A direct solver for the phase retrieval problem in ptychographic imaging

Nada Sissouno[a,b,*], Florian Boßmann[d], Frank Filbir[b], Mark Iwen[e], Maik Kahnt[c], Rayan Saab[f], Christian Schroer[c], Wolfgang zu Castell[b]

[a] *Technical University of Munich, Faculty of Mathematics, Boltzmannstraße 3, Garching, 85748, Germany*
[b] *Helmholtz Zentrum München, Scientific Computing Research Unit, Neuherberg, 85674, Germany*
[c] *Deutsches Elektronen Synchrotron DESY, Hamburg, 22607, Germany*
[d] *Harbin Institute of Technology, Department of Mathematics, Harbin 150001, China*
[e] *Michigan State University, Department of Mathematics and Department of CMSE, East Lansing, MI, 48824, USA*
[f] *University of California San Diego, Department of Mathematics, La Jolla, CA, 92093, USA*

Abstract

Measurements achieved with ptychographic imaging are a special case of diffraction measurements. They are generated by illuminating small parts of a sample with, e.g., a focused X-ray beam. By shifting the sample, a set of far-field diffraction patterns of the whole sample are then obtained. From a mathematical point of view those measurements are the squared modulus of the windowed Fourier transform of the sample. Thus, we have a phase retrieval problem for local Fourier measurements. A direct solver for this problem was introduced by Iwen, Viswanathan and Wang in 2016 and improved by Iwen, Preskitt, Saab and Viswanathan in 2018. Motivated by the applied perspective of ptychographic imaging, we present a generalization of this method and compare the different versions in numerical experiments. The new method proposed herein turns out to be more stable, particularly in the case of missing data.

Keywords: Phase Retrieval, Ptychography, Image Reconstruction, Diffraction Imaging

2010 MSC: 49N45, 49N30, 42A38, 65T50

*Corresponding author
Email address: sissouno@ma.tum.de (Nada Sissouno)

1. Introduction

Ptychography refers to a diffraction imaging technique which collects a number of diffraction patterns of an object in the far-field, where each pattern is generated by illuminating a small subregion one at a time [1, 2, 3, 4]. The selection of the small and necessarily overlapping subregions is managed by using a mask, or window, placed between the X-ray beam and the object. For every shift of the mask, a diffraction pattern is recorded and the imaging task consists of reconstructing the object function from this collection of measurements. Since the measurement takes place in the far-field it is given as the squared modulus of the windowed Fourier transform of the object. This, in particular, results in a loss of the phase information of the signal. For the image formation we therefore have to reconstruct the object function from phaseless localized Fourier transform data. The imaging task in ptychography can thus be formulated in mathematical terms as follows. For an unknown object modeled as a complex-valued function $f \in L^2(\mathbb{R}^d)$ supported on a compact set $\Omega \subset \mathbb{R}^d$ we measure

$$y(\tau,\omega) = \left|\mathcal{F}[f\, T_\tau w](\omega)\right|^2 = \left|\mathcal{F}[T_\tau^* f\, w](\omega)\right|^2, \tag{1}$$

where $w \in L^2(\mathbb{R}^d)$ is a compactly supported window function, $T_\tau f(x) = f(x-\tau)$ is the translation operator with adjoint $T_\tau^* f(x) = f(x+\tau)$, and $\mathcal{F}$ is the Fourier transform defined as $\mathcal{F}f(\omega) = \int_{\mathbb{R}^d} f(x)\, \mathrm{e}^{-2\pi \mathrm{i}\, \omega\cdot x}\, \mathrm{d}x$. We have to invert the nonlinear mapping $f \to y$. Note that the latter formulation in (1) more closely represents the real experimental situation where the object is shifted instead of the window. Of course, in concrete applications we are given only samples $\{y(\tau_\ell, \xi_k)\}$ and the inversion problem has to be formulated in a fully discrete regime. Discretization of (1) on a grid $\{n/N : n = 0, \ldots, N-1\}$ leads to

$$y_{\ell,k} = |\langle \boldsymbol{f}, M_k T_\ell \boldsymbol{w}\rangle|^2, \tag{2}$$

where $\boldsymbol{f} = (f(0), \ldots, f(N-1))^T$ and $\boldsymbol{w} = (w(0), \ldots, w(N-1))^T \in \mathbb{C}^N$ are vectors containing the sampled values of the function f resp. w. The translation

operator T_ℓ and modulation operator M_k are acting on the entries of the vectors by $T_\ell f(n) = f(n-\ell \mod N)$ resp. $M_k f(n) = \mathrm{e}^{2\pi \mathrm{i} nk/N} f(n)$. They have obvious matrix representations. Recovery of $\boldsymbol{f}$ from data (2) falls in the class of problems where a vector $\boldsymbol{x}$ has to be recovered from data of the form $|\langle \boldsymbol{x}, \boldsymbol{\varphi}_m \rangle|$ with $\{\boldsymbol{\varphi}_m\}$ being a frame for $\mathbb{C}^N$. We are considering here the case where frame is given as a discrete Gabor frame $\{M_k T_\ell \boldsymbol{w}\}$ for $k = 0, \ldots, K$ and $\ell \in \mathcal{L} \subseteq \{0, \ldots, N-1\}$ with $\mathrm{card}(\mathcal{L}) = L + 1$ for K, $L \leq N - 1$.

The general Phase Retrieval Problem has a long history and it was tackled in many different ways and under various assumptions. We will make no attempt to review these developments in detail here but refer to [5, 6] and references cited there. Among the diverse techniques one method seems to be particularly popular among practitioners. In the physics literature this method is known as the Ptychographic Interative Engine (PIE), and is regarded as a standard approach for image formation from ptychographic data within the physics community [7, 8]. Mathematically this approach is an alternating projection method which goes back to the work of Gerchberg and Saxton [9] and Fienup [10]. The method projects alternately on the set which consists of functions with support in Ω and the set of functions which agree with the measurement [11]. The latter set is non-convex which makes the problem notoriously difficult to analyze theoretically. Due to the non-convexity, the alternating projection method can converge to a stationary point that differs from the true solution. Determining a good starting point in general is also not easy and some attempts were made to come up with a good initial guess, see for example [12] for a recent approach in this regard.

The method we present in this paper is an adoption of a fast direct solver for the phase retrieval problem (2) as it was developed by Iwen et. al. in [13, 14]. The method is based on a lifting scheme as used in the PhaseLift algorithm [15] which transforms the discretized non-linear problem (2) into a linear problem for the lifted variables $\boldsymbol{f}\boldsymbol{f}^*$. After recovering some of the entries of the lifted variables $\boldsymbol{f}\boldsymbol{f}^*$ the phase of the individual entries of $\boldsymbol{f}$ can be determined by an angular synchronization approach. This finally results in the reconstruction

of the function f on the grid up to a global phase multiple provided that the measurements are sufficiently informative.

In this paper we will demonstrate how to apply a new algorithm for phase retrieval from short-time Fourier measurements to the concrete experimental setup of ptychography. The paper is organized as follows. We state the Algorithm in the Section 2. Here, we first introduce the 1D case and later give its 2D version. In Section 3 we demonstrate the method with numerical examples. Finally, Section 4 concludes with a discussion of the proposed technique.

2. Phase Retrieval from Localized Fourier Measurements

2.1. Description of the Algorithm

Let $f, w \in L^2(\mathbb{R}^d)$ be compactly supported functions. Without loss of generality we may assume $\operatorname{supp}(f) \subseteq [0,1]^d$. The short-time Fourier transform of f with window w is defined as

$$V_w f(\tau,\omega) := \int_{\mathbb{R}^d} f(t)\,\bar{w}(t-\tau)\,\mathrm{e}^{-2\pi\mathrm{i}\omega\tau}\,\mathrm{d}t = \langle f, M_\omega T_\tau w\rangle_{L^2(\mathbb{R}^n)}, \tag{3}$$

where $T_\tau w(t) = w(t-\tau)$, $M_\omega w(t) = \mathrm{e}^{2\pi\mathrm{i}\omega t} w(t)$ are the translation resp. modulation operator. We concentrate on the cases $d \in \{1,2\}$. For clarity of presentation, we first restrict ourselves to the case $d = 1$. The case $d = 2$ will be addressed in Subsection 2.2. For discretization let $N \in \mathbb{N}$ and consider the grid $\Gamma = \{n/N : n = 0, \dots N-1\}$. Discretization of the Fourier integral in (3) on Γ and subsequential evaluation of the resulting semi-discrete transform w.r.t. τ on the grid Γ leads to the fully discretized transform which can be considered as a short-time Fourier transform on the cyclic group $\mathbb{Z}_N$. It is given as

$$V_w f(\ell,k) = \frac{1}{N}\sum_{n\in\mathbb{Z}_N} f(n)\bar{w}(m-\ell)\,\mathrm{e}^{-2\pi\mathrm{i}k\cdot n/N}, \quad \ell,k \in \mathbb{Z}_N. \tag{4}$$

This can be expressed in a more condensed form as

$$V_w f(\ell,k) = \frac{1}{N}\langle \boldsymbol{f}, M_k T_\ell \boldsymbol{w}\rangle_{\mathbb{C}^N}. \tag{5}$$

where $\boldsymbol{f}$, $\boldsymbol{w} \in \mathbb{C}^N$, $T_\ell f(n) = f(n-\ell \mod N)$, and $M_k f(n) = \mathrm{e}^{2\pi\mathrm{i}nk/N} f(n)$ as already defined after (2). The scaling factor $1/N$ in (5) is of no relevance for

our consideration and will therefore be neglected henceforth.

The reconstruction problem can now be formulated as follows. We have to reconstruct the vector $\boldsymbol{f}$ from data

$$y_{\ell,k} = \left|\langle \boldsymbol{f}, T_\ell \bar{\boldsymbol{w}}_k \rangle_{\mathbb{C}^N}\right|^2, \quad k = 0, \dots, K, \ \ell \in \mathcal{L} \subseteq \{0, \dots, N-1\}, \tag{6}$$

with $\mathrm{card}(\mathcal{L}) = L+1$ and $K, L \leq N-1$, where $\boldsymbol{w}_k = M_k \boldsymbol{w}$ is the modulation of the vector of window $\boldsymbol{w}$. We will henceforth assume that $\boldsymbol{w}$ is supported on the first s entries, i.e., $w(n) \neq 0$ for $0 \leq n < s$ and 0 elsewhere.

These nonlinear measurements can be "lifted" to linear measurements on the space of matrices as follows

$$\left|\langle \boldsymbol{f}, T_\ell \bar{\boldsymbol{w}}_k \rangle_{\mathbb{C}^N}\right|^2 = \langle \boldsymbol{f}, T_\ell \bar{\boldsymbol{w}}_k \rangle_{\mathbb{C}^N} \, \overline{\langle \boldsymbol{f}, T_\ell \bar{\boldsymbol{w}}_k \rangle_{\mathbb{C}^N}} = \langle \boldsymbol{f}\boldsymbol{f}^*, T_\ell \bar{\boldsymbol{w}}_k (T_\ell \bar{\boldsymbol{w}}_k)^* \rangle_{HS}. \tag{7}$$

Here $\langle \cdot, \cdot \rangle_{HS}$ is the Hilbert-Schmidt inner product $\langle \boldsymbol{A}, \boldsymbol{B} \rangle_{HS} = \mathrm{trace}(\boldsymbol{A}^* \boldsymbol{B})$.

For $\boldsymbol{X} \in \mathbb{C}^{N \times N}$ we arrange the numbers $\langle \boldsymbol{X}, T_\ell \bar{\boldsymbol{w}}_k (T_\ell \bar{\boldsymbol{w}}_k)^* \rangle_{HS}$ as a vector in $\mathbb{C}^D$ for $D = (K+1)(L+1)$ and define a linear operator $\mathcal{A} : \mathbb{C}^{N \times N} \to \mathbb{C}^D$ by

$$\mathcal{A}(\boldsymbol{X}) = \left(\langle \boldsymbol{X}, T_\ell \bar{\boldsymbol{w}}_k (T_\ell \bar{\boldsymbol{w}}_k)^* \rangle_{HS}\right)_{\alpha=1}^{D}, \quad \alpha = k + 1 + \ell(K+1) \tag{8}$$

and we will call this operator the measurement operator as it coincides with the vector of measurements if $\boldsymbol{X} = \boldsymbol{f}\boldsymbol{f}^*$ according to (7). The operator (8) can not be injective on $\mathbb{C}^{N \times N}$ when $\boldsymbol{w}$ is supported on the first s entries of N. But, depending on the choice of $\boldsymbol{w}$, it might be stably invertible if we restrict it to the space

$$\mathrm{span}\{T_\ell \bar{\boldsymbol{w}}_k \bar{\boldsymbol{w}}_k^* T_\ell^* : \ell \in \mathcal{L}, \ k = 0, \dots, K\} =: \mathbb{P}_{LK}. \tag{9}$$

We denote the corresponding projection operator by $\mathcal{T}_{LK} : \mathbb{C}^{N \times N} \to \mathbb{P}_{LK}$. We will refer to this operator as the *tight projector*. Let $\mathcal{A}_{LK} = \mathcal{A}|_{\mathbb{P}_{LK}}$ be the restriction of $\mathcal{A}$ to the space $\mathbb{P}_{LK}$. Depending of the choice of $\boldsymbol{w}$, the matrix of this linear operator with respect to the generating system $\{T_\ell \bar{\boldsymbol{w}}_k \bar{\boldsymbol{w}}_k^* T_\ell^*\}_{\ell,k}$ is a symmetric positive definite matrix. Hence, the operator is (stably) invertible and we can determine $\boldsymbol{X} = \mathcal{T}_{LK}(\boldsymbol{f}\boldsymbol{f}^*)$ from the data uniquely solving

$$\mathcal{A}_{LK}(\boldsymbol{X}) = \boldsymbol{y}, \quad \boldsymbol{y} = (y_{\ell,k}) \tag{10}$$

since

$$y_{\ell,k} = \langle \mathcal{T}_{LK}(\boldsymbol{f}\boldsymbol{f}^*), T_\ell \bar{\boldsymbol{w}}_k \bar{\boldsymbol{w}}_k^* T_\ell^* \rangle_{HS}. \tag{11}$$

Note that for $L = N - 1$ we have

$$\mathbb{P}_{LK} = \operatorname{span}\{\boldsymbol{E}_{i,j} : |i - j \mod N| < s\} =: \mathbb{V}_s,$$

where the $\boldsymbol{E}_{i,j}$'s constitute the standard basis of $\mathbb{C}^{N\times N}$. This case coincides with the projection considered in [14]. We have $\dim \mathbb{V}_s = (2s-1)N$. Thus, invertibility can only be achieved in general for $K + 1 = 2s - 1$. If $L < N - 1$, $\mathbb{P}_{LK}$ can not be expressed in terms of $(L + 1)(K + 1)$ standard basis vectors $\boldsymbol{E}_{i,j}$. Therefore, the algorithm in [14] needs to be adapted in our context. In Section 3 we will compare the two methods where we will refer to the projection operator given in terms of $\boldsymbol{E}_{i,j}$'s in (10) and (11) as *pattern projector* $\mathcal{P}_{LK}$, since the projector reflects the zero pattern generated by the support of the windows.

In order to determine an approximation of the vector $\boldsymbol{f}$ from $\boldsymbol{X} = \mathcal{T}_{LK}(\boldsymbol{f}\boldsymbol{f}^*)$ we consider the representation $\boldsymbol{Z}$ of $\boldsymbol{X}$ in the standard basis. The amplitudes can now be determined by taking the square-roots of the main-diagonal of $\boldsymbol{Z}$. To reconstruct the phases we can use the following angular synchronization technique: Let $\boldsymbol{z} \in \mathbb{C}^N$ with $z_n = |z_n|\mathrm{e}^{\mathrm{i}\Theta_n}$, let $\tilde{\boldsymbol{z}} := \boldsymbol{z}/|\boldsymbol{z}| = (\mathrm{e}^{\mathrm{i}\Theta_n})_n$ and define $\tilde{\boldsymbol{Z}} := \boldsymbol{Z}/|\boldsymbol{Z}|$, where the operations are considered elementwise for the non-zero entries. The phases of $\boldsymbol{z}$ are then given by the first eigenvector of $\tilde{\boldsymbol{Z}}$. In case of $L = N-1$ it can be shown [16] that the reconstruction is exact. In other words, up to a global phase shift, the phases of $\boldsymbol{z}$ coincide with the phases of $\boldsymbol{f}$. For $L < N - 1$ this process can be stabilized by considering the first eigenvector of $\boldsymbol{D}^{-1/2}\tilde{\boldsymbol{Z}}\boldsymbol{D}^{-1/2}$. Here $\boldsymbol{D}$ is the degree matrix which is a diagonal matrix whose diagonal entries coincide with the number of non-zero entries in the corresponding row of $\tilde{\boldsymbol{Z}}$ (see [17] for more details).

The algorithm for recovering $\boldsymbol{f}$ from measurements $y_{\ell,k}$ consists of five steps which can be summarized as follows.

Algorithm

Input:	Measurements $\boldsymbol{y} = (y_{\ell,k}) \in \mathbb{C}^D$
Output:	$\tilde{\boldsymbol{f}} \approx \boldsymbol{f}$
1.	Compute $\boldsymbol{X} = \mathcal{A}_{LK}^{-1}(\boldsymbol{y})$.
2.	Calculate $\boldsymbol{Z}$, the representation of $\boldsymbol{X}$ in the standard basis.
3.	Form $\tilde{\boldsymbol{Z}} = \boldsymbol{Z}/\lvert\boldsymbol{Z}\rvert$, normalizing non-zero entries of $\boldsymbol{Z}$.
4.	Compute the eigenvector $\tilde{\boldsymbol{z}}$ of $\tilde{\boldsymbol{Z}}$ corresponding to the largest eigenvalue.
5.	Form $\tilde{\boldsymbol{f}}$ via $f_n = \sqrt{Z_{n,n}}\ \tilde{z}_n$.

The crucial step of the algorithm is the first, i.e., we have to ensure by a suitable choice of the window $\boldsymbol{w}$ that (10) has a unique solution which can be computed in a stable manner. In [13, 14] it has been shown that this is the case for the following choice for $\boldsymbol{w}$:

$$w_a(n) = \begin{cases} (2s-1)^{-1/4}\ \mathrm{e}^{-n/a}, & n < s, \\ 0, & otherwise, \end{cases} \tag{12}$$

where $a \in [4, \infty)$. In the numerical examples we will call these *exponential windows* (EW).

However, this window does not reflect the concrete experimental situation. Closer to physical reality is a shape of the window function which is given by an Airy function, i.e., $w(t) = (J_1(t)/t)^2$ where J_1 is the Bessel function of first order. Experiments show that a Gaussian function seems to be an acceptable approximation to this type of window function (see, e.g., [3]). Further, the Gaussian function should be (i) centered at the midpoint of the support and (ii) normalized such that the norm coincides the number of photons in the experiment denoted by $n(p)$. Therefore, we propose to use the *Gaussian windows* (GW) constructed as follows: we consider a Gaussian window function with normalization $c_{n(p)}$ such that (ii) is satisfied. To ensure (i), we choose the support according to an α-quantile t_α for some $\alpha \in (0, 1)$, i.e., the continuous window

function is given by $w_\alpha(t) = c_{n(p)}\, e^{-t^2/2} \cdot \chi([-t_\alpha, t_\alpha])$. By uniform sampling on $[-t_\alpha, t_\alpha]$ we get $\boldsymbol{w}$ with entries

$$w_\alpha(n) = \begin{cases} c_{n(p)}\, e^{-\frac{t_\alpha^2}{2(s-1)^2}(2n-s+1)^2} & \text{for } n < s \\ 0 & \text{for } n \geq s \end{cases}. \tag{13}$$

In Section 3 we will compare reconstructions using the exponential window (EW) and Gaussian window (GW) as well as the use of the pattern projector and tight projector for 2D data. Thus, we first give some remarks on the 2D case.

2.2. The 2D Case

The discretization of (3) for the case $d = 2$ uses the cartesian grid $\Gamma \times \Gamma$ and finally leads in an analogous manner as for the 1D case to

$$V_w f(\ell, k) = \frac{1}{N^2} \sum_{n_1, n_2 \in \mathbb{Z}_N} f(n_1, n_2)\, w(n_1 - \ell_1, n_2 - \ell_2)\, \mathrm{e}^{-2\pi \mathrm{i}(k_1 n_1 + k_2 n_2)/N^2}, \tag{14}$$

with $\ell = (\ell_1, \ell_2), k = (k_1, k_2) \in \mathbb{Z}_N \times \mathbb{Z}_N$. Analogously to (5) we can express this as

$$V_w f(\ell, k) = \frac{1}{N^2} \langle \boldsymbol{F}, M_k T_\ell \boldsymbol{W} \rangle_{HS}, \tag{15}$$

where $\boldsymbol{F} = \big(f(n_1, n_2)\big)_{n_1, n_2 = 0}^{N-1}$, $\boldsymbol{W} = \big(w(n_1, n_2)\big)_{n_1, n_2 = 0}^{N-1} \in \mathbb{C}^{N \times N}$ and with the usual convention that the translation resp. modulation operators are acting entrywise. As in the 1D case we will ignore the factor $1/N^2$ henceforth.

Following [18], for the 2D case we make the assumption that the window function w in (3) separates w.r.t. the variables, i.e., $w(t) = w(t_1, t_2) = u(t_1)\, v(t_2)$. We assume moreover that $\mathrm{supp}(u) = \mathrm{supp}(v)$. Clearly, the Gaussian window $w(t) = \mathrm{e}^{-|t|^2}$ satisfies this assumption, but note that the Airy function does not. With this assumption we have $\boldsymbol{W} = \boldsymbol{u}\boldsymbol{v}^*$ with $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^N$ and, moreover, $M_k T_\ell \boldsymbol{W} = (M_{k_1} T_{\ell_1} \boldsymbol{u})\, (M_{k_2} T_{\ell_2} \boldsymbol{v})^*$. In order to recover $\boldsymbol{F}$ from measurements

$$Y_{\ell,k} = \big|\langle \boldsymbol{F}, M_k T_\ell \boldsymbol{W} \rangle_{HS}\big|^2 = \big|\langle \boldsymbol{F}, T_\ell \boldsymbol{W}_k \rangle_{HS}\big|^2 \tag{16}$$

with $\boldsymbol{W}_k = M_k \boldsymbol{W}$ we adapt the algorithm for the 1D case as follows. For $\boldsymbol{X} \in \mathbb{C}^{N \times N}$ let $\vec{\boldsymbol{X}} = \mathrm{vec}(\boldsymbol{X})$ be its vectorization. Note that for $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{C}^{N \times N}$

we have $\langle \boldsymbol{X}, \boldsymbol{Y} \rangle_{HS} = \langle \vec{\boldsymbol{X}}, \vec{\boldsymbol{Y}} \rangle_{\mathbb{C}^{N^2}}$. With this preparation it is now obvious how to transfer the algorithm presented in Section 2.1 to the 2D case. Applying the same lifting step as in (7) we obtain

$$Y_{\ell,k} = \left\langle \vec{\boldsymbol{F}} \vec{\boldsymbol{F}}^*, \overrightarrow{(T_\ell \boldsymbol{W}_k)}\, \overrightarrow{(T_\ell \boldsymbol{W}_k)}^* \right\rangle_{HS} \tag{17}$$

where according to $\boldsymbol{W} = \boldsymbol{u}\boldsymbol{v}^*$, $T_\ell \boldsymbol{W}_k = (T_{\ell_1} M_{k_1} \boldsymbol{u})\ (T_{\ell_2} M_{-k_2} \boldsymbol{v})^*$.

3. Numerical examples

We test the proposed method on simulated ptychographic data using the object shown in Figure 1. For an easier comparison the color range of all plots has been set to $[0, 1]$ for amplitudes and $[-\pi, \pi]$ for phases. However, keep in mind that the phase is only reconstructed up to a global shift. For the experiments we use an object size of 128×128 pixels and each pixel is about 30×30 nm. The size of the Fourier measurements per shift was chosen to be 15×15 pixels, i.e., $K = 15$ frequencies in both directions. This coincides to a window size of 8x8 pixels ($s = 8$ in each direction). To generate the measurement data a Gaussian beam with main-focus over the support of the window was simulated.

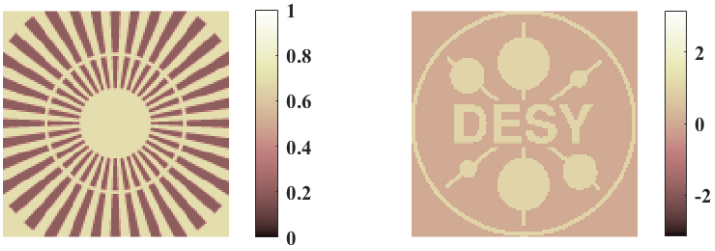

Figure 1: Original amplitude (left) and original phase (right) of the simulated object

In a first experiment we reconstruct the object using all shifts of the window function that fit into the 128x128 pattern of the object. Given a window size of 8x8 pixels, this corresponds to $L = 121$ shifts for each dimension. Note that this already differs from the setup given in [14] where also circulant shifts are considered, i.e., $L = N - 1 = 127$. We compare the reconstruction quality

using the exponential window (12) analyzed in [13, 14] against the Gaussian window (13) with $\alpha = 0.99$, which more closely approximates experimental setups. The results are shown in Figure 2. Since we do not consider circulant shifts, both reconstructions show artifacts at the sides of the images. Because the exponential window is not centered, these artifacts concentrate at the lower and left side of the reconstruction. Moreover, the exponential window shows strong artifacts especially in the phase of the reconstruction as it does not fit the window form given in applications.

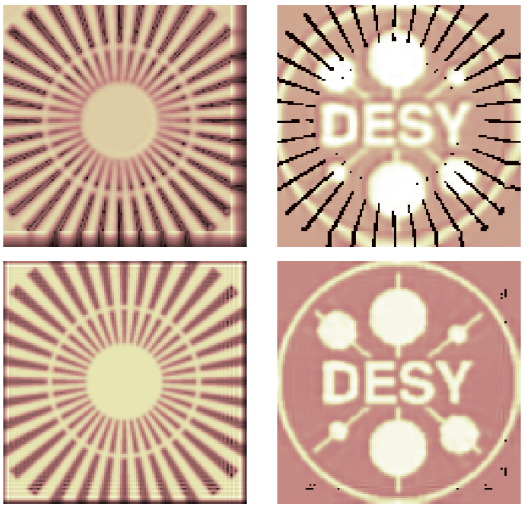

Figure 2: Reconstructed amplitude (left) and phase (right) using exponential window (top) and Gaussian window (bottom)

In our next experiment we compare the proposed tight projector against the pattern projector used in [14]. We already know that both projection spaces coincide if all shifts are taken into account. Thus, we now consider the case where only every κ shift is used for the reconstruction. Therefore let $\mathcal{T}^{\kappa}_{LK}$ be the projector onto the subspace

$$\operatorname{span}\{T_{\kappa\ell}\bar{\boldsymbol{w}}_k\bar{\boldsymbol{w}}_k^* T_{\kappa\ell}^* : \ell = 0,\dots,L,\ k = 0,\dots,K\},$$

and $\mathcal{P}^{\kappa}_{LK}$ the corresponding pattern projector. Note that the pattern projector does not necessarily return Hermitian matrices as required for the angular syn-

chronization. Thus, the algorithm has to be extended to include an update step $\boldsymbol{Z} \leftarrow (\boldsymbol{Z} + \boldsymbol{Z}^*)/2$. A careful analysis regarding this approximation with respect to the phase retrieval problem was made by Iwen et.al. in [14].

For the experiment, we set L, K as above, and $\kappa = 4$. The results are illustrated in Figure 3. For both projectors a Gaussian window (13) with $\alpha = 0.99$ is used. As seen before, the reconstructed amplitude of both projectors is similar. However, the phase reconstruction is much more stable using the proposed tight projector. The pattern projector reconstruction shows strong artifacts almost dominating the original phase.

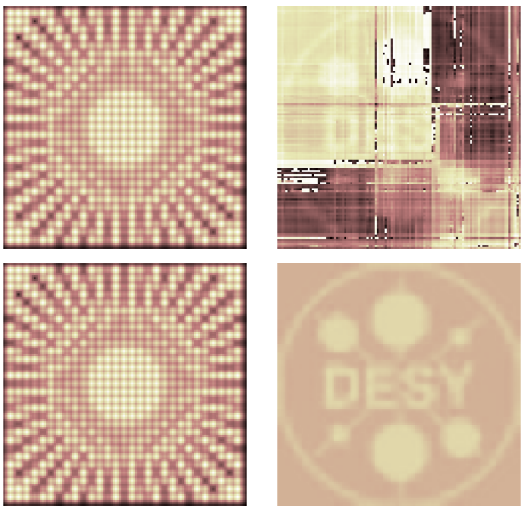

Figure 3: Reconstructed amplitude (left) and phase (right) using the pattern projector (top) and tight projector (bottom)

We verified these observations with extensive numerical experiments applying the reconstruction technique with both projectors using different window functions. We simulated measurements of the object in Figure 1 using seven different Gaussian windows. For the reconstruction we used four window functions, one exponential window (12) and three Gaussian windows (13) that differ from the windows used for simulation. Figure 4 shows the mean squared error (MSE) averaged over all 28 combinations. Here we define the MSE as $\|\boldsymbol{F} - \tilde{\boldsymbol{F}}\|_F^2/N^2$. The original data has an amplitude range of $[0.2, 0.7]$, a phase

range of $[0, \pi/2]$ and a Frobenius norm of $\|\boldsymbol{F}\|_F = 65.31$. Clearly, the tight projector leads to a much more stable technique. As was expected, when the shift is equal to the support size of the windows (i.e., $\kappa = s = 8$) both methods fail. In Figure 4 the total error and the error considering only the reconstruction of the phase is shown. The error in amplitude is not illustrated since it basically coincides for both methods. Thus, the tight operator especially stabilizes the phase reconstruction.

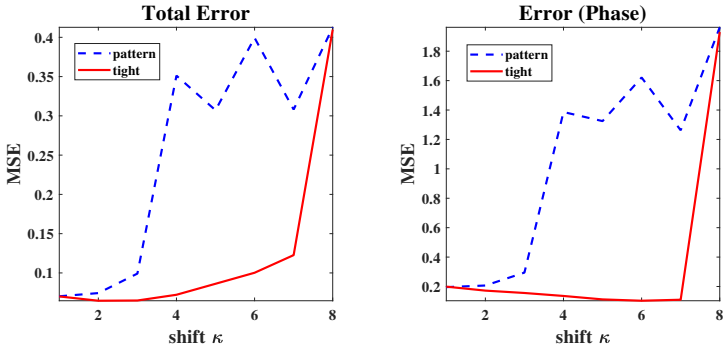

Figure 4: Reconstruction error using pattern and tight projector for different shift sizes κ. Left: total error, right: phase error.

Next, we compare the reconstruction error of the tight projector for different types of window functions. The mean squared error for different shift sizes κ is shown in Figure 5. Here we used Gaussian windows (13) with $\alpha = 0.9, 0.95, 0.99$ and an exponential window (12). We observe that the reconstruction is stable up to a shift of $\kappa = s = 8$ independent of the window. The clipped Gaussian windows (13) result in a smaller error since they more closely approximate the ptychographic windows used to generate the measurements.

Note that some window functions, such as GW with $\alpha = 0.9$, even seem to perform better when not all shifts are taken into account. (Also compare the results shown in Figure 4.) This is possibly due to the the hard cut-off of the window function for small parameters $\alpha < 0.95$ which appears to contribute to Gibbs-like oscillations in the reconstructions. This effect appears to diminish when the shift increases slightly.

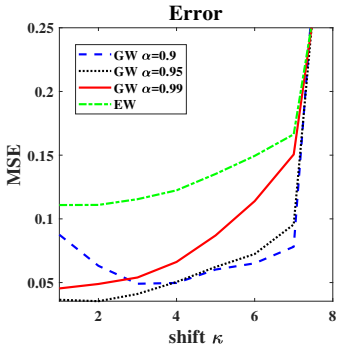

Figure 5: Reconstruction error for different types of window functions using the tight projector.

4. Conclusion

We presented a direct algorithm for ptychographic reconstruction for known windows. We have shown numerically that using a window based on the normal distribution results in good reconstructions even if the real window is not fully known but is assumed to be approximately Gaussian. In addition, although the algorithm was originally designed for reconstructions based on full circulant shifts, i.e., $L = N-1$, we have also shown that our proposed modifications result in good reconstructions of phase and amplitude when fewer shifts are used.

Acknowledgements

Mark Iwen was supported in part by NSF CCF 1615489. Rayan Saab was supported in part by NSF DMS 1517204. Nada Sissouno acknowledges support by the German Science Foundation (DFG) in the context of the Emmy-Noether-Junior Research Group *Randomized Sensing and Quantization of Signals and Images* (KR 4512/1-1).

References

References

[1] R. Hegerl, W. Hoppe, Dynamic theory of crystalline structure analysis by electron diffraction in homogeneous primary wave field, Berichte der

Bunsen-Gesellschaft für physikalische Chemie 74 (1970) 1148.

[2] J. M. Rodenburg, Ptychography and Related Diffractive Imaging Methods, Advances in Imaging and Electron Physics, Elsevier 74 (2008) 87–184.

[3] F. Seiboth, A. Schropp, M. Scholz, F. Wittwer, C. Rdel, M. Wnsche, T. Ullsperger, S. Nolte, J. Rahomki, K. Parfeniukas, S. Giakoumidis, U. Vogt, U. Wagner, C. Rau, U. Boesenberg, J. Garrevoet, G. Falkenberg, E. C. Galtier, H. J. Lee, B. Nagler, C. G. Schroer, Perfect X-ray focusing via fitting corrective glasses to aberrated optics, Nature Communications 8 (2017).

[4] F. Pfeiffer, X-ray ptychography, Nature Photonics 12 (2018) 9–17.

[5] D. R. Luke, Phase Retrieval. What's New?, SIAM SIAG/OPT Views and News 25 (2017) 1–6.

[6] Y. C. Eldar, N. Hammen, D. G. Mixon, Recent Advances in Phase Retrieval, IEEE Signal Processing Magazine 33 (2016) 158–162.

[7] A. M. Maiden, J. M. Rodenburg, An improved ptychographical phase retrieval algorithm for diffractive imaging, Ultramicroscopy 109 (2009) 1256–1262.

[8] A. Maiden, D. Johnson, P. Li, Further improvements to the ptychographical iterative engine, Optica 4 (2017) 736–745.

[9] R. W. Gerchberg, W. O. Saxton, A practical algorithm for the determination of phase from image and diffraction plane pictures, Optik 35 (1972) 237–246.

[10] J. R. Fienup, Reconstruction of a complex-valued object from the modulus of its Fourier transform using support constraint, J. Opt. Soc. Am. A 4 (1986) 118–123.

[11] H. H. Bauschke, P. L. Combettes, D. R. Luke, Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization, J. Opt. Soc. Am. A 19 (2002) 1334–1345.

[12] S. Marchesini, Y.-C. Tu, H.-T. Wu, Alternating projection, ptychographic imaging and phase synchronization, Applied and Computational Harmonic Analysis 41 (2016) 815–851.

[13] M. A. Iwen, A. Viswanathan, Y. Wang, Fast Phase Retrieval from Local Correlation Measurements 9 (2016) 1655–1688.

[14] M. A. Iwen, B. Preskitt, R. Saab, A. Viswanathan, Phase retrieval from local measurements: improved robustness via eigenvector-based angular synchronization, Applied and Computational Harmonic Analysis (2018). In press.

[15] E. J. Cands, T. Strohmer, V. Voroninski, PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex Programming, Communications on Pure and Applied Mathematics 66 (2013) 1241–1274.

[16] A. Viswanathan, M. Iwen, Fast angular synchronization for phase retrieval via incomplete information, 2015. URL: https://doi.org/10.1117/12.2186336. doi:10.1117/12.2186336.

[17] B. P. Preskitt, Phase Retrieval from Locally Supported Measurements, Ph.D. thesis, UC San Diego, 2018.

[18] M. Iwen, B. Preskitt, R. Saab, A. Viswanathan, Phase retrieval from local measurements in two dimensions, in: Wavelets and Sparsity XVII, volume 10394, International Society for Optics and Photonics, 2017, p. 103940X.