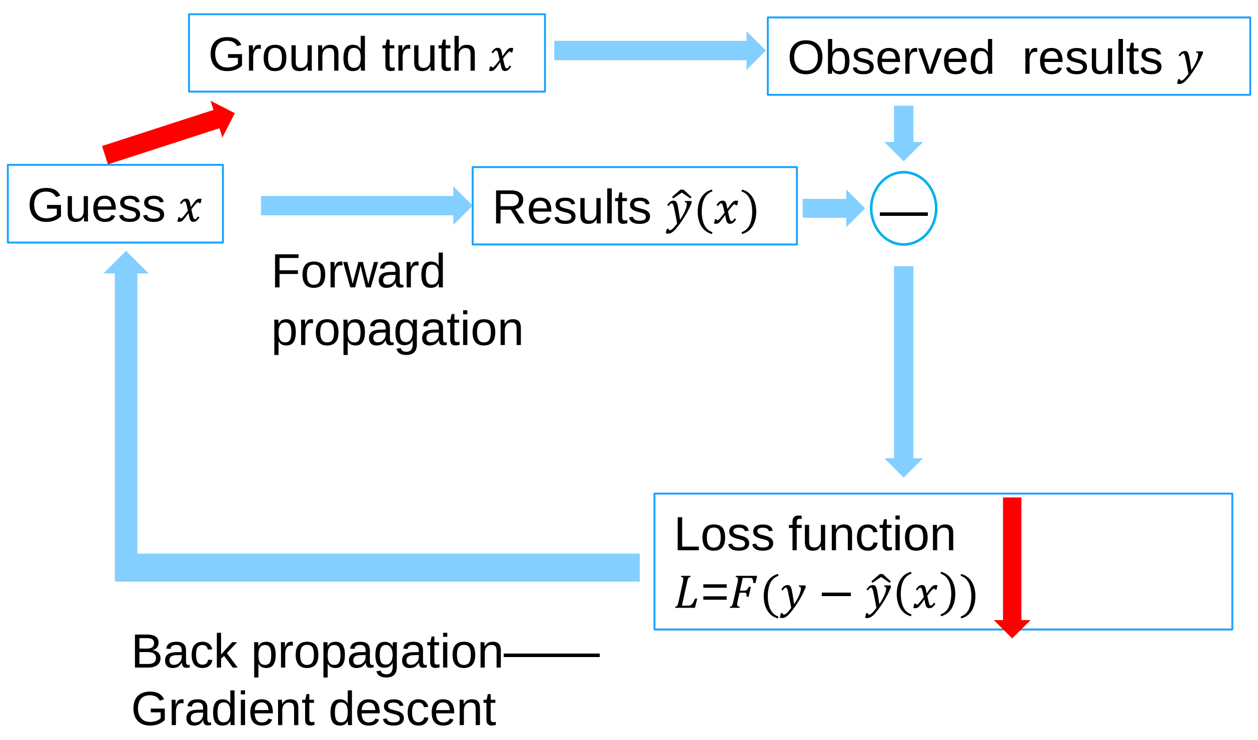
Springer Handbook of Microscopy. (Springer International Publishing, Cham, 2019).

Drawbacks Of Conventional Algorithms

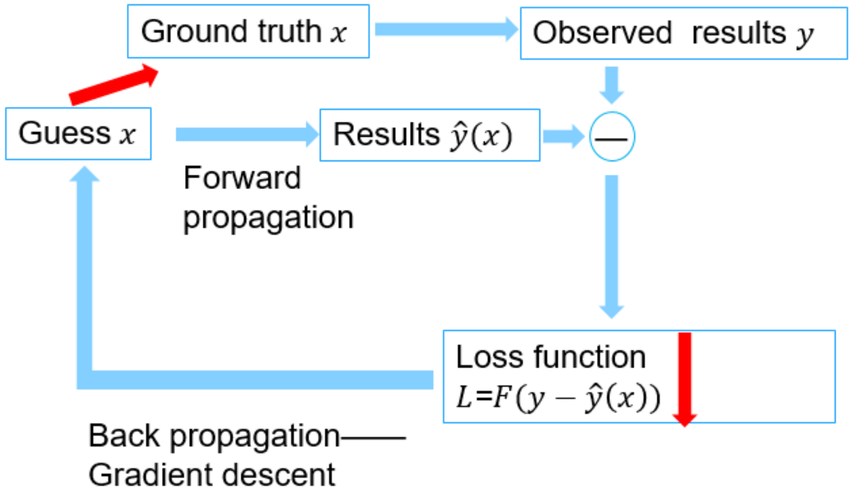
Inflexible:
Change of the experimental condition need manual re-derivation of the calculation strategy

Gradient Descent Method

Flexible



How does gradient descent actually work?

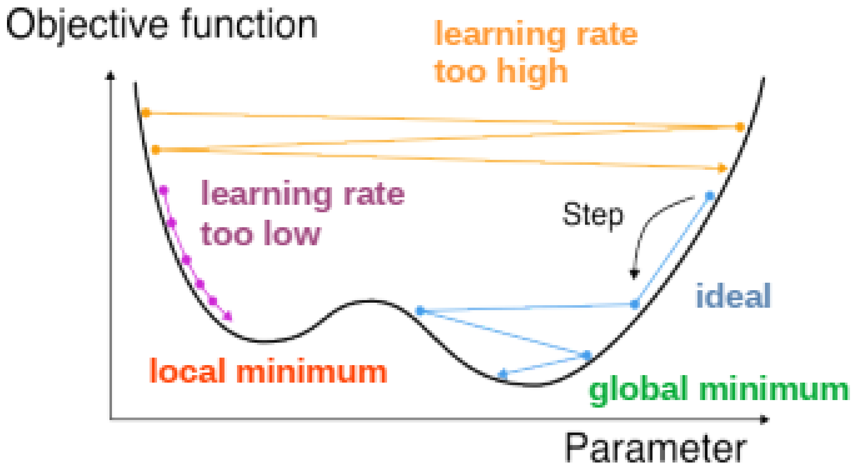


$$x_{new} = x_{old} - \boxed{\alpha} \cdot \nabla_x L(x)$$

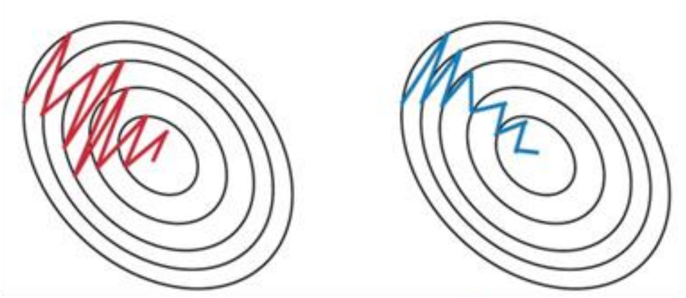
- update in the direction that reduces the loss
- Automatic differentiation to calculate $\nabla_x L(x)$

Chain rule:

$$\frac{\partial L(g_1(g_2(\ldots g_n(x)\ldots)))}{\partial x} = \frac{\partial L}{\partial g_n} \cdot \frac{\partial g_n}{\partial g_{n-1}} \cdots \frac{\partial g_1}{\partial x}$$



- Learning rate α: step size
- Advanced: adapt the learning rate with momentum ("ADAM")--More suitable for non-convex optimization problems



(a) Without momentum (b) With momentum

Framework of simulation

- Randomly initialize two matrices as object and probe, respectively

- Calculate exit wavefront:

$$\Phi_i(\boldsymbol{r}) = P(\boldsymbol{r} - \boldsymbol{r}_i) \cdot O(\boldsymbol{r})$$

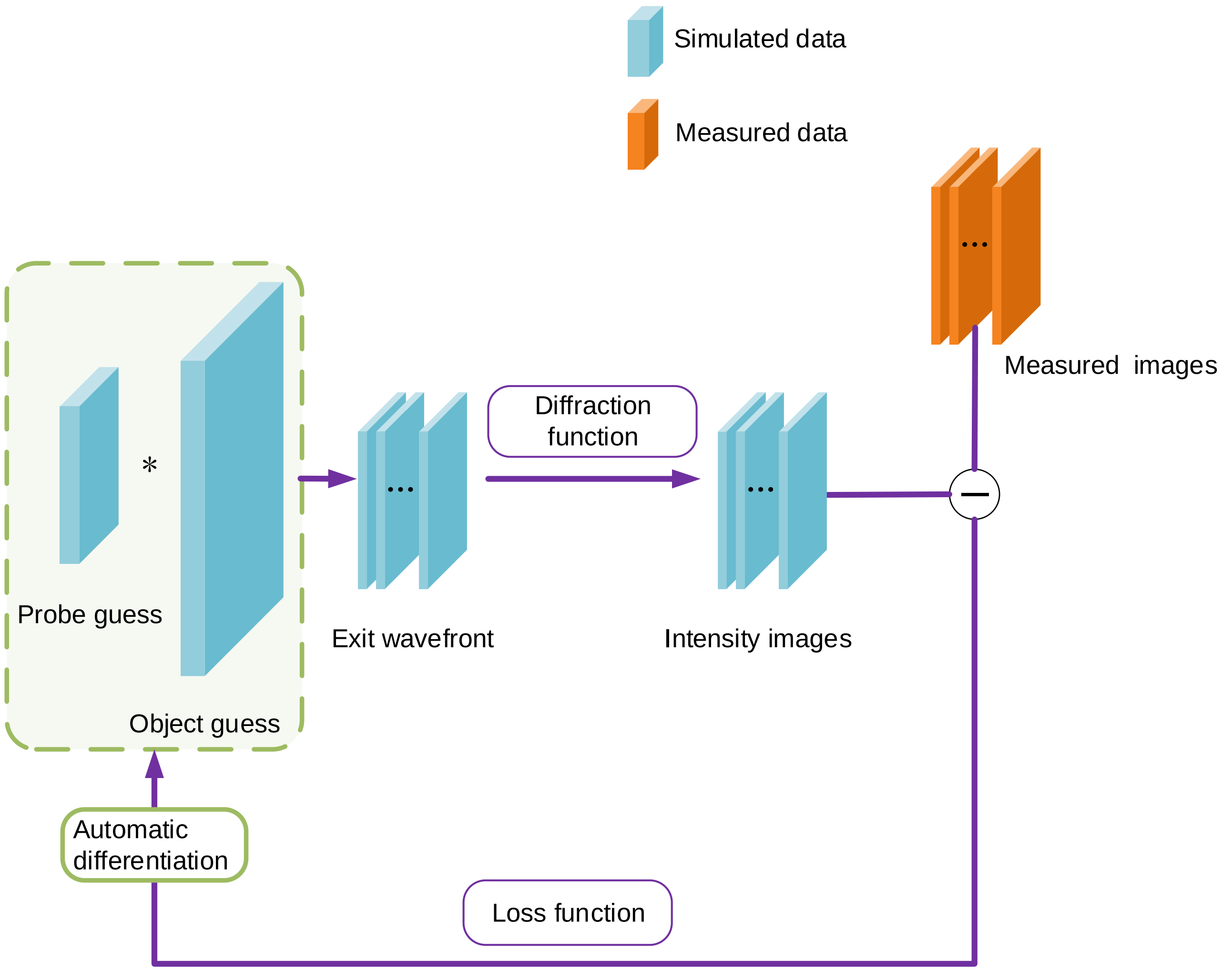
- Propagate the exit wavefront to get the intensity patterns

- $Loss\ function : MSE$ (mean squared error)

$$L(x) = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i(x))^2$$

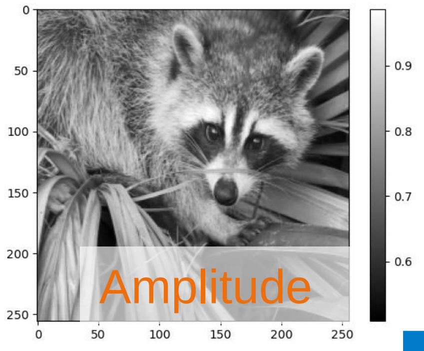
where y is amplitude of images

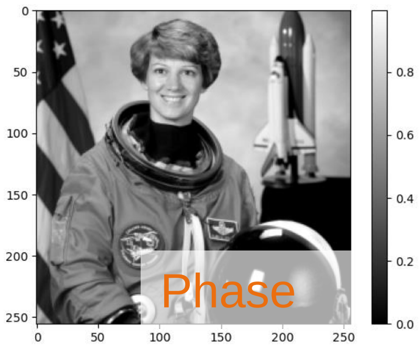
- Update: $P' = P - \alpha\frac{\partial L}{\partial P} \quad O' = O - \alpha\frac{\partial L}{\partial O}$

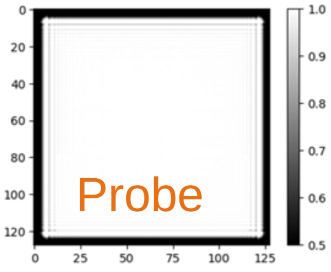


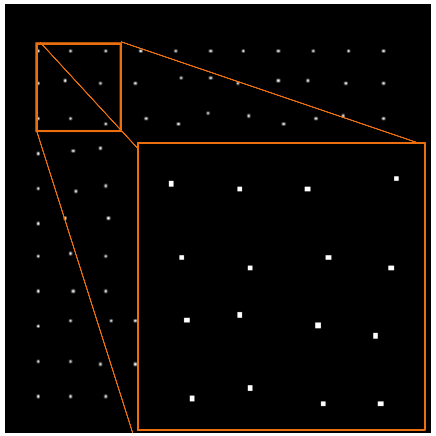
How to generate digital data as ground truth for testing?

Probe size(128×128)
Object size(256×256)
Probe overlap:90%



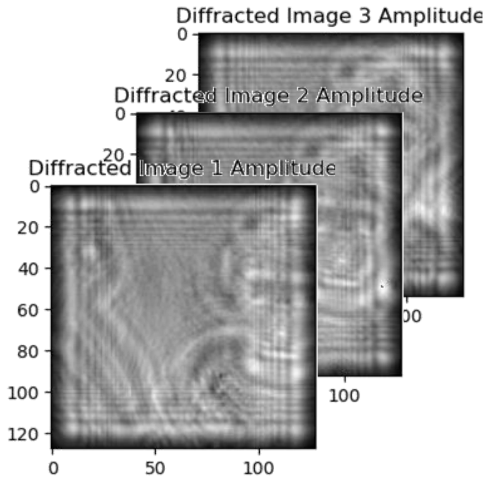






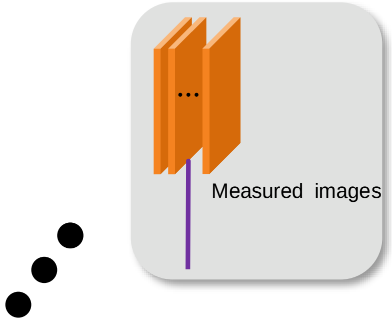
Scan positions with randomness

propagation



Amplitude

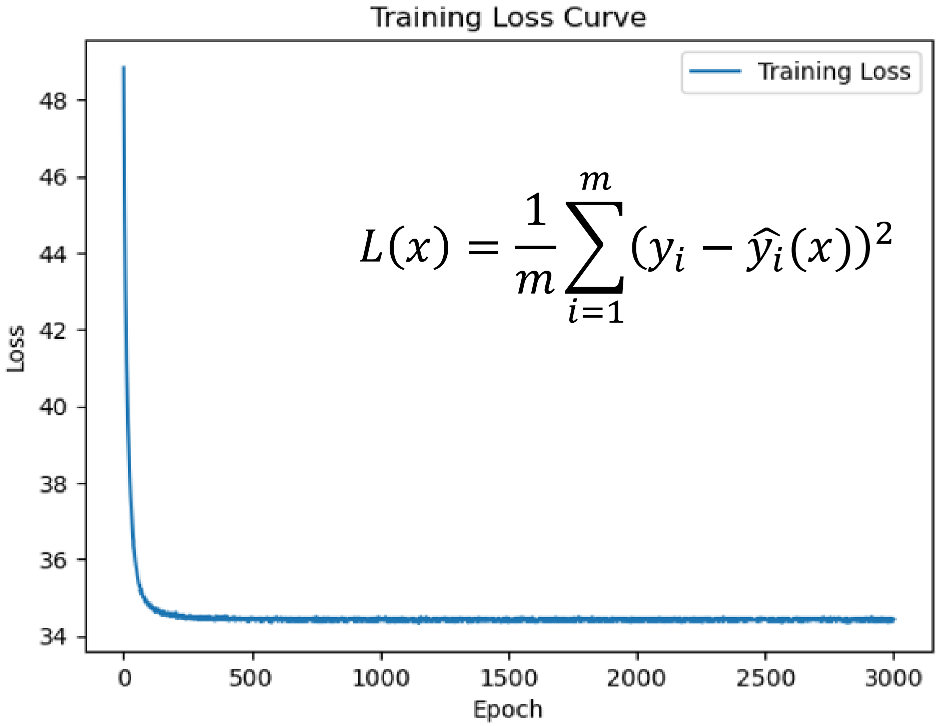
Diffraction images



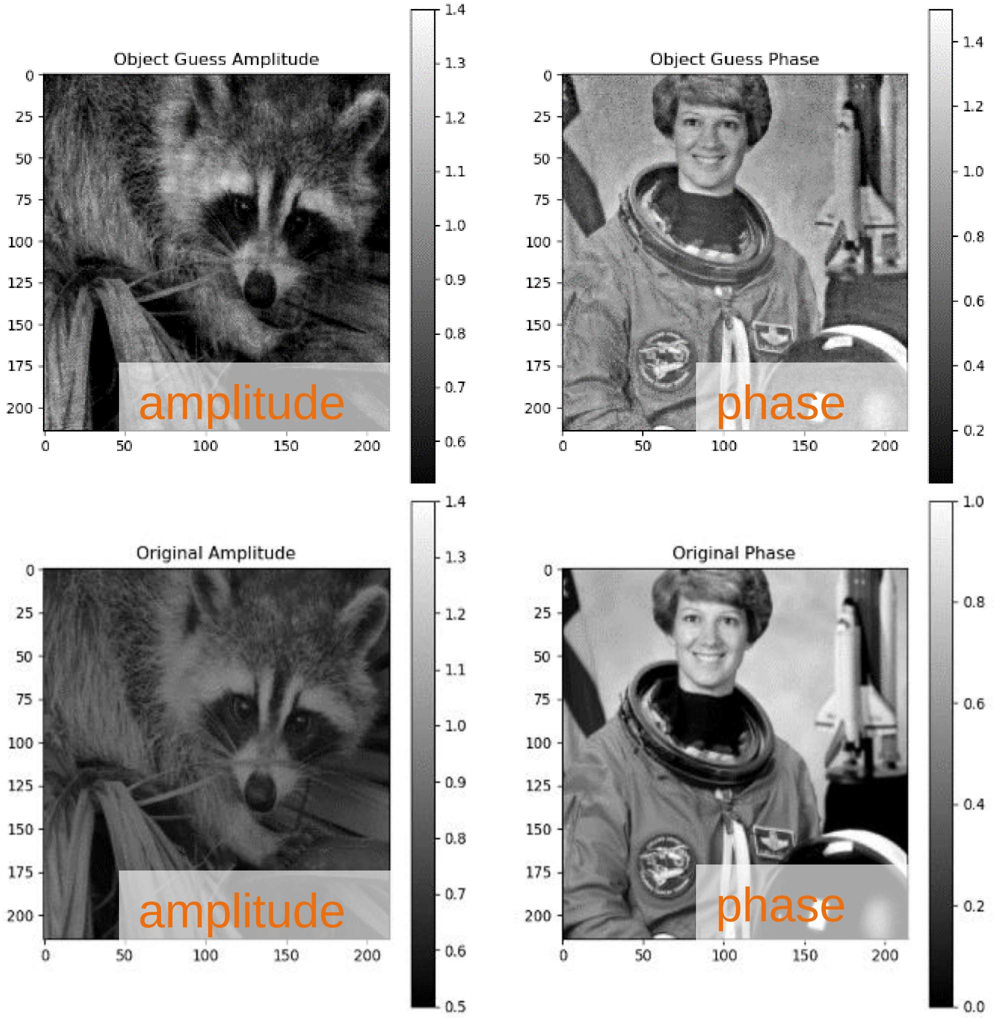
Result

Under high overlap conditions(90%),
the reconstruction quality is acceptable

Initial learning rate=0.01
Iteration = 3000



$$L(x) = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i(x))^2$$



Reconstructed object

Ground truth object

Result

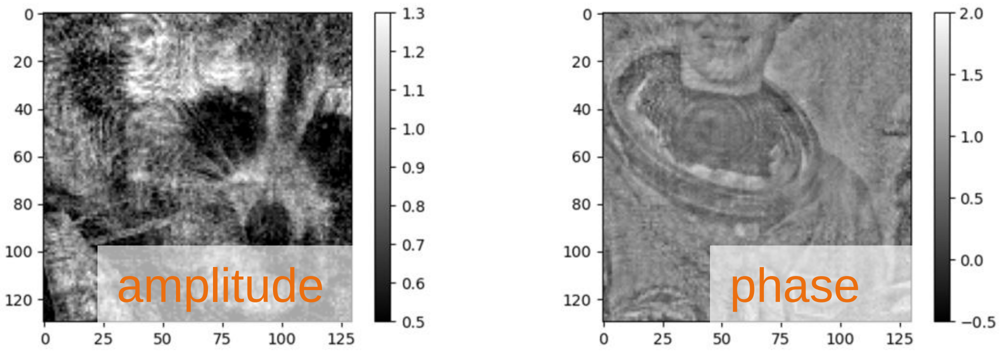
Initial learning rate=0.01
Iteration = 3000

Under low overlap conditions(70%), the reconstruction quality becomes worse

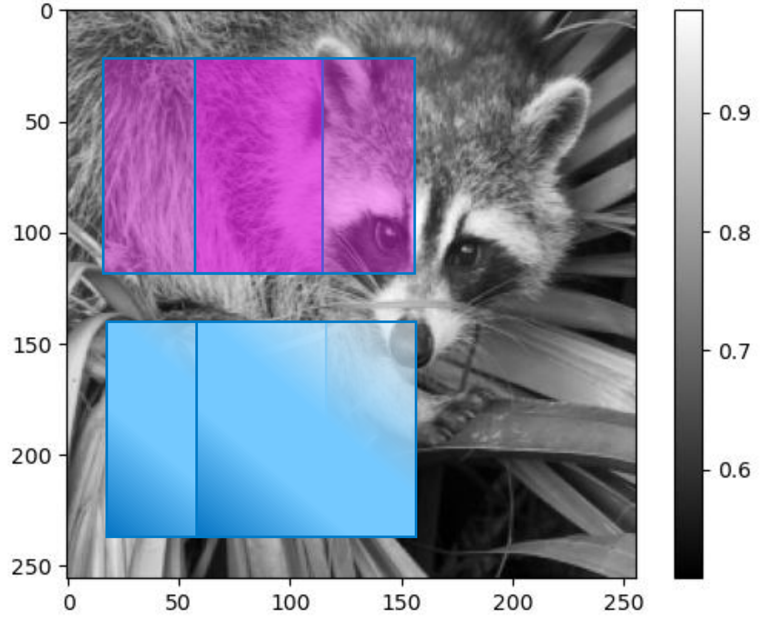
solution

Structured illumination – to add some diversity to the data

Reconstructed object



Why diversity is provided ?



$$Exit\ wavefront = O(r) * P(r - r_i)$$

The Fourier Shift Theorem:

$$F\{g(x-a, y-b)\} = G(f_x, f_y)\exp[-j2\pi(af_x + bf_y)]$$

The intensity:

$$|G(f_x, f_y)\exp[-j2\pi(af_x + bf_y)]|^2 = |G(f_x, f_y)|^2$$

If probe is normal as pink shows, for the overlapping region of object between two probe positions, the **intensities** of diffraction pattern have same information.

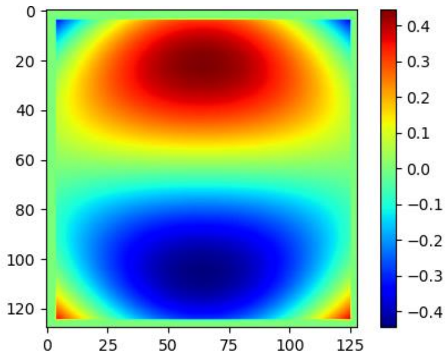
If probe is structured as blue shows, the overlapping region is illuminated by different area of probe. So it can provide more information and diversity.

Add some aberration

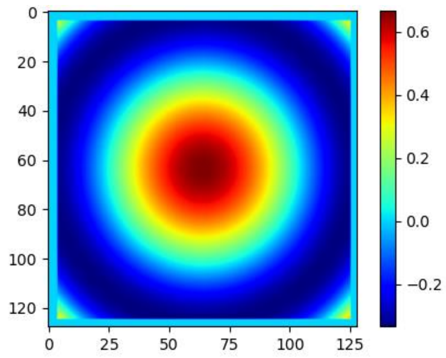
$$P = A \times \exp(i * \phi)$$

- Where P is the probe, A is the amplitude and ϕ is the phase
- Add some aberration to phase

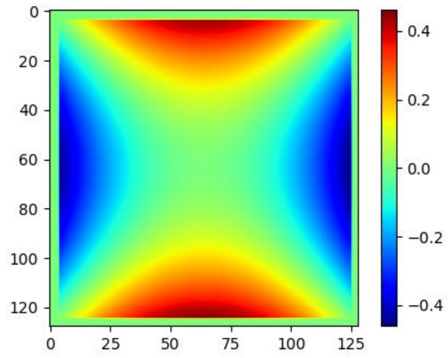
Coma



Spherical

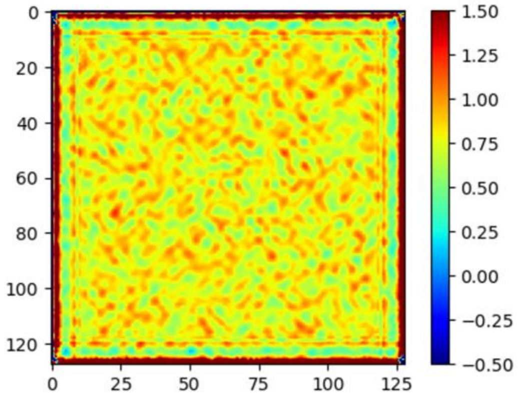


Astigmatism



- Try Zernike aberration but reconstruction doesn't become better

Diffuser



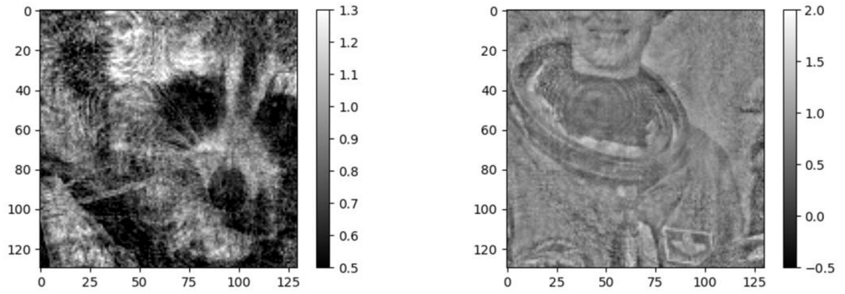
- Diffuser improves reconstruction

Improvement

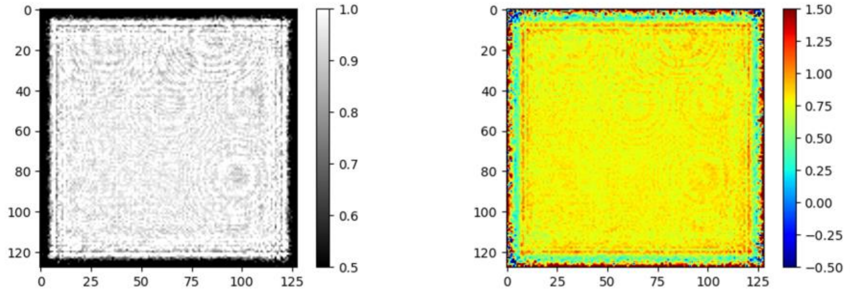
- The overlap rate is 70%

The improvement is quite significant

Without diffuser

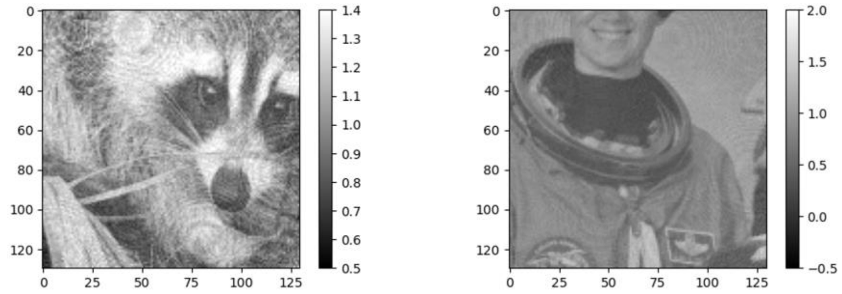


Reconstructed object

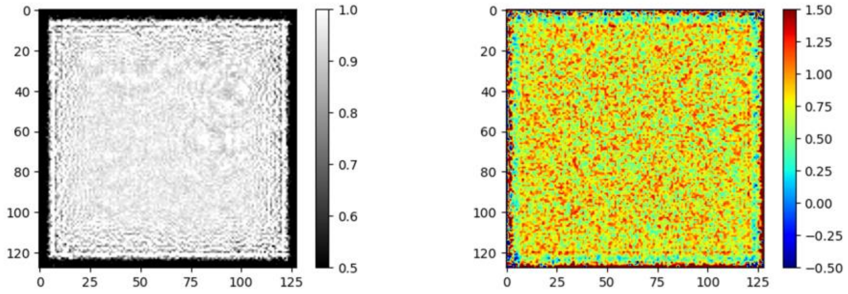


Reconstructed probe

With diffuser



Reconstructed object



Reconstructed probe

Aberration with different spatial frequency components

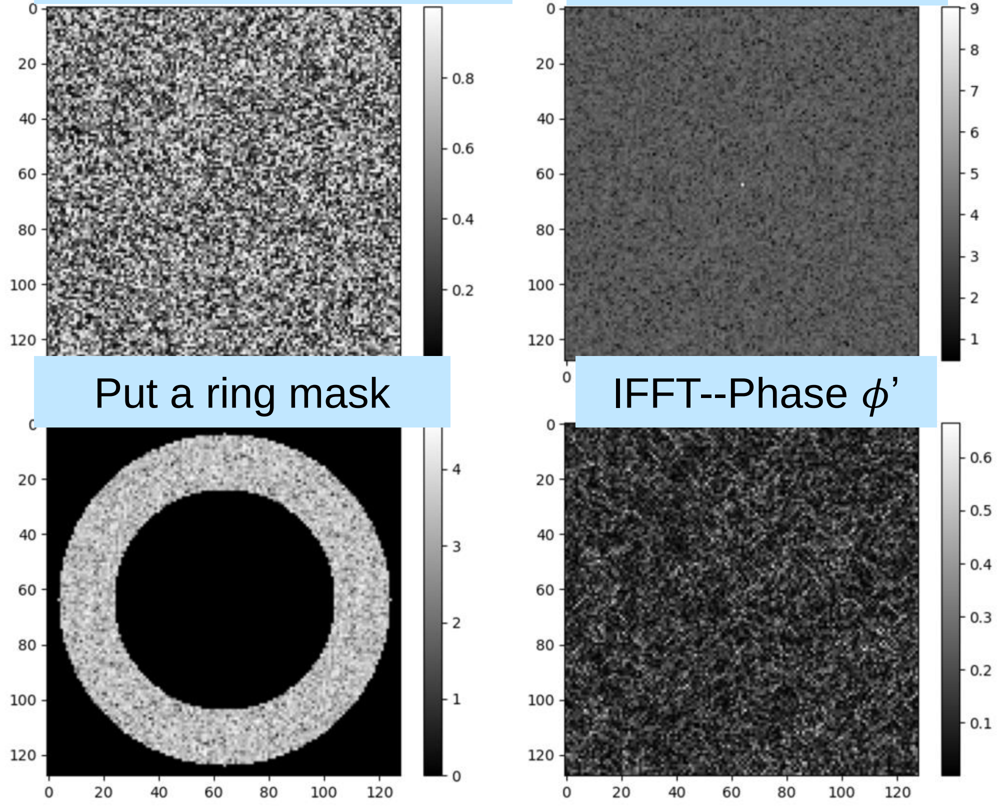
Put a ring mask to select
specific frequency components

Phase ϕ of the diffuser

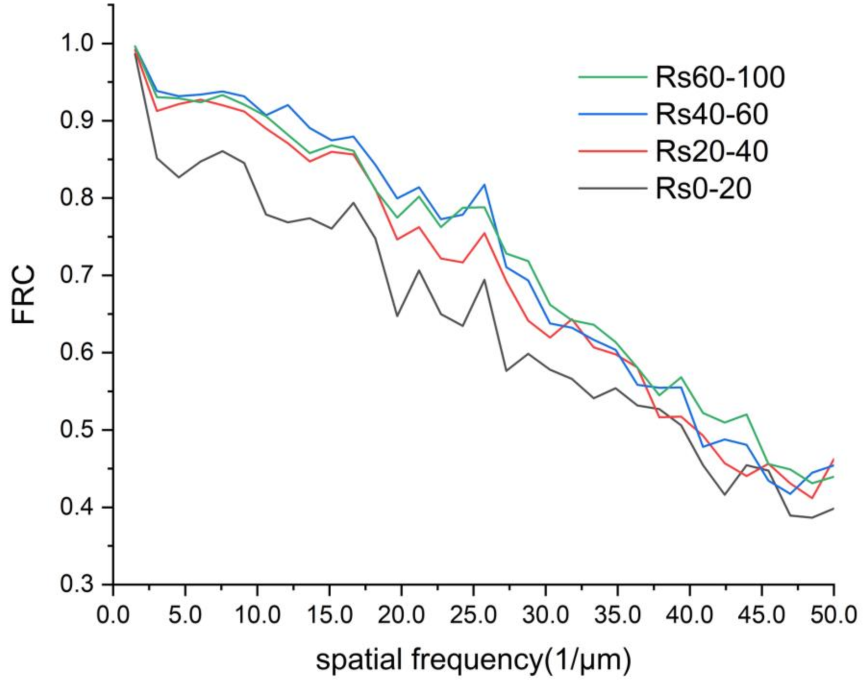
FFT of ϕ

Put a ring mask

IFFT--Phase ϕ'



Draw FRC(fourier ring correlation curve)
of different radius of ring mask



The results are improved with more high spatial frequency components.

Summary

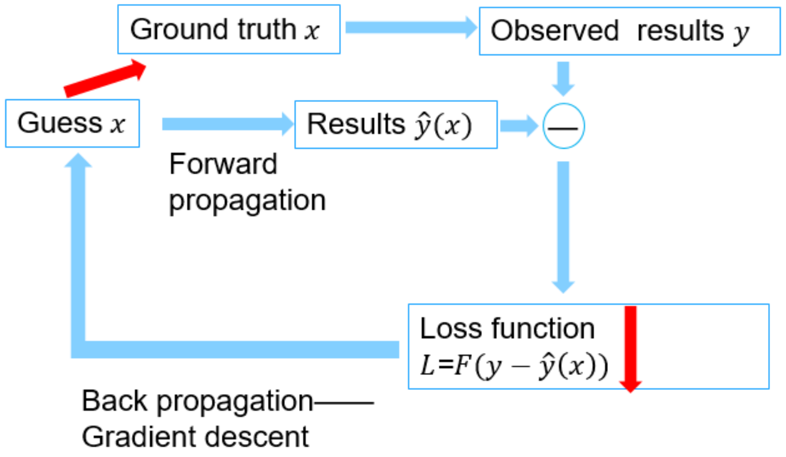
Ptychography:
- **Conventional Algorithms: PIE\projection**

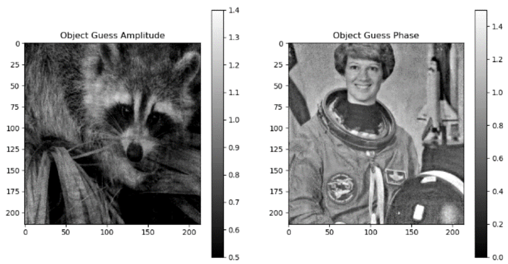
Gradient Descent:
- **Update in the opposite direction of gradient**
- **Use AD to calculate gradient**

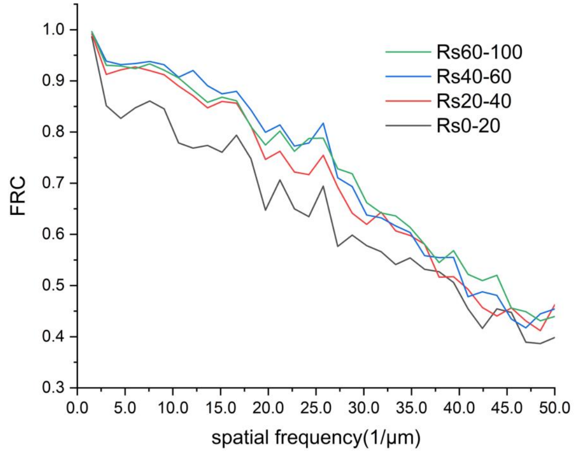
Simulation result:
- **Work with high overlap rate**
- **Bad when overlap rate is low**

Structured illumination :
- **diversity**
- **higher frequency plays an important role**

More flexible and adaptive







Thank you

Contact

Deutsches Elektronen-
Synchrotron DESY

www.desy.de

Zhang Yiwei

yiwei.zhang@desy.de