

PAPER

Coded aperture ptychography: uniqueness and reconstruction

To cite this article: Pengwen Chen and Albert Fannjiang 2018 *Inverse Problems* **34** 025003

View the [article online](#) for updates and enhancements.

Coded aperture ptychography: uniqueness and reconstruction

Pengwen Chen¹ and Albert Fannjiang^{2,3} 

¹ Applied Mathematics, National Chung Hsing University, Taiwan

² Department of Mathematics, University of California, Davis, CA, United States of America

E-mail: pengwen@nchu.edu.tw and fannjiang@math.ucdavis.edu

Received 5 September 2017, revised 21 November 2017

Accepted for publication 4 December 2017

Published 10 January 2018



CrossMark

Abstract

Uniqueness of solution is proved for any ptychographic scheme with a random mask under a minimum overlap condition and local geometric convergence analysis is given for the alternating projection (AP) and Douglas–Rachford (DR) algorithms. DR is shown to possess a unique fixed point in the object domain and for AP a simple criterion for distinguishing the true solution among possibly many fixed points is given.

A minimalist scheme, where the adjacent masks overlap 50% of the area and each pixel of the object is illuminated by exactly four illuminations, is conveniently parametrized by the number q of shifted masks in each direction. The lower bound $1 - C/q^2$ is proved for the geometric convergence rate of the minimalist scheme, predicting a poor performance with large q which is confirmed by numerical experiments. The twin-image ambiguity is shown to arise for certain Fresnel masks and degrade the performance of reconstruction.

Extensive numerical experiments are performed to explore the general features of a well-performing mask, the optimal value of q and the robustness with respect to measurement noise.

Keywords: ptychography, phase retrieval, uniqueness, geometric convergence, twin-image ambiguity

(Some figures may appear in colour only in the online journal)

³ Author to whom any correspondence should be addressed.

1. Introduction

Optical or x-ray ptychography is a coherent diffractive imaging method that uses multiple micro-diffraction patterns obtained through the scan of a localized illumination on the specimen. As such ptychography is a synthetic aperture technique and, along with advances in detection and computation techniques, has enabled microscopies with enhanced resolution and robustness without the need for lenses [3, 6, 20, 27, 28].

Ptychography was initially proposed by Hoppe for transmission electron diffraction microscopy [15]. In his pioneering work Hoppe showed that recording diffraction patterns at two positions removes the remaining ambiguity between the correct solution and its complex conjugate. Hoppe [16] has considered the extension to non-periodic objects with phase-shifting plates as well.

However, only after Faulkner and Rodenburg proposed the so called ptychographical iterative engine (PIE) [10, 11, 25], the redundant information collected via overlapping illuminations was effectively harnessed (see also [12, 21, 23, 28, 29]). A key to success of ptychographic reconstruction is that the adjacent illuminated areas overlap substantially, around 60–70% in each direction [2, 23].

The first question for any inverse problems, including ptychography, is uniqueness of solution. This has been resolved in [18] for the ptychographic scheme where all possible shifts of a damped and windowed Fourier transform are used, i.e. with the maximum overlap between adjacent illuminated areas (see more discussion in section 1.3). However, maximum overlap requires overly redundant measurements and hence the uniqueness question remains open for practical ptychographic schemes with a significantly less overlap.

To demonstrate that the uniqueness issue is relevant in practice and nontrivial in theory, we show that the twin-image ambiguity can be present in a realistic ptychographic setting (appendix), resulting in poor numerical performance (figure 2(a)).

Another mathematical question surrounding ptychography is convergence analysis of reconstruction algorithms. Few results in the literature offer concrete conditions for verifying convergence to the true solution and give an explicit estimate for the convergence rate [14, 18, 30]. In particular, (global or local) geometric convergence to the *true* ptychographic solution has not been established for any ptychographic reconstruction that assumes less than the maximum overlap between adjacent illuminated areas.

On the other hand, the dynamic range of x-ray detectors is a key factor limiting both the spatial resolution and sensitivity of x-ray ptychography with a zero-order component several orders of magnitude more intense than the scattered field. To this end, a beam-stop may be introduced to block the zero-order component. Alternatively, a randomly phased mask (a diffuser) can be deployed to reduce the dynamic range of the recorded diffraction patterns by more than one order of magnitude [8, 22, 24, 27].

With these motivations, a main purpose of the present work is to establish the uniqueness theorem for ptychography with a random mask under a minimum overlap condition and to prove local geometric convergence to the true solution for the widely used alternating projections (AP) and Douglas–Rachford (DR) algorithms. Moreover, we give an explicit bound for the rate of convergence for both algorithms with the minimalist ptychographic scheme introduced below.

First we describe how each constituent diffraction pattern is measured in our ptychographic scheme.

1.1. Oversampled diffraction pattern

Let f^0 be a part of the unknown object f restricted to the initial subdomain

$$\mathcal{M}^0 = \{\mathbf{n} = (n_1, n_2) \in \mathbb{Z}^2 : 0 \leq n_1, n_2 \leq m\}.$$

Let the Fourier transform of f be written as

$$F^0(\mathbf{w}) = \sum_{\mathbf{m} \in \mathcal{M}^0} e^{-i2\pi\mathbf{m}\cdot\mathbf{w}} f^0(\mathbf{m}), \quad \mathbf{w} = (w_1, w_2).$$

Under the Fraunhofer approximation, the diffraction pattern is proportional to $|F^0(\mathbf{w})|^2$ which can be written as

$$I^0(\mathbf{w}) = \sum_{\mathbf{n} = -(m,m)}^{(m,m)} \left\{ \sum_{\mathbf{m} \in \mathcal{M}^0} f^0(\mathbf{m} + \mathbf{n}) \overline{f^0(\mathbf{m})} \right\} e^{-i2\pi\mathbf{n}\cdot\mathbf{w}}, \quad \mathbf{w} \in [0, 1]^2. \quad (1)$$

Here and below the over-line notation means complex conjugacy.

The expression in the parentheses in (1) is the autocorrelation function of f^0 and the summation over \mathbf{n} takes the form of Fourier transform on the enlarged grid

$$\widetilde{\mathcal{M}}^0 = \{(m_1, m_2) \in \mathbb{Z}^2 : -m \leq m_1 \leq m, -m \leq m_2 \leq m\}$$

which suggests sampling $I^0(\mathbf{w})$ on the grid

$$\mathcal{L} = \left\{ (w_1, w_2) \mid w_j = 0, \frac{1}{2m+1}, \frac{2}{2m+1}, \dots, \frac{2m}{2m+1} \right\}. \quad (2)$$

Let $\Phi^0 : \mathbb{C}^{|\mathcal{M}^0|} \rightarrow \mathbb{C}^{|\widetilde{\mathcal{M}}^0|}$ be the \mathcal{L} -sampled discrete Fourier transform (ODFT) defined on \mathcal{M}^0 . We can write $I^0(\mathbf{w}) = |\Phi^0 f^0(\mathbf{w})|^2$ for all $\mathbf{w} \in \mathcal{L}$.

A randomly coded diffraction pattern measured with a mask is the diffraction pattern for the *masked object* $g^0(\mathbf{n}) = f^0(\mathbf{n})\mu^0(\mathbf{n})$ where the mask function μ^0 is a finite array of random variables. With $\mu^0(\mathbf{n}) = |\mu^0(\mathbf{n})|e^{i\phi(\mathbf{n})}$ we will focus on the effect of *random phase* ϕ . For the uniqueness theorem we will assume $\phi(\mathbf{n})$ to be independent, continuous real-valued random variables. In other words, each $\phi(\mathbf{n})$ is independently distributed with a probability density function on $[0, 2\pi]$ that may depend on \mathbf{n} .

The continuity assumption on ϕ is a technical one for proving *almost sure* uniqueness. If ϕ are discrete random variables, then we would have to settle for uniqueness with high probability. Continuous phase modulation can be experimentally realized with various techniques depending on the wavelength. See [17, 19, 22, 30, 33] for recent innovation and development of random phase modulation techniques.

We also assume that $|\mu^0(\mathbf{n})| \neq 0, \forall \mathbf{n} \in \mathcal{M}$ (i.e. the mask is transparent). This is necessary for unique reconstruction of the object as any opaque pixel of the mask where $\mu^0(\mathbf{n}) = 0$ would block the transmission of the information $f^0(\mathbf{n})$. By absorbing $|\mu^0(\mathbf{n})|$ into the object function we can assume, without loss of generality, that $|\mu^0(\mathbf{n})| = 1, \forall \mathbf{n} \in \mathcal{M}^0$, i.e. μ^0 represents a *phase* mask.

Now consider the simplest, 2-part ptychographic set-up: the object domain is the union of two overlapping square grids, one of which is the translate of the other square grid. Denote the two square grids by \mathcal{M}^0 and $\mathcal{M}^{\mathbf{t}}$ which is the shift of \mathcal{M}^0 by the displacement vector $\mathbf{t} = (t_1, t_2) \in \mathbb{Z}^2$. We shall make the overlap assumption

$$|\mathcal{M}^0 \cap \mathcal{M}^{\mathbf{t}} \cap \text{supp}(f)| \geq 2, \quad (3)$$

where $|\cdot|$ denotes the cardinality of a set, i.e. the intersection of the two grids contains at least two points from the support of the object.

Let $f^{\mathbf{t}}$ be the unknown object restricted to $\mathcal{M}^{\mathbf{t}}$ and $\Phi^{\mathbf{t}}$ the ODFT on $\mathcal{M}^{\mathbf{t}}$. We write the object function as $f = f^0 \cup f^{\mathbf{t}}$ where $f^0(\mathbf{m}) = f^{\mathbf{t}}(\mathbf{m})$ for all $\mathbf{m} \in \mathcal{M}^0 \cap \mathcal{M}^{\mathbf{t}}$. Let f^0 and $f^{\mathbf{t}}$ be

respectively illuminated with the mask μ^0 and the mask $\mu^{\mathbf{t}}$ on where $\mu^{\mathbf{t}}(\mathbf{n}) = \mu^0(\mathbf{n} - \mathbf{t})$, for all $\mathbf{n} \in \mathcal{M}^{\mathbf{t}}$.

For multi-part ptychography, let the object domain be contained in the union of the shifted square grids:

$$\text{supp}(f) \subseteq \bigcup_{\mathbf{t} \in \mathcal{T}} \mathcal{M}^{\mathbf{t}} \quad (4)$$

where \mathcal{T} is a set of shifts. Under (4) we can write

$$f = \bigcup_{\mathbf{t} \in \mathcal{T}} f^{\mathbf{t}}. \quad (5)$$

Analogous to (3) we assume that for every $\mathcal{M}^{\mathbf{t}_1} \cap \text{supp}(f) \neq \emptyset$, $\mathbf{t}_1 \in \mathcal{T}$, there is another $\mathcal{M}^{\mathbf{t}_2}$, $\mathbf{t}_2 \in \mathcal{T}$, $\mathbf{t}_2 \neq \mathbf{t}_1$ such that

$$|\mathcal{M}^{\mathbf{t}_1} \cap \mathcal{M}^{\mathbf{t}_2} \cap \text{supp}(f)| \geq 2. \quad (6)$$

In other words, *every connected component of the object is contained in the union of at least two distinct masks (i.e. ptychographically measured) whose intersection contains at least two points of the object support.* Since the support of the object is often not known precisely, some illuminations may totally miss the object and produce no useful information. These illuminations and the resulting diffraction data are easily recognized and should be discarded.

1.2. A minimalist ptychographic scheme

Although our uniqueness theorem and local convergence analysis are for general ptychographic measurement schemes satisfying (6), we will consider the following lattice scheme for numerical experiments and explicit estimation of convergence rate. We call this lattice scheme *the minimalist scheme* because each object pixel participates exactly in four diffraction patterns (two in each direction) with 50% overlap between two adjacent mask domains. Any less overlap would result in non-uniform coverage of the object.

Suppose the initial mask domain \mathcal{M}^0 is $m \times m$ (m is an even integer) and an adjacent domain is obtained by shifting $m/2$ in either direction. To cover each object pixel exactly four times (two in each direction) we assume $m = 2n/q$ with an integer q . This amounts to q^2 diffraction patterns. In other words, we consider the shift T^{kl} corresponding to the displacement $\mathbf{t}_{kl} = \frac{m}{2}(k, l)$, with $k \in \{0, 1, \dots, q-1\}$, $l \in \{0, 1, \dots, q-1\}$. When $k = q-1$ or $l = q-1$ we assume for simplicity that the shifted mask is wrapped around into the other end of the object domain (i.e. the periodic boundary condition). Four times coverage and four times oversampling in each diffraction pattern together produce the total number $q^2(2m-1)^2 \approx 16n^2$ of data.

We emphasize two features of the minimalist scheme: (i) It has a fixed total oversampling ratio (≈ 16) independent of the total number of shifted masks, q^2 ; (ii) It has the minimum (50%) overlap between two adjacent masks while maintaining the same number of coverage (i.e. 4) for every pixel of the object. Note that the 50% overlap is lower than the required overlap found empirically (i.e. 60%–70%) from previous studies [2, 23].

With this minimalist ptychographic scheme, we study analytically and numerically how q and the structure of the mask affect ptychographic reconstruction.

1.3. Main contributions

The first result of the present work is the almost sure uniqueness of ptychographic solution with a random mask under the minimum overlap condition (6) (theorem 2.3 and corollary 2.5).

To demonstrate how certain symmetry of a deterministic mask can spoil uniqueness of solution, we show in appendix that the minimalist scheme with a certain Fresnel mask admits both the true image and its twin-like image as solutions, resulting in poor numerical performance (figure 2(a)).

In this connection, Iwen *et al* [18] proved a uniqueness theorem for the ptychographic scheme where all possible shifts of a damped and windowed Fourier transform are used, i.e. the overlap percentage between two adjacent mask domains is at the maximum. In our notation, this amounts to n^2 oversampled diffraction patterns totaling $(2m-1)^2 n^2$ number of data. Their uniqueness theorem also holds with probability $1 - \mathcal{O}(\ln^{-2} n \ln^{-3}(\ln n))$ after randomly selecting a subset of $\mathcal{O}(n^2 \ln^2 n \ln^3(\ln n))$ data (assuming m is at least poly-log in n).

In section 3, we establish local, geometric convergence (theorem 3.4) for the AP and the DR algorithms under the minimum overlap condition (6). We also prove the uniqueness of the DR fixed point in the object domain (proposition 3.1) and give an easily verifiable criterion for distinguishing the true solution among many AP fixed points (proposition 3.3).

In comparison, Wen *et al* [30] proposed alternating direction methods (ADM), including DR, for ptychographic reconstruction and demonstrated good numerical performance. Hesse *et al* [14] proved global convergence to a critical point for a proximal-regularized alternating minimization formulation of blind ptychography. Many critical points, however, may co-exist and there is no easy way of distinguishing the true solution from the rest. Neither of these papers establishes uniqueness, convergence to the *true* solution or the geometric sense of convergence.

We also give a bound on the convergence rate of AP and DR for the minimalist scheme introduced in section 1.2 (proposition 4.1). The bound shows that the convergence rate can deteriorate rapidly as q becomes large, indicating that the best performing q are in the small and medium ranges. Our numerical experiments in section 5 bear this prediction out nicely, focusing on two kinds of masks: (independent or correlated) random masks and the Fresnel mask.

We prove that twin image exists in ptychography with the Fresnel mask at certain values of the Fresnel number (propositions A.1 and A.2) and causes the reconstruction error to spike (figure 2(a)).

Finally, we perform extensive numerical experiments to explore what the general features of a well-performing mask, the optimal value of q and the noise stability of the minimalist scheme. We summarize our numerical findings in the Conclusion (section 6).

2. Uniqueness of ptychographic solution

First we recall some basic results from *nonptychographic* phase retrieval where the mask μ and the unknown object f have the same dimension.

The z -transform

$$F(\mathbf{z}) = \sum_{\mathbf{n}} \mathbf{f}(\mathbf{n}) \mathbf{z}^{-\mathbf{n}}$$

of f is a polynomial in \mathbf{z}^{-1} and can be factorized uniquely into the product of irreducible polynomials $F_k(\mathbf{z})$ and a monomial in \mathbf{z}^{-1}

$$F(\mathbf{z}) = \alpha \mathbf{z}^{-\mathbf{n}_0} \prod_{\mathbf{k}=1}^{\mathbf{p}} \mathbf{F}_{\mathbf{k}}(\mathbf{z}), \quad (7)$$

where \mathbf{n}_0 is a vector of nonnegative integers and α is a complex coefficient.

Proposition 2.1. [13] Let the z -transform $F(\mathbf{z})$ of a finite complex-valued array $\{f(\mathbf{n})\}$ be given by

$$F(\mathbf{z}) = \alpha \mathbf{z}^{-\mathbf{m}} \prod_{k=1}^p \mathbf{F}_k(\mathbf{z}), \quad \mathbf{m} \in \mathbb{N}^d, \alpha \in \mathbb{C} \quad (8)$$

where $F_k, k = 1, \dots, p$ are nontrivial irreducible polynomials. Let $G(\mathbf{z})$ be the \mathbf{z} -transform of another finite array $g(\mathbf{n})$. Suppose $|F(\mathbf{w})| = |G(\mathbf{w})|, \forall \mathbf{w} \in [0, 1]^d$. Then $G(\mathbf{z})$ must have the form

$$G(\mathbf{z}) = |\alpha| e^{i\theta} \mathbf{z}^{-\mathbf{p}} \left(\prod_{k \in I} \mathbf{F}_k(\mathbf{z}) \right) \left(\prod_{k \in I^c} \mathbf{F}_k^*(1/\mathbf{z}^*) \right), \quad \mathbf{p} \in \mathbb{N}^d, \theta \in \mathbb{R}$$

where I is a subset of $\{1, 2, \dots, p\}$.

2.1. Line object

f is a *line object* if the convex hull of the object support in \mathbb{R}^d is a line segment.

Proposition 2.2. [7] Suppose f is not a line object and let the mask μ 's phase be continuously and independently distributed. Then with probability one the only irreducible factor of the z -transform of the masked object $\tilde{f}(\mathbf{n}) = f(\mathbf{n})\mu(\mathbf{n})$ is a monomial of \mathbf{z}^{-1} .

The following uniqueness theorem is our first theoretical result.

Theorem 2.3. Suppose that the assumptions of proposition 2.2 hold and that

$$|\mathcal{M}^0 \cap \mathcal{M}^t \cap \text{supp}(f)| \geq 2.$$

Then with probability one f is uniquely determined, up to a global phase factor, by the Ptychographic data $b = |A^*f|$.

Remark 2.4. The assumption of a random mask is probably unnecessary for uniqueness of solution. But the counterexample presented in appendix shows that the twin-image ambiguity can arise for certain deterministic, Fresnel masks.

Proof. Let $g(\mathbf{n})$ be another array that vanishes outside $\mathcal{M}^0 \cup \mathcal{M}^t$ and produces the same masked Fourier magnitude data. By propositions 2.1 and 2.2, g has the following possibilities: In \mathcal{M}^0 , g has two alternatives

$$g(\mathbf{n}) = \begin{cases} e^{i\theta_1} f^0(\mathbf{n} + \mathbf{m}_1) \mu^0(\mathbf{n} + \mathbf{m}_1) / \mu^0(\mathbf{n}) \\ e^{i\theta_1} \bar{f}^0(\mathbf{N} - \mathbf{n} + \mathbf{m}_1) \bar{\mu}^0(\mathbf{N} - \mathbf{n} + \mathbf{m}_1) / \mu^0(\mathbf{n}), \end{cases} \quad \forall \mathbf{n} \in \mathcal{M}^0 \quad (9)$$

and

$$g(\mathbf{n}) = \begin{cases} e^{i\theta_2} f^t(\mathbf{n} + \mathbf{m}_2) \mu^t(\mathbf{n} + \mathbf{m}_2) / \mu^t(\mathbf{n}) \\ e^{i\theta_2} \bar{f}^t(\mathbf{N} - \mathbf{n} + \mathbf{m}_2) \bar{\mu}^t(\mathbf{N} - \mathbf{n} + \mathbf{m}_2) / \mu^t(\mathbf{n}), \end{cases} \quad \forall \mathbf{n} \in \mathcal{M}^t \quad (10)$$

for some $\mathbf{m}_1, \mathbf{m}_2 \in \mathbb{Z}^d, \theta_1, \theta_2 \in \mathbb{R}$.

We now focus on the intersection $\mathcal{M}^0 \cap \mathcal{M}^t$ where (9) and (10) are both defined. We have

then four scenarios from the crossover of the alternatives in (9) and (10).

First of all, if, for all $\mathbf{n} \in \mathcal{M}^0 \cap \mathcal{M}^t$,

$$\begin{aligned} g(\mathbf{n}) &= e^{i\theta_1} f^0(\mathbf{n} + \mathbf{m}_1) \mu^0(\mathbf{n} + \mathbf{m}_1) / \mu^0(\mathbf{n}) \\ &= e^{i\theta_2} f^t(\mathbf{n} + \mathbf{m}_2) \mu^t(\mathbf{n} + \mathbf{m}_2) / \mu^t(\mathbf{n}) \end{aligned} \quad (11)$$

then

$$e^{i\theta_1} f^0(\mathbf{n} + \mathbf{m}_1) \mu^0(\mathbf{n} + \mathbf{m}_1) / \mu^0(\mathbf{n}) = e^{i\theta_2} f^t(\mathbf{n} + \mathbf{m}_2) \mu^0(\mathbf{n} - \mathbf{t} + \mathbf{m}_2) / \mu^0(\mathbf{n} - \mathbf{t}). \quad (12)$$

Clearly, $f(\mathbf{n} + \mathbf{m}_1)$ and $f^t(\mathbf{n} + \mathbf{m}_2)$ must simultaneously be zero or nonzero. When they are nonzero, we obtain by taking logarithm on both sides

$$\begin{aligned} i\theta_1 + \ln f^0(\mathbf{n} + \mathbf{m}_1) + \ln \mu^0(\mathbf{n} + \mathbf{m}_1) + \ln \mu^0(\mathbf{n} - \mathbf{t}) \\ = i\theta_2 + \ln f^t(\mathbf{n} + \mathbf{m}_2) + \ln \mu^0(\mathbf{n} - \mathbf{t} + \mathbf{m}_2) + \ln \mu^0(\mathbf{n}) \end{aligned} \quad (13)$$

which holds up to a multiple of 2π . The four random variables

$$\ln \mu^0(\mathbf{n} + \mathbf{m}_1), \ln \mu^0(\mathbf{n} - \mathbf{t}), \ln \mu^0(\mathbf{n} - \mathbf{t} + \mathbf{m}_2), \ln \mu^0(\mathbf{n})$$

can not cancel one another unless either $\mathbf{m}_1 = \mathbf{m}_2 = 0$ or $(\mathbf{t} = 0 \ \& \ \mathbf{m}_1 = \mathbf{m}_2)$. When the continuous random variables do not cancel one another, (13) fails to hold true almost surely.

On the other hand, for $\mathbf{m}_1 = \mathbf{m}_2 = 0$ (since $\mathbf{t} \neq 0$), it follows from (11) that

$$g(\mathbf{n}) = e^{i\theta_1} f^0(\mathbf{n}) = e^{i\theta_2} f^t(\mathbf{n}), \quad \mathbf{n} \in \mathcal{M}^0 \cap \mathcal{M}^t.$$

Since $f^0(\mathbf{n}) = f^t(\mathbf{n})$, we have $\theta_1 = \theta_2$. It follows then from (9)–(10) that $g = e^{i\theta_1} (f^0 \cup f^t)$.

The other three scenarios can be similarly dealt with. Consider the next scenario where for $\mathbf{n} \in \mathcal{M}^0 \cap \mathcal{M}^t$

$$\begin{aligned} g(\mathbf{n}) &= e^{i\theta_1} f^0(\mathbf{n} + \mathbf{m}_1) \mu^0(\mathbf{n} + \mathbf{m}_1) / \mu^0(\mathbf{n}) \\ &= e^{i\theta_2} \bar{f}^t(\mathbf{N} - \mathbf{n} + \mathbf{m}_2) \bar{\mu}^t(\mathbf{N} - \mathbf{n} + \mathbf{m}_2) / \mu^t(\mathbf{n}). \end{aligned}$$

Taking logarithm and rearranging terms we have

$$\begin{aligned} i\theta_1 + \ln f^0(\mathbf{n} + \mathbf{m}_1) + \ln \mu^0(\mathbf{n} + \mathbf{m}_1) + \ln \mu^0(\mathbf{n} - \mathbf{t}) + \ln \mu^0(\mathbf{N} - \mathbf{n} - \mathbf{t} + \mathbf{m}_2) \\ = i\theta_2 + \ln \bar{f}^t(\mathbf{N} - \mathbf{n} + \mathbf{m}_2) + \ln \mu^0(\mathbf{n}) \end{aligned} \quad (14)$$

The four random variables

$$\ln \mu^0(\mathbf{n} + \mathbf{m}_1), \ln \mu^0(\mathbf{n} - \mathbf{t}), \ln \mu^0(\mathbf{N} - \mathbf{n} - \mathbf{t} + \mathbf{m}_2), \ln \mu^0(\mathbf{n})$$

can not cancel one another since $\mathbf{t} \neq 0$. As a result, (14) holds true with probability zero.

The argument for ruling out the third scenario

$$\begin{aligned} g(\mathbf{n}) &= e^{i\theta_1} \bar{f}^0(\mathbf{N} - \mathbf{n} + \mathbf{m}_1) \bar{\mu}^0(\mathbf{N} - \mathbf{n} + \mathbf{m}_1) / \mu^0(\mathbf{n}) \\ &= e^{i\theta_2} f^t(\mathbf{n} + \mathbf{m}_2) \mu^t(\mathbf{n} + \mathbf{m}_2) / \mu^t(\mathbf{n}) \end{aligned}$$

is the same as for the second scenario.

Now consider the fourth scenario

$$\begin{aligned} g(\mathbf{n}) &= e^{i\theta_1 \bar{f}^0(\mathbf{N} - \mathbf{n} + \mathbf{m}_1)} \bar{\mu}^0(\mathbf{N} - \mathbf{n} + \mathbf{m}_1) / \mu^0(\mathbf{n}) \\ &= e^{i\theta_2 \bar{f}^t(\mathbf{N} - \mathbf{n} + \mathbf{m}_2)} \bar{\mu}^t(\mathbf{N} - \mathbf{n} + \mathbf{m}_2) / \mu^t(\mathbf{n}) \end{aligned}$$

which after taking logarithm and rearranging terms becomes

$$\begin{aligned} i\theta_1 + \ln \bar{f}^0(\mathbf{N} - \mathbf{n} + \mathbf{m}_1) + \ln \mu^0(\mathbf{N} - \mathbf{n} - \mathbf{t} + \mathbf{m}_2) + \ln \mu^0(\mathbf{n} - \mathbf{t}) \\ = i\theta_2 + \ln \bar{f}^t(\mathbf{N} - \mathbf{n} + \mathbf{m}_2) + \ln \mu^0(\mathbf{N} - \mathbf{n} + \mathbf{m}_1) + \ln \mu^0(\mathbf{n}). \end{aligned} \quad (15)$$

Since $\mathbf{t} \neq \mathbf{0}$, the four random variables

$$\ln \mu^0(\mathbf{N} - \mathbf{n} - \mathbf{t} + \mathbf{m}_2), \ln \mu^0(\mathbf{n} - \mathbf{t}), \ln \mu^0(\mathbf{N} - \mathbf{n} + \mathbf{m}_1), \ln \mu^0(\mathbf{n})$$

cancel one another only when

$$\begin{aligned} \mathbf{N} - \mathbf{n} - \mathbf{t} + \mathbf{m}_2 &= \mathbf{n} \\ \mathbf{n} - \mathbf{t} &= \mathbf{N} - \mathbf{n} + \mathbf{m}_1 \end{aligned}$$

or equivalently

$$\begin{aligned} \mathbf{m}_2 &= 2\mathbf{n} - \mathbf{N} + \mathbf{t} \\ \mathbf{m}_1 &= 2\mathbf{n} - \mathbf{N} - \mathbf{t} \end{aligned}$$

which can not hold true simultaneously for more than one \mathbf{n} for any given $\mathbf{m}_1, \mathbf{m}_2$. This is ruled out by the assumption that $\mathcal{M}^0 \cap \mathcal{M}^t \cap \text{supp}(f)$ contains at least two points.

In summary, the only possibility is that

$$g = e^{i\theta} (f^0 \cup f^t) = e^{i\theta} f$$

for some $\theta \in \mathbb{R}$, which is what we set out to prove. \square

The divide-overlap-and-conquer strategy is readily extendable to the multi-part setting.

Corollary 2.5. *Consider the multi-part ptychography (4) and (5). Suppose that the assumptions of proposition 2.2 hold and that for every $\mathcal{M}^{t_1} \cap \text{supp}(f) \neq \emptyset, \mathbf{t}_1 \in \mathcal{T}$ there is another $\mathbf{t}_2 \in \mathcal{T}, \mathbf{t}_2 \neq \mathbf{t}_1$, such that*

$$|\mathcal{M}^{t_1} \cap \mathcal{M}^{t_2} \cap \text{supp}(f)| \geq 2. \quad (16)$$

Then with probability one f is determined uniquely, up to a constant phase factor for each connected component of f , by the ptychographic data

$$\{|\Phi^t(\mu^t \odot f^t)| : \mathbf{t} \in \mathcal{T}\}. \quad (17)$$

The constant phase factors of individual connected components becomes the same for the whole object, if

$$\bigcup \{\mathcal{M}^t : \mathcal{M}^t \cap f \neq \emptyset, \mathbf{t} \in \mathcal{T}\} \text{ is a connected set.} \quad (18)$$

Remark 2.6. Clearly the result still holds when some of the shifted masks do not intersect with the object. This has a practical relevance as the support of the object is often not precisely known and some illuminations can totally miss the object. Of course, these illuminations produce no useful information and should be discarded.

When the condition (18) fails, the whole ptychographic problem breaks up into a set of separate independent subproblems, each with its ptychographic data corresponding to a connected component of $\bigcup \{\mathcal{M}^{\mathbf{t}} : \mathcal{M}^{\mathbf{t}} \cap f \neq \emptyset, \mathbf{t} \in \mathcal{T}\}$.

Proof. Let $\mathbf{t}_1, \mathbf{t}_2 \in \mathcal{T}$ be any pair of shifts satisfying the overlapping property (16). Then by theorem 2.3, $f^{\mathbf{t}_1} \cup f^{\mathbf{t}_2}$ is uniquely determined, up to a constant phase factor, by the data

$$\{|\Phi^{\mathbf{t}_j}(\mu^{\mathbf{t}_j} \odot f^{\mathbf{t}_j})| : j = 1, 2\}$$

with probability one where \odot denotes the Hadamard (i.e. componentwise) product. Since f is the union of all such pairs $f^{\mathbf{t}_1} \cup f^{\mathbf{t}_2}$, f is uniquely determined, up to a constant phase factor for each connected component of f , by the data (17), with probability one.

The constant phase factor for individual connected components may be different since some masks may have no intersection with the object. Under (18), however, this ambiguity can not take place. \square

3. Fixed point algorithms

To describe the reconstruction algorithms, it is most convenient to resort to the vector-matrix notation where we use \mathbb{C}^N ($N =$ the total number of pixels in the object $= n^2$) as the object space and \mathbb{C}^M ($M =$ the total number of measurement data $= m^2$) as the the data space *before* taking the modulus of the diffracted field. We use $\|\cdot\|$ to denote the vector norm as well as the Frobenius norm when the object is written as a matrix.

A phase-masked measurement gives rise to an *isometric* matrix in the non-ptychographic setting

$$(1\text{-pattern nonptychographic matrix}) \quad A^* = c\Phi \text{diag}\{\mu\}, \quad (19)$$

where the constant c is chosen to normalize A^* such that $AA^* = I$. The 2-pattern ptychography matrix A^* can be written as

$$(2\text{-pattern ptychography matrix}) \quad A^* = c \begin{bmatrix} \Phi^0 \text{diag}\{\mu^0\} & 0 \\ 0 & \Phi^{\mathbf{t}} \text{diag}\{\mu^{\mathbf{t}}\} \end{bmatrix} \quad (20)$$

where the first and second mask domains overlap due to the nature of a ptychographic scheme.

The propagation matrix A^* for multi-part ptychography is constructed analogous to (20) by stacking $\Phi^{\mathbf{t}} \text{diag}\{\mu^{\mathbf{t}}\}$, $\forall \mathbf{t} \in \mathcal{T}$ in the proper order. For algorithmic analysis, we normalize the columns of A^* so that A^* is isometric.

Let $b \equiv |A^*f| \in \mathbb{R}^M$. For any $y \in \mathbb{C}^M$, $\text{sgn}(y) \in \mathbb{C}^M$ is defined as

$$\text{sgn}(y)[j] = \begin{cases} 1 & \text{if } y[j] = 0 \\ y[j]/|y[j]| & \text{else.} \end{cases}$$

Ptychography can be formulated as the following feasibility problem in the Fourier domain

$$\text{Find } \hat{y} \in A^* \mathcal{X} \cap \mathcal{Y}, \quad \mathcal{Y} := \{y \in \mathbb{C}^M : |y| = b\}. \quad (21)$$

Let P_1 be the projection onto $A^*\mathcal{X}$ and P_2 the projection onto \mathcal{Y} :

$$P_1y = A^*Ay, \quad P_2y = b \odot \text{sgn}(y). \quad (22)$$

The following are two of the most widely used iterative algorithms for solving feasibility problems.

Alternating projections

$$y^{(k+1)} = P_1P_2y^{(k)}. \quad (23)$$

Douglas–Rachford algorithm

$$y^{(k+1)} = y^{(k)} + P_1(2P_2 - I)y^{(k)} - P_2y^{(k)}. \quad (24)$$

As the final output of either algorithm, the object estimate is given by $x^{(k)} = Ay^{(k)}$.

Each iteration of DR or AP makes simultaneous use of all the ptychographic data, represented by P_2 . In contrast, PIE uses the diffraction patterns one-by-one at each iteration [10, 25].

3.1. Fixed point

To accommodate the arbitrariness of the phase of zero components, we call y_* a *Fourier-domain DR fixed point* if there exists

$$u \in U = \{u = (u[i]) \in \mathbb{C}^M : |u[i]| = 1, \forall i\}$$

satisfying

$$u \in U, \quad u[j] = 1, \quad \text{whenever } y_*[j] \neq 0 \quad (25)$$

such that the DR fixed point equation holds

$$A^*A(2b \odot \text{sgn}(y_*) \odot u - y_*) = b \odot \text{sgn}(y_*) \odot u. \quad (26)$$

Note that if the sequence of iterates $y^{(k)}$ converges a limit y_∞ that has no zero component, then the limit y_∞ is a Fourier domain DR fixed point with $u \equiv 1$.

Let $x_* = Ay_*$ be the corresponding object-domain fixed point. Define another object estimate

$$\hat{x} = A(2b \odot \text{sgn}(y_*) \odot u - y_*) \quad (27)$$

for some u satisfying (25).

We have from (26)

$$A^*\hat{x} = b \odot \text{sgn}(y_*) \odot u \quad (28)$$

which implies

$$|A^*\hat{x}| = |A^*f| \quad (29)$$

$$\arg(A^*\hat{x}) = \arg(\text{sgn}(y_*) \odot u) \quad \text{on } \text{supp}(b). \quad (30)$$

Proposition 3.1. *Let x_* be the object-domain fixed point of DR. Under the assumptions of corollary 2.5 including (18), $\hat{x} = x_* = e^{i\theta}f$ for some constant $\theta \in \mathbb{R}$ almost surely.*

Proof. By theorem 2.3 and corollary 2.5, (29) implies that $\hat{x} = e^{i\theta}f$ for some constant $\theta \in \mathbb{R}$. To complete the proof, we only need to show $e^{i\theta}f = x_*$.

By (30) and the identity $\hat{x} = e^{i\theta}f$, we have

$$e^{i\theta}\text{sgn}(A^*f) = \text{sgn}(y_*) \odot u \quad \text{on } \text{supp}(b). \quad (31)$$

Substituting (31) into (27) we obtain

$$e^{i\theta}f = A(2b \odot e^{i\theta}\text{sgn}(A^*f) - y_*) = 2e^{i\theta}A(b \odot \text{sgn}(A^*f)) - Ay_* = 2e^{i\theta}f - x_*$$

where the last identity follows from the isometry of A^* and the definition of x_* . Hence $e^{i\theta}f = x_*$ as claimed. \square

Since (23) can be recast in the object domain as

$$x_{(k+1)} = A(b \odot \text{sgn}(A^*x_{(k)})),$$

we call x_* an AP fixed point if for some $u \in U$

$$x_* = A(b \odot u \odot \text{sgn}(A^*x_*)), \quad (32)$$

where u is defined through $y_* = A^*x_*$. The following result identifies any AP limit point with an AP fixed point.

Proposition 3.2. *Under the assumptions of corollary 2.5, every limit point of AP iterates $\{x^{(k)}\}$ is an AP fixed point in the sense (32).*

The proof of proposition 3.2 can be adapted from [5] *verbatim* and is omitted.

How do we distinguish the true ptychographic solution from the possibly many AP fixed points, as evidenced by numerical stagnation from random initialization?

Consider the inequality

$$\|x_*\| = \|A(\text{sgn}\{A^*x_*\} \odot b \odot u)\| \leq \|\text{sgn}\{A^*x_*\} \odot b \odot u\| = \|b\|. \quad (33)$$

Clearly $\|x_*\| = \|b\|$ holds if and only if the inequality in equation (33) is an equality, which is true only when

$$\text{sgn}\{A^*x_*\} \odot b \odot u = A^*z \quad \text{for some } z \in \mathbb{C}^n. \quad (34)$$

Since $AA^* = I$ the fixed point equation (32) implies $z = x_*$ and hence

$$\text{sgn}\{A^*x_*\} \odot b \odot u = A^*x_*.$$

Thus $b = |A^*x_*|$ implying x_* is the ptychographic solution by corollary 2.5. Therefore

Proposition 3.3. *Under the assumptions of corollary 2.5 including (18), all AP fixed points x_* satisfy $\|x_*\| \leq \|b\|$ and $x_* = e^{i\theta}f$, for some $\theta \in \mathbb{R}$, is the only AP fixed point satisfying $\|x_*\| = \|b\|$.*

While we do not have the assurance of a unique AP fixed point in comparison with DR, AP has a better convergence rate than DR as we discuss next.

3.2. Local convergence

Theorem 3.4. *Under the assumptions of corollary 2.5 including (18), let A^* be the measurement matrix and*

$$B := A \text{diag}\{\text{sgn}(A^*f)\} \in \mathbb{C}^{N \times M}. \quad (35)$$

Then

$$\gamma = \max\{\|\Im(B^*u)\| : u \in \mathbb{C}^N, u \perp if, \|u\| = 1\} < 1. \quad (36)$$

Moreover, for any given $0 < \epsilon < 1 - \gamma$, if the initial point $y^{(1)}$ is chosen such that

$$\|\alpha^{(1)}x^{(1)} - f\| := \min_{\substack{\alpha \in \mathbb{C} \\ |\alpha|=1}} \|\alpha x^{(1)} - f\| \quad \text{is sufficiently small,}$$

then we have the geometric convergence

$$DR : \min_{\substack{\alpha \in \mathbb{C} \\ |\alpha|=1}} \|\alpha x^{(k)} - f\| \leq (\gamma + \epsilon)^{k-1} \|\alpha^{(1)}x^{(1)} - f\|, \quad (37)$$

$$AP : \min_{\substack{\alpha \in \mathbb{C} \\ |\alpha|=1}} \|\alpha x^{(k)} - f\| \leq (\gamma^2 + \epsilon)^{k-1} \|\alpha^{(1)}x^{(1)} - f\|. \quad (38)$$

For $\gamma < 1$, the convergence rate γ^2 of AP is better than the convergence rate γ of DR. The proof is omitted as can be adapted to the ptychographic setting from the nonptychographic setting of [4, 5] without major changes. However, we will elaborate on the meaning of and give an estimate for (36) below.

First let us explain the connection between the matrix B in (35) and the subdifferential of the iterative map. To this end, we consider the isomorphism $\mathbb{C}^N \cong \mathbb{R}^{2N}$ via the map

$$G(x) := \begin{bmatrix} \Re(x) \\ \Im(x) \end{bmatrix}, \quad G(-ix) = \begin{bmatrix} \Im(x) \\ -\Re(x) \end{bmatrix}$$

and define the real-valued matrix

$$B = \begin{bmatrix} \Re(B) \\ \Im(B) \end{bmatrix} \in \mathbb{R}^{2N \times M}. \quad (39)$$

Denote the AP map by

$$F_{AP} = P_1 P_2$$

and the DR map by

$$F_{DR} = I + P_1(2P_2 - I) - P_2.$$

From straightforward but somewhat tedious algebra, we have

$$G(dF_{AP}(f)\xi) = G(iB\Im(B^*\xi)), \quad \forall \xi \in \mathbb{C}^N \quad (40)$$

or equivalently

$$G(-idF_{AP}(f)\xi) = BB^\top G(-i\xi), \quad \forall \xi \in \mathbb{C}^N \quad (41)$$

and

$$dF_{DR}(f)\eta = \text{diag}[\text{sgn}(A^*f)]J\text{diag}[\overline{\text{sgn}(A^*f)}]\eta \quad (42)$$

where

$$Jy = (I - B^*B)\Re(y) + iB^*B\Im(y). \quad (43)$$

Equations (40)–(43) exhibit the central role of B in the subdifferentials dF_{AP} , dF_{DR} at the point f . For detailed derivation we refer the reader to [4, 5].

Next we explain the meaning of the variational principle (36). Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{2N} \geq 0$ be the singular values of

$$\mathcal{B}^\top = [\Re(\mathcal{B})^\top \quad \Im(\mathcal{B})^\top] \in \mathbb{R}^{M \times 2N}.$$

Since the complex matrix B^* is isometric, we have $\lambda_k^2 + \lambda_{2N+1-k}^2 = 1, \forall k = 1, \dots, 2N$.

By definition, for any $x \in \mathbb{C}^N$

$$B^*x = \text{diag} \left[\overline{\text{sgn}(A^*f)} \right] A^*x$$

and hence

$$\mathcal{B}^\top G(f) = \Re[B^*f] = |A^*f|. \quad (44)$$

On the other hand, we have by isometry of A^*

$$\mathcal{B}|A^*f| = \begin{bmatrix} \Re(\mathcal{B}|A^*f|) \\ \Im(\mathcal{B}|A^*f|) \end{bmatrix} = \begin{bmatrix} \Re(AA^*f) \\ \Im(AA^*f) \end{bmatrix} = \begin{bmatrix} \Re(f) \\ \Im(f) \end{bmatrix} = G(f). \quad (45)$$

Equations (44) and (45) imply $\lambda_1 = 1$ and $G(f)$ is a leading singular vector of \mathcal{B}^\top . We can also easily verify

$$\mathcal{B}^\top G(-if) = \Im[B^*f] = 0 \quad (46)$$

and hence $G(-if)$ is a corresponding singular vector to $\lambda_{2N} = 0$.

Note again $\Im[B^*u] = \mathcal{B}^\top G(-iu)$. The orthogonality condition $iu \perp f$ is equivalent to $G(f) \perp G(-iu)$. Therefore γ defined in (36) is the second largest singular value λ_2 of \mathcal{B}^\top and admits the variational principle

$$\gamma = \max\{\|\mathcal{B}^\top u\| : u \in \mathbb{R}^{2N}, u \perp G(f), \|u\| = 1\}. \quad (47)$$

It is now straightforward to verify that the two variational principles, (36) and (47), are equivalent. We will, however, continue to use (36) which is more convenient than (47).

3.3. Spectral gap

Finally, how do we see that $\gamma < 1$?

From

$$\Im(B^*x) = \Im \left(\overline{\text{sgn}[A^*f]} \odot A^*x \right) = \sum_{j=1}^M \frac{\Re(a_j^*f)\Im(a_j^*x) - \Im(a_j^*f)\Re(a_j^*x)}{(\Re^2(a_j^*f) + \Im^2(a_j^*f))^{1/2}} \quad (48)$$

we have by the Cauchy–Schwartz inequality and the isometry of A^*

$$\|\Im(B^*x)\|^2 \leq \sum_{j=1}^M \Re^2(a_j^*x) + \Im^2(a_j^*x) = \sum_{j=1}^M |a_j^*x|^2 = \|A^*x\|^2 = \|x\|^2. \quad (49)$$

In view of (48), the inequality becomes an equality if and only if

$$\Re(a_j^*x)\Re(a_j^*f) + \Im(a_j^*x)\Im(a_j^*f) = 0, \quad \forall j = 1, \dots, N, \quad (50)$$

where a_j are the columns of A , or equivalently

$$\text{sgn}\{A^*x\} = \sigma \odot \omega_0 \quad (51)$$

where the components of σ are either 1 or -1, i.e. $\sigma[j] \in \{1, -1\}, \quad \forall j = 1, \dots, N$.

Now we recall from [4] the following uniqueness theorem for the non-ptychographic setting.

Proposition 3.5. (*Uniqueness of Fourier magnitude retrieval*) Suppose f is not a line object and let the mask μ 's phase be continuously and independently distributed. If for the matrix (19) we have

$$\angle A^* \hat{x} = \pm \angle A^* f \quad (52)$$

where the \pm sign may be pixel-dependent, then almost surely $\hat{x} = cf$ for some constant $c \in \mathbb{R}$.

This result implies that from the *Fourier phase* data, up to a \pm sign, with *each* mask separately we can identify the illuminated part of the object, up to a real constant. Now for any ptychographic scheme under the minimum overlap condition (16), the constants associated with all the masked domains must be the same. Hence (49) is a strict inequality and $\gamma < 1$.

4. Convergence rate bound

In this section, we give an estimate of γ and exhibit an explicit dependence of γ on the parameter q of the minimalist scheme introduced in section 1.2.

We divide the initial mask domain \mathcal{M}^0 , now denoted as \mathcal{M}^{00} , into four equal blocks $\mathcal{M}^{00} = \bigcup_{i,j=0}^1 \mathcal{M}_{ij}^{00}$ or in the matrix form

$$\mathcal{M}^{00} = \begin{bmatrix} \mathcal{M}_{00}^{00} & \mathcal{M}_{10}^{00} \\ \mathcal{M}_{01}^{00} & \mathcal{M}_{11}^{00} \end{bmatrix}, \quad \mathcal{M}_{ij}^{00} \in \mathbb{C}^{m/2 \times m/2}, \quad i, j = 0, 1.$$

For $m = 2n/q$, let $T^{kl} = \frac{m}{2}(k, l)$. Denoting the T^{kl} -shift of \mathcal{M}^{00} by \mathcal{M}^{kl} we have $\mathcal{M}^{kl} = \bigcup_{j,k=0}^1 \mathcal{M}_{ij}^{kl}$ where \mathcal{M}_{ij}^{kl} is the T^{kl} -shift of \mathcal{M}_{ij}^{00} . The corresponding partition of the initial mask μ^{00} and T^{kl} -shifted mask μ^{kl} can be written as

$$\mu^{00} = \begin{bmatrix} \mu_{00}^{00} & \mu_{10}^{00} \\ \mu_{01}^{00} & \mu_{11}^{00} \end{bmatrix}, \quad \mu^{kl} = \begin{bmatrix} \mu_{00}^{kl} & \mu_{10}^{kl} \\ \mu_{01}^{kl} & \mu_{11}^{kl} \end{bmatrix}.$$

For convenience, we consider the periodic boundary condition on the whole object domain, i.e.

$$\mathcal{M}_{10}^{q-1,l} = \mathcal{M}_{00}^{0l}, \quad \mathcal{M}_{11}^{q-1,l} = \mathcal{M}_{01}^{0l}, \quad (53)$$

$$\mathcal{M}_{01}^{k,q-1} = \mathcal{M}_{00}^{k0}, \quad \mathcal{M}_{11}^{k,q-1} = \mathcal{M}_{10}^{k0}, \quad (54)$$

$$\mu_{10}^{q-1,l} = \mu_{00}^{0l}, \quad \mu_{11}^{q-1,l} = \mu_{01}^{0l}, \quad (55)$$

$$\mu_{01}^{k,q-1} = \mu_{00}^{k0}, \quad \mu_{11}^{k,q-1} = \mu_{10}^{k0}, \quad (56)$$

for all $j, k = 1, \dots, q-1$.

Accordingly, we divide the object f into q^2 non-overlapping blocks

$$f = \begin{bmatrix} f_{11} & \cdots & f_{1q} \\ \vdots & \vdots & \vdots \\ f_{q1} & \cdots & f_{qq} \end{bmatrix}, \quad f_{ij} \in \mathbb{C}^{m/2 \times m/2}. \quad (57)$$

Let the ODFT Φ^{kl} defined on \mathcal{M}^{kl} be divided into four equal blocks

$$\Phi^{kl} = \begin{bmatrix} \Phi_{00}^{kl} & \Phi_{10}^{kl} \\ \Phi_{01}^{kl} & \Phi_{11}^{kl} \end{bmatrix}$$

where each $\Phi_{ij}^{kl} : \mathbb{C}^{m/2 \times m/2} \rightarrow \mathbb{C}^{2m \times 2m}$ is a rank-3 tensor defined on \mathcal{M}_{ij}^{kl} and normalized such that

$$\Phi_{ij}^{kl*} \Phi_{i'j'}^{kl} = \frac{\delta_{ij,i'j'}}{4} I_{m/2 \times m/2}, \quad i, j, i', j' = 0, 1. \quad (58)$$

Analogous to (57) the diffracted field $h = A^*f$ can be partitioned into $q \times q$ blocks, $[h_{kl}]$, where

$$h_{kl} = \sum_{i,j=0}^1 \Phi_{ij}^{k-1,l-1} (\mu_{ij}^{k-1,l-1} \odot f_{i+k,j+l}), \quad k, l = 1, \dots, q$$

where $f_{i+k,j+l}$ is cyclically defined with respect to the subscript.

Proposition 4.1. For the minimalist scheme, γ defined in (36) satisfies

$$\gamma > 1 - C/q^2 \quad (59)$$

for some constant C depending on f , but independent of q .

Remark 4.2. The derivation of the bound (59) does not assume a random mask and is valid for the minimalist scheme with any mask, so it is a bound for the worst case scenario. However, our numerical results with deterministic as well as random masks appear to be consistent with the prediction of worsening performance with large q , which we believe has to do with diminishing diversity of the mask as q becomes large.

Proof. For simplicity, we assume $\|f\| = 1$. Analogous to (36), we have from (46) the variational principle

$$\lambda_{2N-1} = \min\{\|\Im(B^*g)\| : g \in \mathbb{C}^{n \times n}, g \perp f, \|g\| = 1\} \quad (60)$$

Denote $\omega = \text{sgn}(A^*f)$ and consider the test function for (60)

$$g = \begin{bmatrix} v_1 f_{11} & \dots & v_1 f_{1q} \\ \vdots & \vdots & \vdots \\ v_q f_{q1} & \dots & v_q f_{qq} \end{bmatrix} \in \mathbb{C}^{n \times n}$$

where

$$v_j = a \sin\left(\frac{2\pi j}{q} - c\right), \quad j = 1, \dots, q, \quad (61)$$

for some real constants a, c to be selected.

Let f_j be the j -th row or column of (57). The orthogonality condition $g \perp f$ leads to

$$0 = \sum_{i=1}^q v_i \sum_{j=1}^q \|f_{ij}\|^2 = \sum_{i=1}^q v_i \|f_i\|^2 = a \sum_{j=1}^q \|f_j\|^2 \sin(2\pi j/q - c) =: p(c). \quad (62)$$

That is, c needs to be a real root of p . Since $p(0) = -p(\pi)$, the existence of a root $c \in [0, \pi]$ follows from the intermediate value theorem. On the other hand, to satisfy $\|g\| = 1$ we need

$$1 = \sum_{j=1}^q v_j^2 \|f_j\|^2 = a^2 \sum_{j=1}^q \sin^2(2\pi j/q - c) \|f_j\|^2 \quad (63)$$

which implies

$$a^2 = \left(\sum_{j=1}^q \sin^2(2\pi j/q - c) \|f_j\|^2 \right)^{-1}. \quad (64)$$

Write

$$A^* g = \begin{pmatrix} h_{11} & \dots & h_{1q} \\ \vdots & \vdots & \vdots \\ h_{q1} & \dots & h_{qq} \end{pmatrix},$$

where

$$h_{kl} = \sum_{i,j=0}^1 \Phi_{ij}^{kl} (\mu_{ij}^{kl} \odot f_{i+k,j+l}) v_{k+j}, \quad k, l = 1, \dots, q.$$

Likewise, we have

$$\omega = \operatorname{sgn}(A^* f) = \begin{pmatrix} \omega_{11} & \dots & \omega_{1q} \\ \vdots & \vdots & \vdots \\ \omega_{q1} & \dots & \omega_{qq} \end{pmatrix}, \quad \omega_{kl} = \operatorname{sgn} \left\{ h_{kl} = \sum_{i,j=0}^1 \Phi_{ij}^{kl} (\mu_{ij}^{kl} \odot f_{i+k,j+l}) \right\}.$$

To calculate $\Im(B^* g) = \Im(\bar{\omega} \odot A^* g)$, we introduce

$$u_{ij} = \Im \left[\bar{\omega}_{ij} \odot (\Phi_{00}^{ij} \mu_{00}^{ij} \odot f_{ij} + \Phi_{10}^{ij} \mu_{10}^{ij} \odot f_{i+1,j}) \right] \in \mathbb{R}^{4n/q \times 4n/q} \quad (65)$$

$$u'_{ij} = \Im \left[\bar{\omega}_{ij} \odot (\Phi_{01}^{ij} \mu_{01}^{ij} \odot f_{i,j+1} + \Phi_{11}^{ij} \mu_{11}^{ij} \odot f_{i+1,j+1}) \right] \in \mathbb{R}^{4n/q \times 4n/q}. \quad (66)$$

Note that

$$u_{ij} + u'_{ij} = \Im \left[\bar{\omega}_{ij} \odot \sum_{k,l=0}^1 \Phi_{kl}^{ij} (\mu_{kl}^{ij} \odot f_{i+k,j+l}) \right] = \Im \left[\sum_{k,l=0}^1 \Phi_{kl}^{ij} (\mu_{kl}^{ij} \odot f_{i+k,j+l}) \right] = 0$$

and hence

$$\|\Im(\bar{\omega}_{ij} \odot h_{ij})\|^2 = \|u_{ij} v_i + u'_{ij} v_{i+1}\|^2 = \|u_{ij}\|^2 (v_i - v_{i+1})^2. \quad (67)$$

Let c_{ij}^{kl} be the norm of the mapping $f_{ij} \rightarrow \Im(\bar{\omega}_{ij} \odot F_{kl}^* (\mu_{kl}^{ij} \odot f_{i+k,j+l}))$. By (58) we have $c_{ij}^{kl} \in [0, 1/2]$. Thus,

$$\sum_{j=1}^q \|u_{ij}\|^2 \leq \sum_{j=1}^q ((c_{ij}^{00})^2 \|f_{ij}\|^2 + (c_{ij}^{10})^2 \|f_{i+1,j}\|^2) \leq \frac{1}{2} \|f_i\|^2. \quad (68)$$

Combining (67), (61), (63), (64) and (68) we have the following calculation

$$\begin{aligned}
 \lambda_{2N-1}^2 &\leq \|\mathfrak{S}(\bar{\omega} \odot (A^*g))\|^2 \\
 &= \sum_{i=1}^q \sum_{j=1}^q \|\mathfrak{S}(\bar{\omega}_{ij} \odot h_{ij})\|^2 \\
 &= \sum_{i=1}^q (v_i - v_{i+1})^2 \sum_{j=1}^q \|u_{ij}\|^2 \\
 &\leq \frac{a^2}{2} \sum_{i=1}^q \|f_i\|^2 |\sin(2\pi i/q - c) - \sin(2\pi(i+1)/q - c)|^2 \\
 &\leq \frac{\sum_{i=1}^q \|f_i\|^2 |\sin(2\pi i/q - c) - \sin(2\pi(i+1)/q - c)|^2}{2 \sum_{j=1}^q \sin^2(2\pi j/q - c) \|f_j\|^2} \\
 &= 4 \sin^2(\pi/q) \frac{\sum_{j=1}^q \|f_j\|^2 \cos^2(2\pi(j+0.5)/q - c)}{2 \sum_{j=1}^q \|f_j\|^2 \sin^2(2\pi j/q - c)}. \tag{69}
 \end{aligned}$$

Note that

$$\begin{aligned}
 \frac{\sum_{j=1}^q \cos^2(2\pi(j+0.5)/q - c) \|f_j\|^2}{2 \sum_{j=1}^q \sin^2(2\pi j/q - c) \|f_j\|^2} &\leq \frac{\max_j \|f_j\|^2 \sum_{j=1}^q \cos^2(2\pi j/q - c + \pi/q)}{2 \min_j \|f_j\|^2 \sum_{j=1}^q \sin^2(2\pi j/q - c)} \\
 &\leq \frac{\max_j \|f_j\|^2}{2 \min_j \|f_j\|^2} := c
 \end{aligned}$$

and hence

$$\lambda_{2N-1}^2 \leq 4c \sin^2(\pi/q).$$

The desired result then follows from the identify $\lambda_2^2 + \lambda_{2N-1}^2 = 1$. \square

5. Numerical experiments

A primary purpose of our numerical experiments is to find out how q affects the numerical reconstruction and propose a practical guideline for using the minimalist scheme. We also want to see how the complexities of the mask and the object affect numerical performance. Finally, we want to test how robust the minimalist scheme is with respect to measurement noise.

As pointed out above, DR has the true solution as the unique fixed point in the object domain (proposition 3.1) while AP has a better convergence rate than DR (theorem 3.4). A natural way to combine their strengths is to use DR as the initialization method for AP. We choose AP and DR as the building blocks of our reconstruction algorithm because of the proven fixed point and convergence properties and also because of no adjustable parameters to be tuned to optimize the performance as in other algorithms [14, 30]. In other words, DR + AP provides a reasonable platform for evaluating the minimalist scheme and answering the aforementioned questions. For other reasons for choosing DR for ptychographic reconstruction, see [27].

We do not claim that this reconstruction method yields the best performance. Quite the contrary, our approach can be easily improved, for example, by initializing AP with the DR

iterate of the *least residual* within a given number of iterations instead of the last iterate (recall that the residual is not monotonically decreasing with the DR iteration).

Our test image f is randomly phased phantom (RPP): the phantom (figure 1(a)) with phase at each pixel being independent and uniformly distributed over a specific range, referred to as the phase range hereafter. RPP is chosen for two reasons: (i) the core image is surrounded by dark pixels and the loose support makes RPP more challenging to reconstruct than an image of a tight support; (ii) the adjustable phase range is a convenient way for controlling the object complexity.

We use the relative error (RE) and residual (RR) as figures of merit for the recovered image \hat{f} :

$$\begin{aligned} \text{RE}(\hat{f}) &= \min_{\alpha \in \mathbb{R}} \frac{\|f - e^{i\alpha\hat{f}}\|}{\|f\|}, \\ \text{RR}(\hat{f}) &= \frac{\|b - |A^*\hat{f}|\|}{\|b\|}. \end{aligned}$$

5.1. Random and Fresnel masks

We consider two kinds of random masks $e^{i\theta(\mathbf{n})}$ where $\theta(\mathbf{n})$ are either independent, identically distributed (i.i.d.) or ℓ -correlated uniform random variables on $[0, 2\pi]$, where ℓ is the correlation length.

The correlated random mask is produced by convolving the i.i.d. mask with the characteristic function of the set $\{(k_1, k_2) \in \mathbb{Z}^2 : |\max\{|k_1|, |k_2|\}| \leq \ell/2\}$ and normalizing pixel-by-pixel to get a phase mask. The i.i.d. mask corresponds to $\ell = 1$.

We also consider the Fresnel mask with

$$\mu^0(k_1, k_2) := \exp\{i\pi\rho((k_1 - \beta_1)^2 + (k_2 - \beta_2)^2)/m\}, \quad k_1, k_2 = 1, \dots, m \quad (= 2n/q) \quad (70)$$

where $\rho, \beta_1, \beta_2 \in \mathbb{R}$ are adjustable parameters, as well as the plain mask ($\rho = 0$). The choice of β_1, β_2 has an insignificant effect on numerical reconstruction. The functional form of the Fresnel phase in (70) is dictated by our goal of keeping the angular aperture of the illumination fixed (with fixed ρ) independent of m as (70) describes a point-source illumination with both the aperture (i.e. the linear size of the mask) and the distance to the object proportional to the parameter m [1].

If we set the distance from the point source to the object to be mL and the object pixel size to be $\delta \times \delta$, then (70) is the Fresnel kernel with $\rho = \delta^2/(\lambda L)$ where λ is the wavelength. For a different ρ , we imagine varying L while keeping λ and δ fixed. The larger L is, the smaller ρ is and hence the slower the mask phase modulation is. With the minimalist scheme described in section 1.2 and the Fresnel mask (70) with ρ fixed, the total space-bandwidth product of the mask and the total number of measurement data are fixed as q varies. Figures 1(b) and (c) shows the real part of the Fresnel mask at two different ρ .

In the same spirit, in the case of random mask, we let the correlation length ℓ be proportional to m as q varies when we turn to the noisy case (figures 5(a) and (d)). Figures 1(d)–(f) shows three examples of correlated masks with the same ratio $m/\ell = 4$.

5.2. Twin images with the Fresnel mask

As proved in appendix, for integer-valued ρ , $g := Qf \odot Q\mu \odot \bar{\mu}$ produces the same ptychographic data set as f where Q is the conjugate inversion operation. Not surprisingly, the twin-image ambiguity degrades the quality of reconstruction.

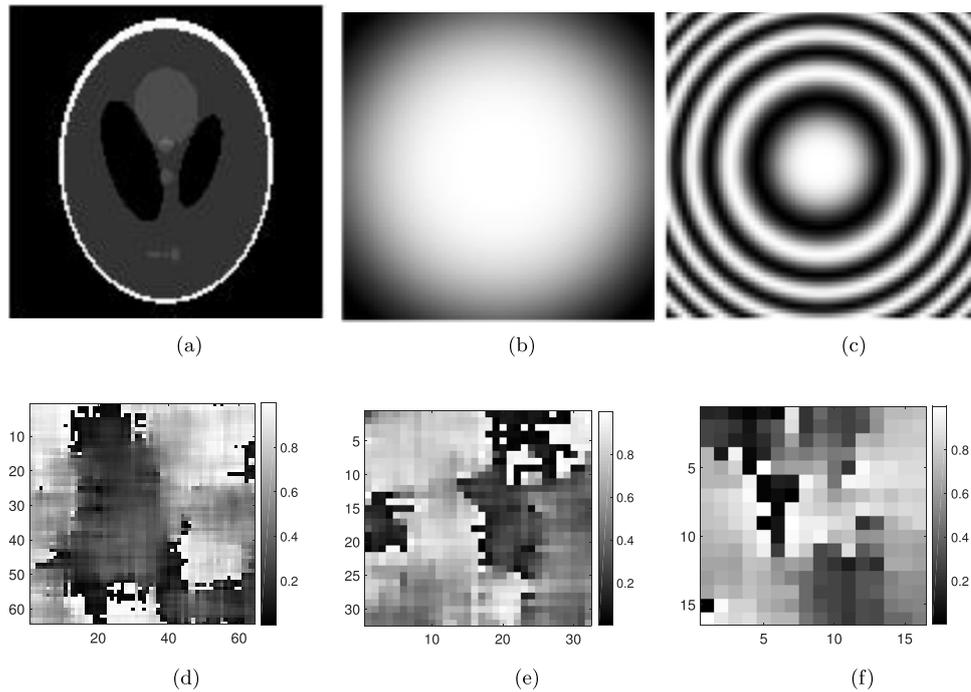


Figure 1. (a) The modulus of 128×128 RPP; (b), (c) The modulus of the real part of the respective Fresnel masks; (d)–(f) The phases of various correlated random masks in the unit of 2π . (a) Modulus of RPP. (b) $\rho = \frac{3}{25\pi} \approx 0.038$. (c) $\rho = \frac{6}{5\pi} \approx 0.38$. (d) $\ell = 16, m = 64$. (e) $\ell = 8, m = 32$. (f) $\ell = 4, m = 16$. Reproduced from [26]. [CC BY 3.0](#).

Figure 2 shows that the error spikes at integer-valued ρ . For $q > 2$, the spikes are much smaller than those of $q = 2$. As q increases from 2 to 6, the order of magnitude of fluctuation from peak to valley decreases from more than four orders of magnitude to about two or less. To avoid twin-like images, we choose irrational values of ρ .

5.3. Effect of the mask

Figure 3 shows RE versus 100 AP iterations after the DR initialization with the i.i.d. mask and two Fresnel masks for RPP of various phase ranges (legend). We see that the i.i.d. mask produces the best initialization and the fastest convergence rate and that the Fresnel mask of a larger ρ produces a better initialization and a better convergence rate than the Fresnel mask of a smaller ρ for all phase ranges.

5.4. Effect of q

Figure 4 shows RE versus 100 AP iterations after the DR initialization with an i.i.d. random mask and two Fresnel masks with various q . Interestingly, we observe that a mask of higher

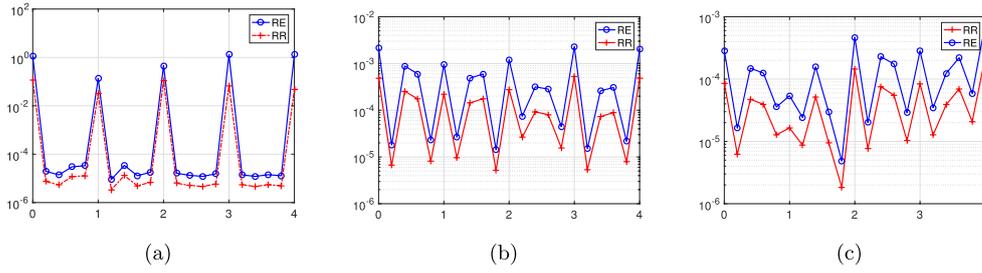


Figure 2. RE and RR on the semi-log scale versus the parameter ρ of the Fresnel mask with 200 FDR iterations followed by 100 AP iterations for 60×60 RPP of the full phase range $[0, 2\pi]$. (a) $q = 2$. (b) $q = 4$. (c) $q = 6$.

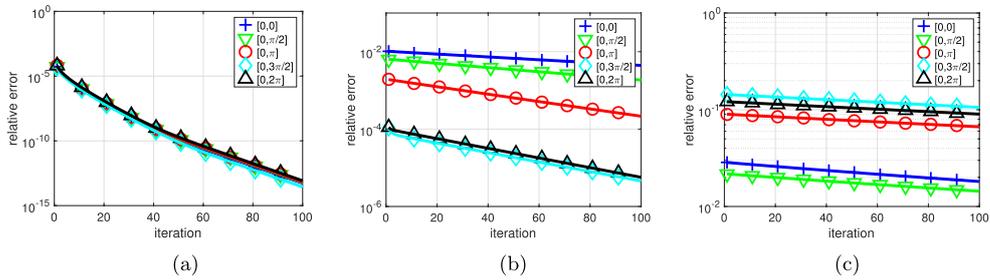


Figure 3. RE on the semi-log scale for 60×60 RPP of various object phase ranges (legend) versus 100 AP iterations with initialization given by 300 DR iterations with $q = 4$. (a) i.i.d. mask. (b) Fresnel mask with $\rho = \frac{6}{5\pi}$. (c) Fresnel mask with $\rho = \frac{3}{25\pi}$.

complexity (random or larger ρ) works better with a smaller value of q . But q is as large as 32, the results are always poor.

5.5. Effect of noise

To introduce both phase and magnitude noises to our signal model, we add complex Gaussian noise to A^*f before taking the modulus as data

$$b = |A^*f + z|$$

where $z \in \mathbb{C}^N$ is an i.i.d. circularly symmetric complex Gaussian random vector. The size of the noise is measured in terms of the noise-to-signal ratio (NSR)

$$\text{NSR} = \frac{\|b - |A^*f|\|}{\|A^*f\|}. \quad (71)$$

We also test the effect of increasing the overlap of adjacent masks from 50% to 75%: With the same relation $m = 2n/q$, 75% overlap between adjacent masks corresponds to $4q^2$ diffraction patterns.

Figure 5 shows RE versus NSR for various masks and q with (top) 50% and (bottom) 75% overlap. The correlated masks used are shown in figures 1(d)–(f) with $m/\ell = 4$ so that the complexity of the mask is independent of q . We see that the performance for the random

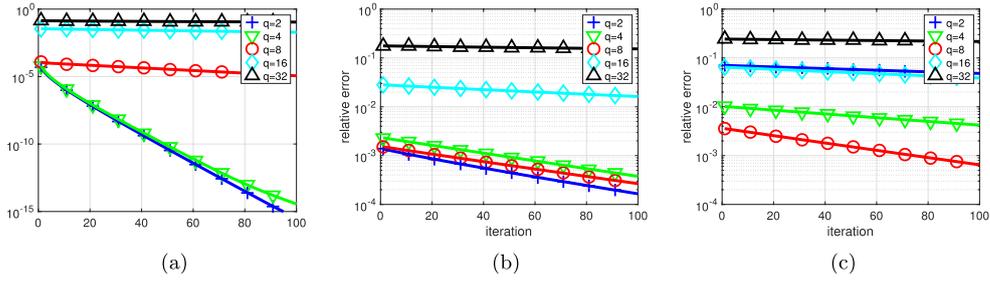


Figure 4. RE on the semi-log scale for the 128×128 RPP of phase range $[0, 2\pi]$ versus 100 AP iterations after initialization given by 300 DR iterations with various q . (a) i.i.d. mask (b) Fresnel mask with $\rho = \frac{6}{5\pi}$. (c) Fresnel mask with $\rho = \frac{3}{25\pi}$.

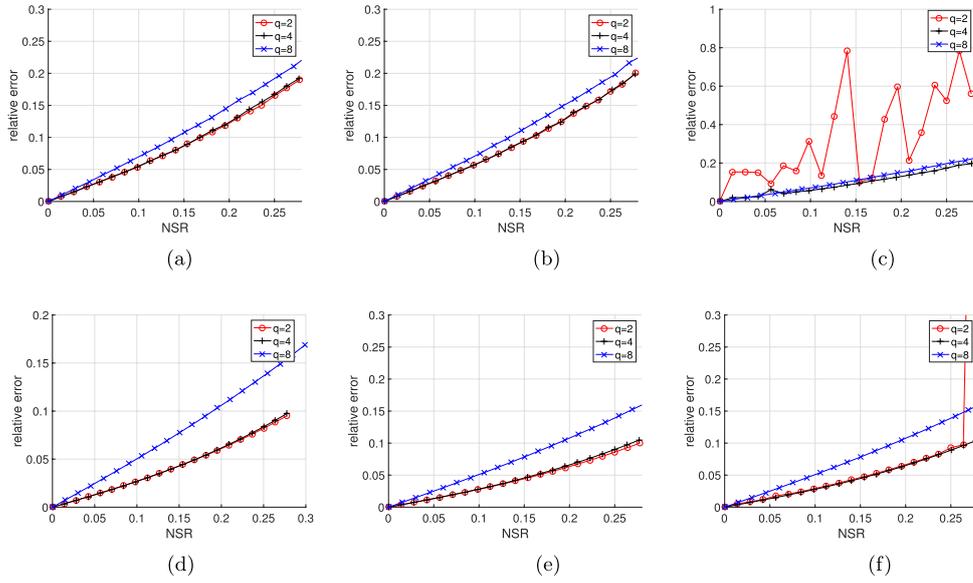


Figure 5. RE versus NSR after 300 FDR iterations followed by 200 AP iterations for 64×64 RPP of phase range $[0, 2\pi]$ with (a)–(c) 50% and (d)–(f) 75% overlap between adjacent masks. The random masks for (a) & (d) have a correlation length $\ell = m/4$ as shown in figures 1(d)–(f). (a) Random with 50% overlap. (b) Fresnel with $\rho = \frac{6}{5\pi}$. (c) Fresnel with $\rho = \frac{3}{25\pi}$. (d) Random with 75% overlap. (e) Fresnel with $\rho = \frac{6}{5\pi}$. (f) Fresnel with $\rho = \frac{3}{25\pi}$.

masks is similar to that for the Fresnel mask with $\rho = \frac{6}{5\pi}$ as well as to that for the Fresnel mask with $\rho = \frac{3}{25\pi}$ and $q = 4, 8$.

For up to 25% NSR, the RE-NSR curves in figure 5 are roughly straight lines of a slope less than 1, with $q = 4$ the best performing value across the board. The violent fluctuations in (c) for $q = 2$ indicates non-convergent behaviors consistent with figure 4(c) with $q = 2$.

Figures 5(d)–(f) shows that the four times number of data with 75% overlap predictably result in a significantly reduced RE, especially for $q = 2, 4$.

6. Conclusion and discussion

In the present work, we have proved the uniqueness theorem (theorem 2.3, corollary 2.5) for any ptychographic scheme with an independent random mask under the minimum overlap condition (16).

We have also given a local geometric convergence analysis for AP and DR algorithms (theorem 3.4). We have shown that DR has a unique fixed point in the object domain (proposition 3.1) and given a simple criterion for distinguishing the true solution among possibly many fixed points of AP (proposition 3.3).

At each iteration of DR or AP, our reconstruction method makes simultaneous use of all the ptychographic data, in what is called the global approach [12, 27]. The other approach called the sequential approach uses the diffraction patterns one-by-one iteratively for reconstruction. For example, the standard ptychographic iterative engine [10, 25] is a sequential version of AP when there is no mask estimation (as in the present setting). A comparison of the two approaches in the setting of Fourier ptychography can be found in [31, 33]. The sequential approach is usually more difficult to analyze. A convergence analysis of the sequential AP given in [5] for conventional phase retrieval (with two coded diffraction patterns) proved that the sequential AP has a better convergence rate than the global AP but is less robust when the noise level is high. We do not know if the convergence analysis in [5] can be extended to the ptychographic setting.

We have proposed a minimalist lattice scheme parametrized by $q = 2n/m$ where m is the number of mask pixels in each direction and given a lower bound on the geometric rate of convergence (proposition 4.1). The bound $\gamma > 1 - C/q^2$ predicts a poor performance for the minimalist scheme with large q which is confirmed by our numerical experiments. We have shown that the twin-image ambiguity can arise in ptychography with certain Fresnel illuminations and degrade the numerical reconstruction (appendix).

From our numerical experiments, we have found that (i) the mask of higher complexity (e.g. random masks or the Fresnel mask of a larger ρ) produces faster convergence in reconstruction than the mask of lower complexity (e.g. the Fresnel mask of a smaller ρ); (ii) the best-performing value of q for a mask of higher complexity is smaller than that for a mask of lower complexity; (iii) ptychographic reconstruction with medium values of q (e.g. $q \in [4, 8]$) is robust with respect to measurement noise regardless of the mask used, with the ratio of RE to NSR less than unity; (iv) increased overlap between adjacent masks generally reduces the reconstruction error as expected.

We have not addressed the important question in ptychography about retrieving the mask and the object simultaneously without knowing precisely the mask function (blind ptychography). This is much more subtle than one addressed in the present work.

For example, in the notation of appendix, if the (deterministic or random) mask μ used in the measurement is unknown and to be recovered simultaneously with the object x , then the ptychographic data set can not distinguish the true solution (x, μ) from (x', μ') defined by

$$x'_j = x_j \odot \epsilon, \quad \mu'_j = \mu_j \odot \epsilon^{-1}, \quad \forall j = 1, 2, 3, 4$$

for any non-vanishing $\epsilon \in \mathbb{C}^{m/2 \times m/2}$. This and other more sophisticated counterexamples point to the subtlety of the uniqueness question for blind ptychography, which is often glossed over in the literature.

To this end, we have previously proved the simultaneous determination of a roughly known mask and the object for the conventional setting of phase retrieval from coded diffraction patterns [9]. Based on [9] and the present work, we will turn to blind ptychography in a forthcoming paper.

Acknowledgments

The research of A Fannjiang is supported in part by the US National Science Foundation grant DMS-1413373.

Appendix. Twin image ambiguity with a Fresnel mask

We give a sufficient condition for the existence of twin-like images for the Fresnel masks with $q = 2, \rho \in \mathbb{Z}$, which we believe explain the spikes in figure 2(a).

Let $Q_m x$ be the conjugate inversion of $x \in \mathbb{C}^{m \times m}$, i.e. $(Q_m x)_{ij} = \bar{x}_{m+1-i, m+1-j}$. For an even integer m , write

$$x = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}, \quad x_j \in \mathbb{C}^{m/2 \times m/2}, \quad j = 1, 2, 3, 4,$$

and we have

$$Q_m x = \begin{bmatrix} Q_{m/2} x_4 & Q_{m/2} x_3 \\ Q_{m/2} x_2 & Q_{m/2} x_1 \end{bmatrix}.$$

For ease of notation, we will omit writing the subscript in Q .

Proposition A.1. Let $\rho \in \mathbb{Z}$ and $\mu \in \mathbb{C}^{m \times m}$ be the Fresnel mask with the elements

$$\exp\{i\pi\rho((k_1 - \beta_1)^2 + (k_2 - \beta_2)^2)/m\}, \quad k_1, k_2 = 1, \dots, m. \quad (\text{A.1})$$

For an even integer m , the matrix

$$\overline{Q\mu} \odot \mu := \begin{pmatrix} h_1 & h_2 \\ h_3 & h_4 \end{pmatrix}, \quad h_j \in \mathbb{C}^{m/2 \times m/2}, \quad j = 1, 2, 3, 4, \quad (\text{A.2})$$

satisfies the symmetry

$$h_1 = h_4 = \sigma h_2 = \sigma h_3, \quad \sigma = (-1)^{\rho(1+m/2)}. \quad (\text{A.3})$$

Proof. With

$$\mu = \begin{pmatrix} \mu_1 & \mu_2 \\ \mu_3 & \mu_4 \end{pmatrix}, \quad \mu_j \in \mathbb{C}^{m/2 \times m/2}, \quad j = 1, 2, 3, 4, \quad (\text{A.4})$$

we have

$$\overline{Q\mu} \odot \mu = \begin{bmatrix} \overline{Q\mu_4} \odot \mu_1 & \overline{Q\mu_3} \odot \mu_2 \\ \overline{Q\mu_2} \odot \mu_3 & \overline{Q\mu_1} \odot \mu_4 \end{bmatrix}.$$

Direct algebra with (A.1) and $\rho \in \mathbb{Z}$ gives

$$\mu_4 \odot \overline{Q\mu_1} = \mu_2 \odot \overline{Q\mu_3} \exp\{(\rho + \rho m/2)\pi i\} = \mu_3 \odot \overline{Q\mu_2} \exp\{(\rho + \rho m/2)\pi i\} = \mu_1 \odot \overline{Q\mu_4}$$

and hence the result in (A.3) if ρ is an integer. \square

Let Φ be the oversampled 2-D Fourier transform as before. We have the ambiguity of conjugate inversion (twin image):

$$|\Phi(x)| = |\Phi(Qx)|, \quad \forall x \in \mathbb{C}^{m \times m}. \quad (\text{A.5})$$

Proposition A.2. *Let $q = 2$ (hence $m = 2n/q = n$). Let h be the matrix given in (A.2) and let $y = Qx \odot \bar{h}$. If the symmetry (A.3) holds, then y and x produce the same ptychographic data set with the mask μ .*

Proof. We have

$$y = \begin{bmatrix} Qx_4 \odot \bar{h}_1 & Qx_3 \odot \bar{h}_2 \\ Qx_2 \odot \bar{h}_3 & Qx_1 \odot \bar{h}_4 \end{bmatrix}$$

and by direct substitution

$$|\Phi(y \odot \mu)| = |\Phi(Q(x \odot \mu))| = |\Phi(x \odot \mu)|. \quad (\text{A.6})$$

Let R_1 and R_2 be the reflectors defined by

$$R_1 x = \begin{bmatrix} x_2 & x_1 \\ x_4 & x_3 \end{bmatrix}, \quad R_2 x = \begin{bmatrix} x_3 & x_4 \\ x_1 & x_2 \end{bmatrix},$$

for all $x \in \mathbb{C}^{m \times m}$. These reflectors describe the periodic boundary condition in the minimalist scheme with $q = 2$.

Let

$$\begin{aligned} y' &= (QR_1 x) \odot \overline{R_1 h}, \\ y'' &= (QR_2 x) \odot \overline{R_2 h}, \\ y''' &= (QR_2 R_1 x) \odot \overline{R_2 R_1 h}. \end{aligned}$$

Analogous to (A.6) we have

$$\begin{aligned} |\Phi(y' \odot \mu)| &= |\Phi(Q(R_1 x \odot \mu))| = |\Phi(R_1 x \odot \mu)|, \\ |\Phi(y'' \odot \mu)| &= |\Phi(Q(R_2 x \odot \mu))| = |\Phi(R_2 x \odot \mu)|, \\ |\Phi(y''' \odot \mu)| &= |\Phi(Q(R_2 R_1 x \odot \mu))| = |\Phi(R_2 R_1 x \odot \mu)| \end{aligned}$$

The symmetry (A.3) implies

$$h = \sigma R_1 h = \sigma R_2 h = R_2 R_1 h = \mu \odot \overline{Q\mu}. \quad (\text{A.7})$$

Using (A.5) and (A.7) we have

$$|\Phi(x \odot \mu)| = |\Phi(Qx \odot Q\mu)| = |\Phi(Qx \odot \bar{h} \odot \mu)| = |\Phi(y \odot \mu)|. \quad (\text{A.8})$$

Similarly, we have

$$|\Phi(R_1 x \odot \mu)| = |\Phi(QR_1 x \odot Q\mu)| = |\Phi(QR_1 x \odot \overline{\gamma R_1 h} \odot \mu)| = |\Phi(y' \odot \mu)|, \quad (\text{A.9})$$

$$|\Phi(R_2 x \odot \mu)| = |\Phi(QR_2 x \odot Q\mu)| = |\Phi(QR_2 x \odot \overline{\gamma R_2 h} \odot \mu)| = |\Phi(y'' \odot \mu)|, \quad (\text{A.10})$$

$$|\Phi(R_2 R_1 x \odot \mu)| = |\Phi(Q R_2 R_1 \odot Q \mu)| = |\Phi(Q R_2 R_1 x \odot \overline{R_2 R_1 h} \odot \mu)| = |\Phi(y''' \odot \mu)|. \quad (\text{A.11})$$

The lefthand and righthand sides of equations (A.8)–(A.11) are precisely the ptychographic data set for $q = 2$ with the mask μ . \square

ORCID iDs

Albert Fannjiang  <https://orcid.org/0000-0003-3150-7392>

References

- [1] Born M and Wolf E 1999 *Principles of Optics* 7th edn (Cambridge: Cambridge University Press) pp 199–201
- [2] Bunk O, Dierolf M, Kynde S, Johnson I, Marti O and Pfeiffer F 2008 Influence of the overlap parameter on the convergence of the ptychographical iterative engine *Ultramicroscopy* **108** 481–7
- [3] Chapman H N 2010 Microscopy: a new phase for x-ray imaging *Nature* **467** 409–10
- [4] Chen P and Fannjiang A 2016 Phase retrieval with a single mask by Douglas–Rachford algorithms *Appl. Comput. Harmon. Anal.* <https://doi.org/10.1016/j.acha.2016.07.003>
- [5] Chen P, Fannjiang A and Liu G 2017 Phase retrieval with one or two coded diffraction patterns by alternating projection with the null initialization *J. Fourier Anal. Appl.* <https://doi.org/10.1007/s00041-017-9536-8>
- [6] Dierolf M, Menzel A, Thibault P, Schneider P, Kewish C M, Wepf R, Bunk O and Pfeiffer F 2010 Ptychographic x-ray computed tomography at the nanoscale *Nature* **467** 436–9
- [7] Fannjiang A 2012 Absolute uniqueness of phase retrieval with random illumination *Inverse Problems* **28** 075008
- [8] Fannjiang A and Liao W 2012 Phase retrieval with random phase illumination *J. Opt. Soc. A* **29** 1847–59
- [9] Fannjiang A and Liao W 2013 Fourier phasing with phase-uncertain mask *Inverse Problems* **29** 125001
- [10] Faulkner H M L and Rodenburg J M 2004 Movable aperture lensless transmission microscopy: a novel phase retrieval algorithm *Phys. Rev. Lett.* **93** 023903
- [11] Faulkner H M L and Rodenburg J M 2005 Error tolerance of an iterative phase retrieval algorithm for moveable illumination microscopy *Ultramicroscopy* **103** 153–64
- [12] Guizar-Sicairos M and Fienup J R 2008 Phase retrieval with transverse translation diversity: a nonlinear optimization approach *Opt. Express* **16** 7264–78
- [13] Hayes M 1982 The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform *IEEE Trans. Acoust. Speech Signal Process.* **30** 140–54
- [14] Hesse R, Luke D R, Sabach S and Tam M K 2015 Proximal heterogeneous block implicit-explicit method and application to blind ptychographic diffraction imaging *SIAM J. Imaging Sci.* **8** 426–57
- [15] Hoppe W 1969 Beugung im inhomogenen Primärstrahlwellenfeld. I. Prinzip einer Phasenmessung von Elektronenbeugungsinterferenzen *Acta Cryst. A* **25** 495
- [16] Hoppe W 1969 Beugung im inhomogenen Primärstrahlwellenfeld. III. Amplituden- und phasenbestimmung bei unperiodischen objekten *Acta Cryst. A* **25** 508
- [17] Hunt J, Driscoll T, Mrozack A, Lipworth G, Reynolds M, Brady D and Smith D R 2013 Metamaterial apertures for computational imaging *Science* **339** 310–3
- [18] Iwen M A, Viswanathan A and Wang Y 2016 Fast phase retrieval from local correlation measurements *SIAM J. Imaging Sci.* **9** 1655–88
- [19] Lipworth G, Mrozack A, Hunt J, Marks D L, Driscoll T and Brady D 2013 Metamaterial apertures for coherent computational imaging on the physical layer *J. Opt. Soc. Am. A* **30** 1603–12
- [20] Maiden A M, Humphry M J, Zhang F and Rodenburg J M 2011 Superresolution imaging via ptychography *J. Opt. Soc. Am. A* **28** 604–12

- [21] Maiden A M, Johnson D and Li P 2017 Further improvements to the ptychographical iterative engine *Optica* **4** 736–45
- [22] Maiden A M, Morrison G R, Kaulich B and Gianoncelli A 2013 Soft x-ray spectromicroscopy using ptychography with randomly phased illumination *Nat. Commun.* **4** 1669
- [23] Maiden A M and Rodenburg J M 2009 An improved ptychographical phase retrieval algorithm for diffractive imaging *Ultramicroscopy* **109** 1256–62
- [24] Maiden A M, Rodenburg J M and Humphry M J 2010 Optical ptychography: a practical implementation with useful resolution *Opt. Lett.* **35** 2585–7
- [25] Rodenburg J M and Faulkner H M L 2004 A phase retrieval algorithm for shifting illumination *Appl. Phys. Lett.* **85** 4795
- [26] Stockmar M, Cloetens P, Zanette I, Enders B, Dierolf M, Pfeiffer F and Thibault P 2013 Near-field ptychography: phase retrieval for inline holography using a structured illumination *Sci. Rep.* **3** 1927
- [27] Thibault P, Dierolf M, Bunk O, Menzel A and Pfeiffer F 2009 Probe retrieval in ptychographic coherent diffractive imaging *Ultramicroscopy* **109** 338–43
- [28] Thibault P, Dierolf M, Menzel A, Bunk O, David C and Pfeiffer F 2008 High-resolution scanning x-ray diffraction microscopy *Science* **321** 379–82
- [29] Watts C M, Shrekenhamer D, Montoya J, Lipworth G, Hunt J, Sleasman T, Krishna S, Smith D R and Padilla W J 2014 Terahertz compressive imaging with metamaterial spatial light modulators *Nat. Photon.* **8** 605–9
- [30] Wen Z, Yang C, Liu X and Marchesini S 2012 Alternating direction methods for classical and ptychographic phase retrieval *Inverse Problems* **28** 115010
- [31] Yeh L-H, Dong J, Zhong J, Tian L, Chen M, Tang G, Soltanolkotabi M and Waller L 2015 Experimental robustness of Fourier ptychography phase retrieval algorithms *Opt. Express* **23** 33214–40
- [32] Zhang X, Jiang J, Xiangli B and Arce G R 2015 Spread spectrum phase modulation for coherent x-ray diffraction imaging *Opt. Express* **23** 25034–47
- [33] Zuo C, Sun J and Chen Q 2016 Adaptive step-size strategy for noise-robust Fourier ptychographic microscopy *Opt. Express* **24** 20724–44