The CRLB and Maximum likelihood in ptychography with Poisson noise

XUKANG WEI,1,* H. PAUL URBACH1
1Optics Research Group, Imaging Physics Department, Delft University of Technology, The Netherlands
*x.x.wei-2@tudelft.nl

Abstract: We investigate the performance of ptychography with noisy data by analyzing the Cramér Rao Lower Bound. The lower bound of ptychography is derived and numerically computed for both plane wave and structured illumination. The influence of Poisson noise on the ptychography reconstruction is discussed. The computation result shows that, if the estimator is unbiased, the minimum variance for Poisson noise is mostly determined by the probe function. Monte Carlo analysis is conducted to validate our calculation results for different signal-to-noise ratios. The performance of the maximum likelihood method and the approach of amplitude-based cost function minimization is studied in the Monte Carlo analysis also.

1. Introduction

Ptychography [1–6] is a scanning coherent diffraction imaging method which aims to reconstruct a complex valued object function from intensity measurements recorded in the Fraunhofer or Fresnel diffraction region. In ptychography the object is partially illuminated multiple times so that the entire object is covered and adjacent illuminations are partially overlapped [7]. The technique is found very suitable for EUV [8, 9] and X-ray imaging applications [10–13] due to its high fidelity and its minimum requirement on optical imaging elements. Moreover, abundant studies show that ptychography is able to provide a wide field-of-view and retrieve the illumination probe also [14, 15]. During the last two decades, ptychography has been successfully demonstrated with X-ray radiation sources [11, 16, 17], electron beams [18] and visible light sources [19]. More recently, many extensions of ptychography have been proposed, including Fourier ptychography [20–22], spatially partial coherent ptychography [23–25], broadband ptychography [26, 27], 3D ptychography [28–30], on-the-fly scanning ptychography [31, 32] and interference probe ptychography [33].

For retrieving the object from a ptychographic data set, the key is to find a solution which fulfills both the ptychographic illumination condition in real space and the corresponding measured diffraction intensities in reciprocal space. A commonly used approach for solving the problem is the ptychography iterative engine [5, 15], which can be derived by sequentially minimizing the distance between the estimated amplitude of the diffracted wavefield and the measurements [6]. Another popular choice is the difference map algorithm, which can be formulated in terms of finding the intersection of two constraint sets [14, 34]. Based on the augmented Lagrangian methods for solving the conventional constrained optimization problems, several interesting ptychographic algorithm have been developed during the past ten years [35–38].

However, obtaining an unique reconstruction and a reconstruction with minimum defect in ptychography is considered difficult and there is still room for improvement. On the one hand, ambiguities due to a constant scaling factor, a global phase shift and raster grid pathology, occur in particular when the probe is undetermined [39]. Although many algorithms have been presented to enhance the robustness of ptychography [19, 35, 37, 38], a good starting point and proper parameter settings (e.g. update step size, regularization factor, etc.) are needed in general. Furthermore, noise in the measurements of the diffracted intensity cause inaccuracies in the reconstructions [40–42]. To prevent the effect caused by the saturation of the detector, dark-field and near-field ptychography have been introduced [43, 44]. Moreover, it was shown that adaptive step size strategies are able to improve the performance of ptychography in the
presence of noise [19, 45]. In general, the most powerful and robust de-noising methods are
based on the maximum likelihood principle [21, 40–42, 46]. The likelihood function used in the
maximum likelihood method depends on the noise model. Common choices for the noise model
in ptychography are Poisson noise, Gaussian noise and the mixed Poisson-Gaussian model. It
has been demonstrated [22, 40, 41, 47] that, by using the variance stabilization transform given
by Bartlett [48] and Anscombe [49], one can approximate the maximum likelihood method of
Poisson noise by the amplitude-based cost minimization algorithm. Therefore both the approach
of maximum likelihood and the amplitude-based cost minimization algorithm can be used as a
refinement method in ptychography with noisy data.

In this work we investigate the Cramér Rao Lower Bound (CRLB) for the variance of any
unbiased estimator in ptychography [50–52]. To the best of our knowledge, this is the first
investigation of the CRLB in ptychography. We begin by deriving the lower bound for Poisson
distributed photon counting noise, which is the most dominant source of noise which occurs even
under the best experimental conditions [40, 41]. In Section 2, we briefly discuss ptychography,
Poisson photon counting noise and the maximum likelihood method. We compute the Fisher
information matrix of ptychography with Poisson noise and introduce the CRLB. In Section 3,
the CRLB is computed for plane-wave and structured illumination. To validate this obtained
CRLB, Monte Carlo analysis is implemented in Section 4. The performance of the maximum
likelihood method and the approach of amplitude-based cost function minimization are also
compared. The paper is concluded with a summary and outlook in the last section.

2. Theory

2.1. Ptychography, Poisson noise, and maximum likelihood method

The goal of ptychography is to reconstruct a complex-valued object \( O \) from a set of diffraction
intensity patterns which are recorded in the Fraunhofer or Fresnel region. Let \( r \) and \( r' \) be 2D
coordinates in the object plane and the detector plane, respectively. The exit wave immediately
behind the object is denoted by \( \Psi(r) \) and the measured diffraction intensity measurement
\( I(r') \). According to the thin object model, the exit wave \( \Psi(r) \) for an illumination with a probe function
\( P(r) \) which is centered on position \( R_m \) is given by

\[
\Psi_m(r) = P(r - R_m) \cdot O(r) = P_m(r) \cdot O(r).
\]

The probe function is assumed to have a support with circular boundary:

\[
P(r) = \begin{cases} 
P(r), & |r| \leq r_0, \\
0, & |r| > r_0.
\end{cases}
\]

For a detector located at distance \( z \) in the far field, the diffraction intensity pattern \( I(r') \) for the
\( m \)th illumination is [53]:

\[
I_m(r') = \left| \mathcal{F} \left( \Psi_m \right) \left( \frac{r'}{\lambda z} \right) \right|^2 \\
= \left| \iint \Psi_m(r) \cdot \exp \left( -i \frac{2\pi}{\lambda z} r \cdot r' \right) \, dr \right|^2.
\]

where \( \mathcal{F} \) is the Fourier transform operator. From \( I_m(r') \), one can estimate the average number of
measured photons:

\[
N_m(r') = \frac{I_m(r')}{\hbar \omega}, \quad \text{where} \quad \omega = \frac{2\pi c}{\lambda}.
\]

The cost function $E$ is defined as the $l_2$-distance between the square roots of the estimated average photon number $N(r')$ and of the number of measured photons $n(r')$:

$$E = \frac{1}{\hbar \omega} \sum_{R_m} \iint \left[ \sqrt{n_m(r')} - \sqrt{N_m(r')} \right]^2 \, dr'$$

$$= \sum_{R_m} \iint \left[ \sqrt{\frac{n_m(r')}{\hbar \omega}} - \sqrt{I_m(r')} \right]^2 \, dr'.$$

(5)

The task of ptychography is to find an object function which takes account of the a priori knowledge, which in our case are the support constraint on the probe function, and the $m$th relative position $R_m$ between the probe and the object, while the cost function $E$ is minimized.

Among various of noise models, we consider Poisson noise. For Poisson noise, the probability distribution is given by:

$$P_p = \prod_{R_m} \prod_{r'} \frac{N_m(r')^{n_m(r')}}{n_m(r')!} e^{-N_m(r')},$$

where the cumulative product is over both the 2-D coordinate $r'$ and the probe position $R_m$. The likelihood functional is defined by:

$$L_p = \ln P_p = - \sum_{R_m} \sum_{r'} \left[ n_m(r') \ln(N_m(r')) - N_m(r') - \ln(n_m(r'))! \right].$$

(7)

Hence, the object function $O(r)$ which maximizes the likelihood functional satisfies for all $\delta O$:

$$\delta L_p(\delta O) = \sum_{R_m} \frac{2}{\hbar \omega} \iint \left( \frac{n_m(r')}{N_m(r')} - 1 \right) \cdot \delta I_m(\delta O) \, dr'$$

$$= \sum_{R_m} \frac{2}{\hbar \omega} \iint \left( n_m(r') \right) \Re \left\{ \mathcal{F} \left( P_m O \right)(r') \cdot \mathcal{F} \left( P_m \delta O \right)^* (r') \right\} \, dr'$$

$$= \sum_{R_m} \frac{2}{\hbar \omega} \cdot \Re \left\{ \int \mathcal{F}^{-1} \left[ \frac{n_m(r')}{N_m(r')} - 1 \right] \cdot \mathcal{F} \left( P_m O \right)(r') \right\} P_m^*(r) \cdot \delta O^*(r) \, dr$$

$$= 0,$$

(8)

where $\Re$ denotes the real part. The solution of Eq. (8) can be found by the method of steepest descent [27, 46, 54]:

$$O_{k+1}(r) = O_k(r) + \alpha_O \sum_{R_m} P_m^* \cdot \mathcal{F}^{-1} \left[ \frac{n_m}{N_m} - 1 \right] \cdot \mathcal{F} \left( P_m \cdot O \right)(r),$$

(9)

where $k$ is the iteration number, and $\alpha_O$ is the step-size, which is normally chosen to be a constant, i.e. it is independent on the iteration number. Alternatively, projection based method or conjugate gradient method can be applied to achieve maximum likelihood [40].

2.2. The CRLB and the Fisher matrix

In estimation theory, the CRLB gives a lower bound on the variance of any unbiased estimator for the parameter which need to be estimated. The estimators that can reach the lower bound are called the minimum variance unbiased estimators. Minimum variance unbiased estimators are often not available.
We recall the definition of the CRLB, using the notation as in [50]. Suppose we have a collection of measurements \( \mathbf{X} = [X_1, X_2, \ldots]^T \), and a set of real parameters \( \Theta = [\theta_1, \theta_2, \ldots]^T \) that have to be retrieved. To determine the lower bound on the variance of estimator \( \hat{\Theta} \), one computes the Fisher information matrix \( I_F \), given by:

\[
I_F (\Theta) = -E \left[ \frac{\partial^2 \ln \mathcal{P}(\mathbf{X}; \Theta)}{\partial \Theta^2} \right],
\]

where \( \mathcal{P}(\mathbf{X}; \Theta) \) is the probability distribution function and \( E \) is the expectation operator. Hence, the element \( i, j \) of \( I_F (\Theta) \) is given by:

\[
I_{F} (\Theta)_{ij} = -E \left[ \frac{\partial^2 \ln \mathcal{P}(\mathbf{X}; \Theta)}{\partial \theta_i \partial \theta_j} \right] = -\sum_{\mathbf{X}} \frac{\partial^2 \ln \mathcal{P}(\mathbf{X}; \Theta)}{\partial \theta_i \partial \theta_j} \mathcal{P}(\mathbf{X}; \Theta).
\]

Note that the sum is over all possible outcomes of all measurements. The CRLB is then given by the diagonal elements of the inverse of matrix \( I_F \), i.e.

\[
\text{Var} \left( \hat{\theta}_i \right) \geq \left[ I_F^{-1} (\Theta) \right]_{ii},
\]

where \( \text{Var} \left( \hat{\theta}_i \right) \) stands for the variance of estimator \( \hat{\theta}_i \) for the unknown parameter \( \theta_i \).

It is important to note that the estimator based on the maximum likelihood principle \( \hat{\Theta}_{ML} \) asymptotically becomes unbiased and achieves the CRLB for large data sets [50], that is:

\[
\hat{\Theta}_{ML} \overset{a}{\sim} \mathcal{N} \left( \Theta, \text{diag} \left[ I_F^{-1} (\Theta) \right] \right),
\]

where \( \mathcal{N} \) stands for the normal distribution and \( \text{diag} \) takes the diagonal elements of a matrix.

### 2.3. The Fisher matrix with Poisson noise in ptychography

To find the Fisher information matrix, we start computing the second order derivative of the likelihood functional \( \mathcal{L}_P \) with respect to \( O(\mathbf{r}) \):

\[
\delta^2 \mathcal{L}_P (\delta O) (\delta \tilde{O}) = -\frac{1}{(\hbar \omega)^2} \sum_{\mathbf{r}_m} \int \int E \left\{ \frac{n_m(\mathbf{r}')}{N_m^2(\mathbf{r}')} \left[ \delta I_m (\delta O) \right] [\delta I_m (\delta \tilde{O})] \right\} d\mathbf{r}'
\]

\[
+ \frac{1}{(\hbar \omega)^2} \sum_{\mathbf{r}_m} \int \int \left( \frac{n_m(\mathbf{r}')}{N_m(\mathbf{r}')} - 1 \right) \delta^2 I_m (\delta O) (\delta \tilde{O}) d\mathbf{r}'.
\]

where \( \delta O \) and \( \delta \tilde{O} \) are perturbations of \( O \). Taking the expectation of Eq. (14), we have:

\[
E \left( \delta^2 \mathcal{L}_P \right) (\delta O) (\delta \tilde{O}) = \frac{1}{(\hbar \omega)^2} \sum_{\mathbf{r}_m} \int \int E \left\{ \frac{n_m(\mathbf{r}')}{N_m^2(\mathbf{r}')} [\delta I_m (\delta O)] \left[ \delta I_m (\delta \tilde{O}) \right] \right\} d\mathbf{r}'
\]

\[
- \frac{1}{(\hbar \omega)^2} \sum_{\mathbf{r}_m} \int \int E \left\{ \left( \frac{n_m(\mathbf{r}')}{N_m(\mathbf{r}')} - 1 \right) \delta^2 I_m (\delta O) (\delta \tilde{O}) \right\} d\mathbf{r}',
\]

in which we commute the expectation and summation because the measurements \( n_m(\mathbf{r}') \) are independent photon measurements. Using the properties of the Poisson distribution [52]:

\[
\begin{align*}
\sum_{n_m} n_m! \frac{n_m}{n_m} e^{-N_m} &= 1, \\
\sum_{n_m} n_m \frac{n_m}{n_m} e^{-N_m} &= N_m.
\end{align*}
\]
and using Eq. (8), we find:

\[
E \left( \delta^2 L_p \right) (\delta O) (\delta \bar{O}) = \frac{1}{(\hbar \omega)^2} \sum_{\text{R}_m} \int \int \frac{1}{N_m(r')} \delta I_m(\delta O) \delta I_m(\delta \bar{O}) \, dr'
\]

\[
= \frac{4}{\hbar \omega} \sum_{\text{R}_m} \int \int \frac{1}{I_m(r')} \Re \left[ \mathcal{F}(\Psi_m) \mathcal{F}(P_m \delta O) \right] \Re \left[ \mathcal{F}(\Psi_m) \mathcal{F}(P_m \delta \bar{O})^* \right] \, dr'
\]

\[
= 2 \frac{\hbar \omega}{\hbar \omega} \sum_{\text{R}_m} \Re \left\{ \int \frac{\mathcal{F}(\Psi_m)(r')^2}{I_m(r')} \mathcal{F}(P_m \delta O) \mathcal{F}(P_m \delta \bar{O})^* \, dr' \right\}
\]

\[
+ 2 \frac{\hbar \omega}{\hbar \omega} \sum_{\text{R}_m} \Re \left\{ \int \frac{|\mathcal{F}(\Psi_m)(r')|^2}{I_m(r')} \mathcal{F}(P_m \delta O) \mathcal{F}(P_m \delta \bar{O})^* \, dr' \right\}.
\]

(17)

If we use a grid to approximate the perturbations \(\delta O\) and \(\delta \bar{O}\) by vectors, and rearrange these complex-value vectors by their real and imaginary parts:

\[
\delta O(r) = \begin{bmatrix} \delta O_R(r_i) \\ i \delta O_I(r_i) \end{bmatrix} \delta (r - r_i), \quad \delta \bar{O}(r) = \begin{bmatrix} \delta O_R(r_j) \\ i \delta O_I(r_j) \end{bmatrix} \delta (r - r_j).
\]

(18)

we obtain the Fisher information matrix with elements:

\[
I_{F,ij} = \begin{bmatrix} (I_F)_{RR,ij} & (I_F)_{RI,ij} \\ (I_F)_{IR,ij} & (I_F)_{II,ij} \end{bmatrix}
\]

\[
= \frac{2}{\hbar \omega} \sum_{\text{R}_m} \Re \begin{bmatrix} f_m(r_i, r_j) \\ f_m(r_i, r_j) \end{bmatrix} \Im \begin{bmatrix} f_m(r_i, r_j) \\ f_m(r_i, r_j) \end{bmatrix} + \frac{2}{\hbar \omega} \sum_{\text{R}_m} \begin{bmatrix} |P_m(r_i)|^2 \delta_{ij} \\ 0 \end{bmatrix}
\]

\[
\begin{bmatrix} 0 \\ |P_m(r_i)|^2 \delta_{ij} \end{bmatrix},
\]

(19)

where we used Eq. (3) and Kronecker’s symbol \(\delta_{ij}\). The auxiliary function \(f\) is given by:

\[
f_m(r_i, r_j) = \mathcal{F}^{-1} \left[ \frac{\mathcal{F}(\Psi_m) \mathcal{F}(P_m \delta O)}{\mathcal{F}(\Psi_m)} \right] (r_i + r_j) \cdot P_m^*(r_i) P_m(r_j).
\]

(20)

In Eq. (19) we see that the first term is symmetric and the second one is diagonal. The analytical expression for the CRLB, which is obtained from the inverse of the Fisher matrix, cannot be easily derived, but this inverse can be computed numerically instead. Since \(I_{F,ij}\) is an symmetric matrix with real elements, one can apply the eigenvalue decomposition to find the inverse of the Fisher matrix. Detailed examples are presented in the next section.

3. Direct calculation of the CRLB

As shown in Eq. (12), the CRLB is given by the diagonal elements of the inverse of matrix \(I_F\), which can be obtained by numerical computations. In this section, we present the results of some computed CRLB. We consider first plane wave illumination with various levels of noise. The result of structured illumination is considered in the last part of this section. Note that only Poisson noise is applied throughout our simulation. Other noise models (e.g. Gaussian noise or Poisson-Gaussian noise [22]) should be included when these are dominant. All of the calculation results given in this section are compared to the Monte Carlo experiment result shown in the next section.
3.1. Plane-wave illumination

We first investigate the CRLB in ptychography when the illumination (i.e. the probe) is a plane wave transmitted by a circular aperture. This probe function is normalized and shown in Fig. 1(a). The probe illuminates a complex-value object with the 'REAL' character as its real part and with the 'IMAG' character as its imaginary part, as shown in Fig. 1(c1) and Fig. 1(c2). In line with the conventional ptychography configuration, the probe moves over the object by a $2 \times 2$ regular grid. The $2 \times 2$ probes are combined to form a synthetic overall illumination pattern as illustrated in Fig. 1(b). The overlap ratio between adjacent illuminated areas is 70%, which is regarded as a prior knowledge and employed in the reconstruction algorithm. The overlap ratio is defined as follows. Suppose the diameter of the circular support is $L$, and the distance between adjacent illumination positions is $d$, where $0 < d < L$. The overlap ratio is then defined by:

$$\text{overlap ratio} = 1 - \frac{d}{L},$$

which is usually chosen between 60% and 85% to achieve optimal performance of the reconstruction algorithm [55].

To make the object function closer to practical situations, its maximum transmission is normalized to 1. The object is zero-padded and discretised to $70 \times 70$ pixels with $1\mu m$ resolution. The illuminated area of the object is roughly $40 \times 40 \mu m^2$. The discretised probe contains $60 \times 60$ pixels, with the same resolution. The circular support of the probe has a radius of $r_0 = 15 \mu m$. The wavelength of the illumination is $30 nm$. The far field intensities are measured with a detector at a propagation distance of $z = 5 cm$ behind the object. The detector has a $60 \times 60$ array of $50 \mu m$ pixels, hence the maximum spatial frequency that can be measured is $1 \mu m^{-1}$.

To compute the CRLB, we first construct the Fisher information matrix $I_F$ using Eq. (19). Although the number of degrees of freedom used to describe the object is very small ($70 \times 70 \times 2$
pixels), the discretised Fisher matrix already includes 9800 × 9800 pixels. The CRLB is obtained by numerically computing the inverse of $I_F$. We utilize the eigenvalue decomposition to compute the inverse of $I_F$. The diagonal elements of the inverse matrix $I_F^{-1}$ is reshaped to a 70 × 70 × 2 array, of which the first 70 × 70 elements correspond to the CRLB of the real part of the object function and the last 70 × 70 elements contain the CRLB of the imaginary part.

An important property of the CRLB is that it is proportional to the reciprocal of the signal-to-noise ratio (SNR), which can be estimated by the square root of the total photon number collected by the detector. Since the number of photons collected is not necessarily the same for every ptychographic measurements, we define the average total photon number (PN) as:

$$PN = \frac{\sum_{m} \text{total photon number for the } m\text{th probe position}}{\sum \text{number of probe positions } m}.$$  \hspace{1cm} (22)

Since PN is proportional to the illumination power, it is proportional to the squared modulus of the probe function. We first assume that $PN = 10^6$, so that the SNR is sufficiently high and the ptychography reconstruction using the maximum likelihood method is asymptotically close to the actual object. How to identify whether the SNR is high or low is addressed in Section 4.

For the case $PN = 10^6$, the computed CRLB for the real and imaginary part of the object are shown in Fig. 1(d1) and Fig. 1(d2), respectively. It is seen that when the probe is a plane wave and is known, the CRLB of the object resembles the overall illumination pattern shown in Fig. 1(b). In particular, the part of the object which is illuminated 4 times reaches a variance approximately 4 times smaller than the part which is illuminated only once, and this conclusion holds for both the real and imaginary part of the object. Although Eq. (12), Eq. (19) indicate that the CRLB in ptychography with Poisson noise is a functional of both the true probe and the true object, the calculation result shows that the illumination pattern most strongly determines the bound.

![Fig. 2. Calculated CRLB of the real part of the object, derived from the Fisher matrix, for noise levels given by (a) PN = 10^2, (b) PN = 10^3 and (c) PN = 10^4 and for a plane wave probe with circular support.](image)

The CRLB of the real part for the case that $PN = 10^2$, $10^3$ and $10^4$ are shown in Fig. 2. In accordance with Fig. 1(d), when the illumination is a (truncated) plane wave, the bound of the imaginary part is almost the same as that of the real part. Therefore we show only the CRLB of the real part in Fig. 2. Fig. 1(d) and Fig. 2 suggest that the bound is linearly proportional to the inverse of PN (i.e. the illumination power), which is in agreement with Eq. (19). In the next section, the CRLB shown in Fig. 1(d) and Fig. 2 are used as references for Monte Carlo experiments.

3.2. Structured illumination

The goal of this subsection is to investigate the influence of the illumination pattern on the CRLB. The real part of the probe now has the shape of the character 'P' as shown in Fig. 3(a1), while
the imaginary part is zero. The real and imaginary part of the object is the same as in Fig. 1(c1) and Fig. 1(c2). The scanning pattern of the probe, the discretization of the probe and the object, the discretization of the far field intensity and the calculation of diagonal elements of the inverse of the Fisher matrix are the same as in the last subsection. The CRLB for the real and imaginary part of the object are shown in Fig. 3(d1) and Fig. 3(d2), respectively. For a fair comparison with the case of a circular uniform probe function studied in the previous section, the average of the detected number of photons is again $10^6$.

Interestingly, for the case of structured illumination, the CRLB shown in Fig. 3(d) again resembles the overall illumination pattern shown in Fig. 3(b). In other words, the more illumination power we apply to the object, the lower the minimum variance of the obtained reconstruction. For the case of PN = $10^6$, it is clear that the CRLB is predominantly determined by the probe function. The maximum of the CRLB in Fig. 3(d) is in the yellow corner, and larger than in Fig. 1(d). This is because for our structured illumination the illuminating power is concentrated in the 'P' character, as shown in Fig. 3(a). In the following section, Monte Carlo simulations are reported for the case of structured illumination, and the result is compared with the CRLB in Fig. 3(d).

4. Monte Carlo analysis

To validate our calculation of the CRLB, Monte Carlo computations have been performed. For consistency, we discretise the probe and the object in the same way as described in Section 3.1. Hence the probe is again a plane wave transmitted by a circular aperture and is moved over the object by a $2 \times 2$ regular grid, and the overlap ratio is fixed at 70%. The probe function is discretised using a $60 \times 60$ grid. Hence, for every individual ptychography simulation 4 far field intensity patterns are detected. The wavelength of the illumination is $30\,nm$. The detector has a $60 \times 60$ array of $50\,\mu m$ pixels at a distance of $z = 5\,cm$ from the object, hence the maximum
spatial frequency that can be measured is \(1 \mu m^{-1}\). Since the probe has a circular support with a radius of \(r_0 = 15 \mu m\), the Fresnel number of the system is 0.15. Hence for this configuration the detector is in the Fraunhofer region.

The ptychographic data with various level of noise is generated as follows. For every ptychography simulation and for every probe position, we first calculate the noise-free diffracted wavefield in the far field. Then, the square modulus of the diffracted wavefields are assigned with corresponding photon numbers in accordance with the PN that is chosen, and the Poisson random number generator in MATLAB is applied to generate the noisy data.

Algorithm 1 Maximum likelihood method with Poisson noise

1: choose \(k_{\text{max}}, \delta_{L}\) and \(\gamma\), 
2: \(O_1 \leftarrow O_o\), 
3: \(k \leftarrow 1\), 
4: repeat 
5: \(\text{for each } R_m \text{ do}\) 
6: \(\text{estimate the number of photon measured: } N_{m,k} \leftarrow \frac{1}{\eta_{\text{tot}}} |F(P_m \cdot O_k)|^2\), 
7: \(g_{k,m} \leftarrow -P_m \cdot F^{-1} \left( \frac{N_{m,k} \gamma - 1}{N_{m,k} \gamma} \right) \cdot F(P_m \cdot O_k)\), 
8: \(\text{end for}\), 
9: \(\text{compute the steepest descent direction: } g_k \leftarrow \sum_m g_{k,m}\), 
10: if \(k = 1\) then 
11: \(\Delta_k \leftarrow g_k\), 
12: else 
13: use the formula of PolakĂŠRibière: \(\beta_k^{\text{PR}} \leftarrow \frac{\langle g_k - g_{k-1} | g_k \rangle}{\|g_k - g_{k-1}\|^2}\), 
14: \(\beta_k \leftarrow \max \left( \beta_k^{\text{PR}}, 0 \right)\), 
15: \(\text{compute the conjugate direction: } \Delta_k \leftarrow g_k + \beta_k \Delta_{k-1}\), 
16: \(\text{end if}\), 
17: \(\text{optimize the update step size: } \alpha_k \leftarrow \arg \min_{\alpha} \mathcal{L}_P (O_k + \alpha \Delta_k)\), 
18: \(\text{update the object function: } O_{k+1} \leftarrow O_k + \alpha_k \Delta_{O,k}\), 
19: \(k \leftarrow k + 1\), 
20: until \(k = k_{\text{max}} \text{ or } |\mathcal{L}_{P,k} - \mathcal{L}_{P,k-1}| \leq \delta_L\).

To verify asymptotically property of the maximum likelihood method in Eq. (13), we first developed and implemented Algorithm 1 as described in the pseudo-code chart shown below. In accordance with Eq. (9), we employ the gradient based optimization method. For the time being, the probe function is assumed to be known. In order to prevent that the algorithm terminates in a local minimum, the initial guess of the object is selected to be the actual object \(O_o(r)\). The denominator \(N_m\) in Eq. (9) is a function of \(r'\), and may be close to zero for some \(r'\). Hence the maximum likelihood method can be unstable. To avoid the instability, a regularization parameter \(\gamma\) is introduced in Algorithm 1, of which the value can be determined depending on the noise level. Unlike Eq. (9), the update step size \(\alpha\) is not a constant anymore in Algorithm 1. Instead, an optimal \(\alpha\) for every iteration \(k\) is obtained in the manner described in [56]: (1), approximate the likelihood function \(\mathcal{L}_P\) by a quadratic function of \(\alpha\); (2), choose the value for \(\alpha\) for which the quadratic function is at its minimum. To shorten the computation time and to improve the convergence, the conjugate gradient method [40,57] is implemented in Algorithm 1 also. The parameter \(\beta_k\) is chosen such that the update direction of the object function are conjugate between two subsequent iterations, for which many proposals exist [58]. Based on the formula of
Polak-\-Ribière [59], we choose $\beta_k = \max \left( \beta_k^{\text{PR}}, 0 \right)$, where $\beta_k^{\text{PR}}$ is given by:

$$
\beta_k^{\text{PR}} = \frac{\langle (g_k - g_{k-1}) | g_k \rangle}{\|g_{k-1}\|^2_2},
$$

(23)

where $g_k$ is the gradient of $\mathcal{L}_P$ with respect to $O(r)$ in the $k$th iteration. $\beta_k$ resets the search direction from the conjugate gradient back to the local decent gradient direction. **Algorithm 1** terminates when the change of the likelihood function between two subsequent iterations is smaller than a threshold $\delta_L$, or when the number of iteration reaches a maximum $k_{\text{max}}$.

For comparison, the performance of another popular method, namely the amplitude-based cost function minimization approach [6], is investigated in the Monte Carlo experiment also. The idea of this algorithm is to retrieve the object by minimizing the cost function defined in Eq. (5). Alternatively, one can also derives this algorithm from the maximum likelihood method by using the variance stabilization transform [22, 40, 41, 48, 49]. The approach is described in **Algorithm 2**, in which the search of the optimal step size $\alpha_k$ and the method of conjugate gradient are added too. Similar to **Algorithm 1**, **Algorithm 2** stops when the change of the cost function between two subsequent iterations is smaller than a threshold $\delta_E$, or when the number of iteration reaches a maximum $k_{\text{max}}$.

**Algorithm 2** Amplitude-based cost function minimization approach

1: choose $k_{\text{max}}$, $\delta_E$ and $\gamma$.
2: $O_1 \leftarrow O_o$.
3: $k \leftarrow 1$.
4: repeat
5: for each $R_m$ do
6: estimate the diffracted wavefield intensity: $I_{E,m,k} \leftarrow |\mathcal{F}(P_m \cdot O_k)|^2$.
7: $g_{k,m} \leftarrow -P_m^* \cdot \mathcal{F}^{-1} \left( \sqrt{\frac{\hbar \omega_m}{I_{E,m,k}}} - 1 \right) \cdot \mathcal{F}(P_m \cdot O_k)$.
8: end for.
9: compute the steepest descent direction: $g_k \leftarrow \sum_m g_{k,m}$.
10: follow 10th-19th steps of **Algorithm 1**.
11: until $k = k_{\text{max}}$ or $|E_k - E_{k-1}| \leq \delta_E$.

To investigate the performance of the above mentioned algorithms, three quantities are evaluated in our Monte Carlo analysis. The first is the mean value of the estimator $\hat{O}$ of the object function:

$$
\text{Mean} \left[ \hat{O}(r) \right] = \langle \hat{O}(r) \rangle = E \left[ \hat{O}(r) \right].
$$

(24)

The second quantity is the variance of the estimator:

$$
\text{Var} \left[ \hat{O}(r) \right] = E \left[ (\hat{O}(r) - \langle \hat{O}(r) \rangle)^2 \right].
$$

(25)

Moreover, the mean square error (MSE) which measures the deviation of the estimator from the actual object $O_o$ is also computed:

$$
\text{MSE} \left[ \hat{O}(r) \right] = E \left[ (\hat{O}(r) - O_o(r))^2 \right].
$$

(26)

The MSE can be regarded as a combination of the variance and the squared bias of the estimator [50]:

$$
\text{MSE} \left[ \hat{O}(r) \right] = \text{Var} \left[ \hat{O}(r) \right] + \text{Bias}^2 \left[ \hat{O}(r) \right].
$$

(27)
In order to compute the expectation accurately, 10000 individual ptychographic data sets are generated for the case of \( PN = 10^2, 10^3, 10^4 \) and \( 10^6 \). These data-sets are post-processed by Algorithm 1 and Algorithm 2, respectively, and the results are discussed next.

4.1. Plane wave illumination

We begin with plane wave illumination. The results for the CRLB obtained by the Fisher matrix in Section 3 are compared with the CRLB obtained by Monte Carlo simulations for different noise levels. When \( PN = 10^6 \), the mean, the MSE and the variance of the ptychography reconstruction are shown in Fig. 4. Fig. 4(a1) and Fig. 4(a2) show the real part and the imaginary part of the averaged object reconstruction, respectively, obtained using Algorithm 1. The MSE of the real and imaginary parts obtained with Algorithm 1 are plotted in Fig. 4(b1) and Fig. 4(b2), respectively. Similarly, the mean value, the MSE and the variance of the real and imaginary parts obtained with Algorithm 2 are shown in Fig. 4(d), Fig. 4(e) and Fig. 4(f), respectively.

It follows from Fig. 4(a) and Fig. 4(c) that for the case of \( PN = 10^6 \), both Algorithm 1 and Algorithm 2 provide reconstructions of which the mean value is approximately equal to the actual object. Moreover, the variance of the reconstruction is almost equal to the MSE for both Algorithm 1 and Algorithm 2. Therefore, in accordance with Eq. (27), we can conclude that both Algorithm 1 and Algorithm 2 are unbiased estimators that asymptotically achieve the CRLB when \( PN = 10^6 \). By comparing Fig. 4(c), Fig. 4(f) and Fig. 1(d), we see that the variance obtained with Algorithm 1 and Algorithm 2 are the same, and have a distribution very close to the CRLB computed from the Fisher matrix. Therefore our Monte Carlo result confirms that, for the same total number of photons, more localized illumination power leads to a lower variance of the same location in the reconstruction. Moreover, the variance of both algorithms are roughly 10% larger than the CRLB computed from the Fisher matrix, hence it can be concluded that Algorithm 1 and Algorithm 2 asymptotically achieve the same variance when the SNR is high. Hence, we conclude that both Algorithm 1 and Algorithm 2 have the asymptotic property described by Eq. (13) when \( PN = 10^6 \). By inspecting Fig. 4(c) and Fig. 4(f), one can furthermore infer that the variance is related to the illuminating power as mentioned in Section 3.1. In particular, the parts of the object that are illuminated 4 times have a variance that is 4 times smaller than the parts that are illuminated only once.

For the cases \( PN = 10^2, 10^3 \) and \( 10^4 \), the results obtained with Algorithm 1 and Algorithm 2 are shown in Fig. 5 and Fig. 6, respectively. Note that we only show the results for the real part of the mean, MSE and variance in Fig. 5 and Fig. 6, because the imaginary parts are almost the same as the real parts.

It is seen from Fig. 5 that the reconstruction obtained with Algorithm 1 deviates increasingly from the actual value when the SNR is decreased. This agrees with the MSE in Fig. 5. On the other hand, the MSE in Fig. 5 partially shows the same pattern as the variance, but a trace of the actual object is also visible. This trace is more observable when the value of PN is lower, which indicates that Algorithm 1 becomes more biased when the SNR is lower. When comparing the variances in Fig. 5 with Fig. 2, we see that the variance of Algorithm 1 is actually smaller than the CRLB computed from the Fisher matrix for the case that \( PN = 10^2 \) and \( 10^3 \). Because the CRLB is the minimum variance for unbiased estimator, the estimator given by Algorithm 1 is biased for smaller PN. Furthermore, the asymptotic property of the maximum likelihood method, given by Eq. (13), barely holds when \( PN = 10^4 \) and fails when \( PN = 10^3 \) and \( PN = 10^2 \).

Not surprisingly, these conclusions are valid also for Algorithm 2 as shown in Fig. 6. The reconstructions obtained with Algorithm 2 becomes increasingly biased when the value of PN is smaller. Furthermore, the variances shown in Fig. 6 and Fig. 2 imply that Algorithm 2 violates the CRLB when \( PN = 10^3 \) and \( 10^2 \). Therefore the asymptotic property given by Eq. (13) of the amplitude-based cost function minimization also barely holds when \( PN = 10^4 \) and fails when \( PN = 10^3 \) and \( PN = 10^2 \).
Fig. 4. The result of Monte Carlo experiment when \( PN = 10^6 \). (a1) and (a2) are the mean of the real part and the imaginary part obtained with Algorithm 1, (b1) and (b2) are the corresponding MSE, respectively, and (c1), (c2) show the variances. Similarly, (d1), (d2) are the mean, (e1), (e2) the MSE and (f1), (f2) the variance of the real and imaginary parts, respectively, obtained with Algorithm 2.
Comparing the variances in Fig. 5 and Fig. 6, it can be seen that Algorithm 1 and Algorithm 2 have different trends when decreasing the SNR. In particular, the variance of Algorithm 2 tends to stabilize at $\text{Var} = 0.01$ when $PN$ is decreased to $10^3$ and $10^2$, whereas Algorithm 1 does not. This can be explained from the fact that minimizing the amplitude-based cost function minimization can approximately be regarded as a variance stabilizing de-noising algorithm [22, 40, 41, 48]. From Fig. 5 and Fig. 6, we see that Algorithm 1 in fact reaches smaller MSEs than Algorithm 2 when $PN = 10^2, 10^3$ and $10^4$. This suggests that the approach based on the maximum likelihood principle can provides less bias than the amplitude-based cost function minimization method.

4.2. Structured illumination

Finally, we consider structured illumination using the same probe function as in Fig. 3(a). The overall illumination pattern is as in Fig. 3(b) and the object function shown in Fig. 3(c). In this part of the Monte Carlo analysis, we only investigate the case of $PN = 10^6$, and we only demonstrate the result by using Algorithm 1. Other scenarios are also interesting and deserve further research in the future. In Fig. 7 the real and imaginary part of the variance of Algorithm 1 are shown. The variance in Fig. 7 is very similar to that in Fig. 3(d) and closely related to the overall illumination pattern in Fig. 3(b). This leads again to the conclusion that the variance of the ptychographic reconstruction is mostly determined by the probe function for the case of
Fig. 6. (a1), (b1), (c1) are the mean, (a2), (b2), (c2) are the MSE and (a3), (b3), (c3) are the variance of the real part of the reconstructed object using Algorithm 2 for $PN = 10^2$, $PN = 10^3$ and $PN = 10^4$, respectively.

Poisson noise and high PN.

Fig. 7. (a1) and (a2) show the variance of the real and imaginary parts, respectively, of the reconstructed object using Algorithm 1 for $PN = 10^6$ and using the structures illumination of Fig. 3(a).
5. Conclusion

We have studied the influence of Poisson noise on ptychography by analyzing the CRLB. The CRLB was theoretically derived and numerically computed from the Fisher matrix for plane wave and structured illumination. It was found that if the estimator is unbiased, the minimum variance in the presence of Poisson noise is mostly determined by the probe function. The calculations of the CRLB suggest that the minimum variance is inversely proportional to the number of photons. The computations of the CRLB using the Fisher matrix were validated with Monte Carlo analysis. Both the maximum likelihood method and the amplitude-based cost function minimization algorithm were applied in the Monte Carlo simulations, using a conjugate gradient based implementation. It was shown that both approaches are asymptotically unbiased with variances that are slightly larger than the CRLB when the SNR is high. For the case of lower SNR, the Monte Carlo analysis showed that the approach using the amplitude-based cost function can provides a lower and a more stabilized variance. The stabilized variance can even be smaller than the CRLB which implies that these reconstructions are biased. On the other hand, the reconstruction method based on maximum likelihood principle is less biased. Furthermore, it was found that the probe function has a strong effect on the variance. Our result can help to understand the defects that occur in the ptychography reconstruction from noisy data. Our conclusions also suggest that more illumination power should be given to the part of object which is of most interest. As next steps of research, the performance of other ptychographic de-noising algorithm [19, 35, 37, 38, 47] deserve further investigation. Investigating the CRLB for Gaussian noise and the mixed Poisson-Gaussian noise is also an interesting topic for further research.

Funding

H2020 Marie Skłodowska-Curie Actions (675745).

Acknowledgments

X. Wei thanks Z. Xi and W. M. J. Coene for fruitful discussions.

Disclosures

The authors declare no conflicts of interest.

References

1. W. Hoppe, “Beugung im inhomogenen primRstrahlwellenfeld. i. prinzip einer phasenmessung von elektronenbeugungsinterferenzen,” Acta Crystallogr. Sect. A 25, 495–501 (1969).
2. J. M. Rodenburg and R. H. T. Bates, “The theory of super-resolution electron microscopy via wigner-distribution deconvolution,” Philos. Transactions Royal Soc. London. Ser. A: Phys. Eng. Sci. 339, 521–553 (1992).
3. H. N. Chapman, “Phase-retrieval x-ray microscopy by wigner-distribution deconvolution,” Ultramicroscopy 66, 153–172 (1996).
4. H. M. L. Faulkner and J. M. Rodenburg, “Movable aperture lensless transmission microscopy: A novel phase retrieval algorithm,” Phys. Rev. Lett. 93, 023903 (2004).
5. J. M. Rodenburg and H. M. L. Faulkner, “A phase retrieval algorithm for shifting illumination,” Appl. Phys. Lett. 85, 4795–4797 (2004).
6. M. Guizar-Sicairos and J. R. Fienup, “Phase retrieval with transverse translation diversity: a nonlinear optimization approach,” Opt. Express 16, 7264–7278 (2008).
7. J. C. da Silva and A. Menzel, “Elementary signals in ptychography,” Opt. Express 23, 33812–33821 (2015).
8. 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 1, 39–44 (2014).
9. M. Odstrcil, J. Bussmann, D. Rudolf, R. Bresenitz, J. Miao, W. S. Brooksby, and L. Juschkin, “Ptychographic imaging with a compact gas–discharge plasma extreme ultraviolet light source,” Opt. Lett. 40, 5574–5577 (2015).
10. J. M. Rodenburg, A. C. Hurst, A. G. Cullis, B. R. Dobson, F. Pfeiffer, O. Bunk, C. David, K. Jefimovs, and I. Johnson, “Hard-x-ray lensless imaging of extended objects,” Phys. Rev. Lett. 98, 034801 (2007).
11. P. Thibault, M. Dierolf, A. Menzel, O. Bunk, C. David, and F. Pfeiffer, “High-resolution scanning x-ray diffraction microscopy,” Science 321, 379–382 (2008).
12. H. N. Chapman and K. A. Nugent, “Coherent lensless x-ray imaging,” Nat. Photonics 4, 833–839 (2010).
13. F. Pfeiffer, “X-ray ptychography,” Nat. Photonics 12, 9–17 (2017).
14. P. Thibault, M. Dierolf, O. Bunk, A. Menzel, and F. Pfeiffer, “Probe retrieval in ptychographic coherent diffractive imaging,” Ultramicroscopy 109, 338–343 (2009).
15. A. M. Maiden and J. M. Rodenburg, “An improved ptychographical phase retrieval algorithm for diffractive imaging,” Ultramicroscopy 109, 1256–1262 (2009).
16. M. Holler, M. Guizar-Sicairos, E. H. R. Tsai, R. Dinapoli, E. Mâjliller, O. Bunk, J. Raabe, and G. Aeppli, “High-resolution non-destructive three-dimensional imaging of integrated circuits,” Nature 543, 402–406 (2017).
17. D. F. Gardner, M. Tanksalvala, E. R. Shanblatt, X. Zhang, B. R. Galloway, C. L. Porter, R. K. Jr, C. Bevis, D. E. Adams, H. C. Kaptyn, M. M. Murnane, and G. F. Mancini, “Subwavelength coherent imaging of periodic samples using a 13.5 nm tabletop high-harmoinic light source,” Nat. Photonics 11, 259–263 (2017).
18. Y. Jiang, Z. Chen, Y. Hao, P. Deh, H. Gao, S. Xie, P. Purohit, M. W. Tate, J. Park, S. M. Gruner, V. Elser, and D. A. Muller, “Electron ptychography of 2d materials to deep sub-Ångström resolution,” Nature 559, 343–349 (2018).
19. A. Maiden, D. Johnson, and P. Li, “Further improvements to the ptychographical iterative engine,” Optica 4, 736–745 (2017).
20. G. Zheng, R. Horstmeyer, and C. Yang, “Wide-field, high-resolution fourier ptychographic microscopy,” Nat. Photonics 7, 739–745 (2013).
21. 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,” Opt. Express 23, 33214–33240 (2015).
22. Y. Zhang, P. Song, and Q. Dai, “Fourier ptychographic microscopy using a generalized anscome transform approximation of the mixed poisson-gaussian likelihood,” Opt. Express 25, 168–179 (2017).
23. P. Thibault and A. Menzel, “Reconstructing state mixtures from diffraction measurements,” Nature 494, 68–71 (2013).
24. N. Burdet, X. Shi, D. Parks, J. N. Clark, X. Huang, S. D. Kevan, and I. K. Robinson, “Evaluation of partial coherence correction in x-ray ptychography,” Opt. Express 23, 5452–5467 (2015).
25. J. Zhong, L. Tian, P. Varma, and L. Waller, “Nonlinear optimization algorithm for partially coherent phase retrieval and source recovery,” IEEE Transactions on Comput. Imaging 2, 310–322 (2016).
26. D. J. Batey, D. Claus, and J. M. Rodenburg, “Information multiplexing in ptychography,” Ultramicroscopy 138, 13–21 (2014).
27. X. Wei and P. Urbach, “Ptychography with multiple wavelength illumination,” Opt. Express 27, 36767–36789 (2019).
28. A. M. Maiden, M. J. Humphry, and J. M. Rodenburg, “Ptychographic transmission microscopy in three dimensions using a multi-slice approach,” J. Opt. Soc. Am. A 29, 1606–1614 (2012).
29. M. A. Gilles, Y. S. G. Nashed, M. Du, C. Jacobsen, and S. M. Wild, “3d x-ray imaging of continuous objects beyond the depth of focus limit,” Optica 5, 1078–1086 (2018).
30. M. Kahnt, J. Becher, D. Bräjkner, Y. Fam, T. Sheppard, T. Weissenberger, F. Wittwer, J.-D. Grunwaldt, W. Schwieger, and C. G. Schroer, “Coupled ptychography and tomography algorithm improves reconstruction of experimental data,” Optica 6, 1282–1289 (2019).
31. P. M. Pelz, M. Guizar-Sicairos, P. Thibault, I. Johnson, M. Holler, and A. Menzel, “On-the-fly scans for x-ray ptychography,” Appl. Phys. Lett. 105, 251101 (2014).
32. J. Deng, Y. S. G. 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,” Opt. Express 23, 5438âʻÁ5451 (2015).
33. D. E. B. Flaes and S. Witte, “Interference probe ptychography for computational amplitude and phase microscopy,” Opt. Express 26, 31372–31390 (2018).
34. V. Elser, “Phase retrieval by iterated projections,” J. Opt. Soc. Am. A 20, 40–55 (2003).
35. Z. Wen, C. Yang, X. Liu, and S. Marchesini, “Alternating direction methods for classical and ptychographic phase retrieval,” Inverse Probl. 28, 115010 (2012).
36. S. Marchesini, A. Schirotzek, C. Yang, H. tieng Wu, and F. Maia, “Augmented projections for ptychographic imaging,” Inverse Probl. 29, 115009 (2013).
37. R. Horstmeyer, R. Y. Chen, X. Ou, B. Ames, J. A. Tropp, and C. Yang, “Solving ptychography with a convex relaxation,” New J. Phys. 17, 053044 (2015).
38. M. Pham, A. Rana, J. Miao, and S. Osher, “Semi-implicit relaxed douglas-rachford algorithm (sDR) for ptychography,” Opt. Express 27, 31246–31260 (2019).
39. A. Fanjiang and P. Chen, “Blind ptychography: Uniqueness and ambiguities,” arXiv: 1806.02674v3 (2018).
40. P. Thibault and M. Guizar-Sicairos, “Maximum-likelihood refinement for coherent diffractive imaging,” New J. Phys. 14, 063004 (2012).
41. P. Godard, M. Allain, V. Chamard, and J. Rodenburg, “Noise models for low counting rate coherent diffraction imaging,” Opt. Express 20, 25914–25934 (2012).
42. 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, 10395–10418 (2019).
43. A. Suzuki and Y. Takahashi, “Dark-field x-ray ptychography,” Opt. Express 23, 16429–16438 (2015).
44. 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,” Sci. Reports 3, 1927 (2013).

45. C. Zuo, J. Sun, and Q. Chen, “Adaptive step-size strategy for noise-robust fourier ptychographic microscopy,” Opt. Express 24, 20724–20744 (2016).

46. M. Ostdrčil, A. Menzel, and M. Guizar-Sicairos, “Iterative least-squares solver for generalized maximum-likelihood ptychography,” Opt. Express 26, 3108–3123 (2018).

47. A. P. Konijnienberg, W. M. J. Coene, and H. P. Urbach, “Model-independent noise-robust extension of ptychography,” Opt. Express 26, 5857–5874 (2018).

48. M. S. Bartlett, “The square root transformation in analysis of variance,” Suppl. to J. Royal Stat. Soc. 3, 68–78 (1936).

49. F. J. Anscombe, “The transformation of poisson, binomial and negative-binomial data,” Biometrika 35, 246–254 (1948).

50. J. Goodman, Introduction to Fourier Optics, McGraw-Hill physical and quantum electronics series (W. H. Freeman, 2005).

51. J. R. Shewchuk, “An introduction to the conjugate gradient method without the agonizing pain,” Tech. rep., Carnegie Mellon University (1994).