

# Blind Ptychography via Blind Deconvolution

Mark Philip Roach

## Abstract

Ptychography involves a sample being illuminated by a coherent, localised probe of illumination. When the probe interacts with the sample, the light is diffracted and a diffraction pattern is detected. Then the probe or sample is shifted laterally in space to illuminate a new area of the sample while ensuring there is sufficient overlap. **Far-field Ptychography** occurs when there is a large enough distance (when the Fresnel number is  $\ll 1$ ) to obtain magnitude-square Fourier transform measurements. In an attempt to remove ambiguities, masks are utilized to ensure unique outputs to any recovery algorithm are unique up to a global phase. In this paper, we assume that both the sample and the mask are unknown, and we apply blind deconvolutional techniques to solve for both. Numerical experiments demonstrate that the technique works well in practice, and is robust under noise. This chapter is comprised of three sections. Section 2 introduces far-field Fourier ptychography, and an algorithm for solving given noisy ptychographic measurements. In particular of use to us will be Theorem 2.2 which reformulates the measurements into a convolution. Section 3 explores a method for solving a blind deconvolution problem, given certain appropriate real-world assumptions. Section 4 combines the previous two sections, taking the reformulated convolutional measurements, assuming that both components are unknown, and then applying the blind deconvolution recovery algorithm. The full algorithm is stated and numerical simulations are summarized, outlining good recovery which is robust under noise.

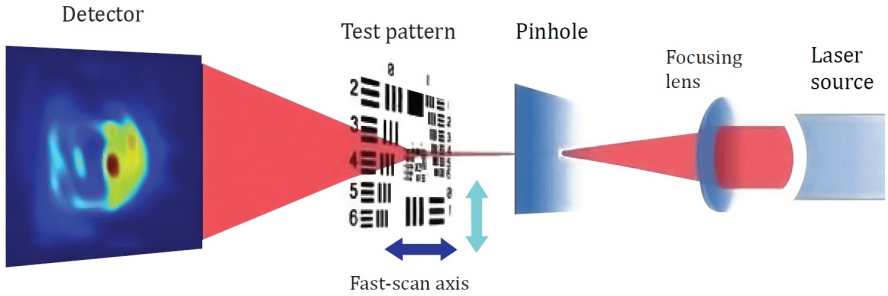
**Keywords:** Blind Ptychography, Blind Deconvolution

# 1 Introduction

Ptychography involves a sample being illuminated by a coherent, localised probe of illumination. When the probe interacts with the sample, the light is diffracted and a diffraction pattern is detected. Then the probe or sample is shifted laterally in space to illuminate a new area of the sample while ensuring there is sufficient overlap. **Far-field Ptychography** occurs when there is a large enough distance (when the Fresnel number is  $\ll 1$ ) to obtain magnitude-square Fourier transform measurements.

Ptychography was initially studied in the late 1960s ([1]), with the problem solidified in 1970 ([2]). The name "Ptychography" was coined in 1972 ([3]), after the Greek word *to fold* because the process involves an interference pattern such that the scattered waves fold into one another (coherent) Fourier diffraction pattern of the object. Initially developed to study crystalline objects under a scanning transmission electron microscope, since then the field has widened to setups such as using visible light ([4], [5], [6]), x-rays ([7], [8], [9]), or electrons ([10],[11][12]). It is benefited from being unaffected by lens-induced aberrations or diffraction effects unlike conventional lens imaging. Various types of ptychography are studied based on the optical configuration of the experiments. For instance, Bragg Ptychography ([13] [14], [15], [16]) measures strain in crystalline specimens by shifting the surface of the specimen.

Fourier ptychography ([17],[18],[19],[20]) consists of taking multiple images at a wide field-of-view then computationally synthesizing into a high-resolution image reconstruction in the Fourier domain. This results in an increased resolution compared to a conventional microscope.



**Fig. 1** [5] Experimental setup for fly-scan ptychography

## 2 Far-field Fourier Ptychography

Let  $\mathbf{x}, \mathbf{m} \in \mathbb{C}^d$  denote the unknown sample and known mask, respectively. We suppose that we have  $d^2$  noisy ptychographic measurements of the form

$$(\mathbf{Y})_{\ell,k} = |(\mathbf{F}(\mathbf{x} \circ S_k \mathbf{m}))_{\ell}|^2 + (\mathbf{N})_{\ell,k}, \quad (\ell, k) \in [d]_0 \times [d]_0, \quad (1)$$

where  $S_k, \circ, \mathbf{F} := \mathbf{F}_d$  denote  $k^{\text{th}}$  circular shift, Hadamard product, and  $d$ -dimensional discrete Fourier transform, and  $\mathbf{N}$  is the matrix of additive noise.

In this section, we will define a discrete Wigner distribution deconvolution method for recovering a discrete signal. A modified Wigner distribution deconvolution approach is used to solve for an estimate of  $\hat{\mathbf{x}}\hat{\mathbf{x}}^* \in \mathbb{C}^{d \times d}$  and then angular synchronization is performed to compute estimate of  $\hat{\mathbf{x}}$  and thus  $\mathbf{x}$ .

In Section 2.1, we introduce definitions and technical lemmas which will be of use. In particular, the decoupling lemma (Lemma 2.2) allows use to effectively 'separate' the mask and object from a convolution. In Section 2.2, these technical lemmas are applied to the ptychographic measurements to write the problem as a decoupled deconvolution problem, the blind variant of which will be studied later on. In Section 2.3, an additional Fourier transform is applied and the measurements have been rewritten to a form in which a pointwise division approach can be applied. Sub-sampled version of this theorem are also given. We then state the full algorithm for recovering the sample.

## 2.1 Properties of the Discrete Fourier Transform

We firstly define the modulation operator.

**Definition 2.1.** Given  $k \in [d]_0$ , define the **modulation operator**  $W_k : \mathbb{C}^d \rightarrow \mathbb{C}^d$  component-wise via

$$(W_k \mathbf{x})_n = x_n e^{2\pi i k n / d}, \quad \forall n \in [d]_0. \quad (2)$$

From this definition, we can develop some useful equalities which we will use in the main proofs of this section.

**Lemma 2.1. (Technical Equalities)** (Lemma 1.3.1., [21]) The following equalities hold for all  $\mathbf{x} \in \mathbb{C}^d, [\ell] \in [d]_0$ :

- (i)  $\mathbf{F}_d \hat{\mathbf{x}} = d \cdot \tilde{\mathbf{x}}$ ;
- (ii)  $\mathbf{F}_d (W_\ell \mathbf{x}) = S_{-\ell} \hat{\mathbf{x}}$ ;
- (iii)  $\mathbf{F}_d (S_\ell \mathbf{x}) = W_\ell \hat{\mathbf{x}}$ ;
- (iv)  $W_{-\ell} \mathbf{F}_d (S_\ell \tilde{\mathbf{x}}) = \hat{\mathbf{x}}$ ;
- (v)  $\overline{S_\ell \mathbf{x}} = S_{-\ell} \tilde{\tilde{\mathbf{x}}}$ ;
- (vi)  $\mathbf{F}_d \tilde{\mathbf{x}} = \mathbf{F}_d \tilde{\tilde{\mathbf{x}}}$ ;
- (vii)  $\tilde{\hat{\mathbf{x}}} = \mathbf{F}_d \tilde{\mathbf{x}}$ .

We wish to be able to convert between the convolution and the Hadamard product, so we will need this useful theorem.

**Theorem 2.1. (Discretized Convolution Theorem)** (Lemma 1.3.2., [21]) Let  $\mathbf{x}, \mathbf{y} \in \mathbb{C}^d$ . We have that

- (i)  $F_d^{-1}(\hat{\mathbf{x}} \circ \hat{\mathbf{y}}) = \mathbf{x} *_d \mathbf{y}$ ;  
(ii)  $(\mathbf{F}_d \mathbf{x}) *_d (\mathbf{F}_d \mathbf{y}) = d \cdot \mathbf{F}_d(\mathbf{x} \circ \mathbf{y})$ .

Currently, the measurements we are dealing with will be having the specimen and the mask intertwined. We introduce the decoupling lemma to essentially detangle the two.

**Lemma 2.2. (Decoupling Lemma)** (Lemma 1.3.3., [21])

Let  $\mathbf{x}, \mathbf{y} \in \mathbb{C}^d$ ,  $\ell, k \in [d]_0$ . Then

$$\left( (\mathbf{x} \circ S_{-\ell} \mathbf{y}) *_d (\bar{\bar{\mathbf{x}}} \circ S_{\ell} \bar{\bar{\mathbf{y}}}) \right)_k = \left( (\mathbf{x} \circ S_{-k} \bar{\mathbf{x}}) *_d (\bar{\mathbf{y}} \circ S_k \bar{\bar{\mathbf{y}}}) \right)_{\ell}. \quad (3)$$

*Proof* Let  $\mathbf{x}, \mathbf{y} \in \mathbb{C}^d$ ,  $\ell, k \in [d]_0$ . By the definitions of the circular convolution, Hadamard product and shift operator, we have that

$$\begin{aligned} \left( (\mathbf{x} \circ S_{-\ell} \mathbf{y}) *_d (\bar{\bar{\mathbf{x}}} \circ S_{\ell} \bar{\bar{\mathbf{y}}}) \right)_k &= \sum_{n=0}^{d-1} (\mathbf{x} \circ S_{-\ell} \mathbf{y})_n ((\bar{\bar{\mathbf{x}}} \circ S_{\ell} \bar{\bar{\mathbf{y}}})_{k-n} \\ &= \sum_{n=0}^{d-1} x_n y_{n-\ell} \bar{\bar{x}}_{k-n} \bar{\bar{y}}_{\ell+k-n} \\ &= \sum_{n=0}^{d-1} x_n \bar{\bar{x}}_{n-k} \bar{\mathbf{y}}_{\ell-n} \bar{\bar{\mathbf{y}}}_{k+\ell-n} \\ &= \sum_{n=0}^{d-1} (\mathbf{x} \circ S_{-k} \bar{\mathbf{x}})_n ((\bar{\mathbf{y}} \circ S_k \bar{\bar{\mathbf{y}}})_{\ell-n} \\ &= \left( (\mathbf{x} \circ S_{-k} \bar{\mathbf{x}}) *_d (\bar{\mathbf{y}} \circ S_k \bar{\bar{\mathbf{y}}}) \right)_{\ell}. \end{aligned} \quad (4)$$

□

Lastly before entering the main part of this subsection, we need a lemma involving looking at how the Fourier squared magnitude measurements will relate to a convolution.

**Lemma 2.3.** Let  $\mathbf{x} \in \mathbb{C}^d$ . We have that

$$|\mathbf{F}_d \mathbf{x}|^2 = \mathbf{F}_d(\mathbf{x} *_d \bar{\bar{\mathbf{x}}}). \quad (5)$$

*Proof* Let  $\mathbf{x} \in \mathbb{C}^d$ . Then we have that

$$|\mathbf{F}_d \mathbf{x}|^2 = (\mathbf{F}_d \mathbf{x}) \circ \overline{(\mathbf{F}_d \mathbf{x})} = (\mathbf{F}_d \mathbf{x}) \circ (\mathbf{F}_d \bar{\bar{\mathbf{x}}}) = \mathbf{F}_d(\mathbf{x} *_d \bar{\bar{\mathbf{x}}}). \quad (6)$$

□

## 2.2 Discretized Wigner Distribution Deconvolution

We now prove the Discretized Wigner Distribution Deconvolution theorem which will allow us to convert the measurements into a form in which we can algorithmically solve.

**Theorem 2.2.** (Lemma 1.3.5., [21]) Let  $\mathbf{x}, \mathbf{m} \in \mathbb{C}^d$  denote the unknown specimen and known mask, respectively. Suppose we have  $d^2$  noisy ptychographic measurements of the form

$$(\mathbf{y}_\ell)_k = \left| \sum_{n=0}^{d-1} x_n m_{n-\ell} e^{-2\pi i n k / d} \right|^2 + (\mathbf{N})_{\ell, k}, \quad (\ell, k) \in [d]_0 \times [d]_0. \quad (7)$$

Let  $\mathbf{Y} \in \mathbb{R}^{d \times d}$ ,  $\mathbf{N} \in \mathbb{C}^{d \times d}$  be the matrices whose  $\ell^{\text{th}}$  column is  $\mathbf{y}_\ell, \mathbf{N}_\ell$  respectively. Then for any  $k \in [d]_0$ ,

$$\left( \mathbf{Y}^T \mathbf{F}_d^T \right)_k = d \cdot (\mathbf{x} \circ S_k \bar{\mathbf{x}}) *_d (\tilde{\mathbf{m}} \circ S_{-k} \tilde{\tilde{\mathbf{m}}}) + \left( \mathbf{N}^T \mathbf{F}_d^T \right)_k. \quad (8)$$

*Proof* Let  $\ell \in [d]_0$ . We have that

$$\mathbf{y}_\ell = |F_d(\mathbf{x} \circ S_{-\ell} \mathbf{m})|^2 + \mathbf{N}_\ell = F_d \left( (\mathbf{x} \circ S_{-\ell} \mathbf{m}) *_d (\bar{\mathbf{x}} \circ S_\ell \tilde{\tilde{\mathbf{m}}}) \right) + \mathbf{N}_\ell. \quad (9)$$

Taking Fourier transform of both sides at  $k \in [d]_0$  and using that  $\mathbf{F}_d \hat{\mathbf{x}} = d \cdot \bar{\mathbf{x}}$  yields

$$\begin{aligned} (\mathbf{F}_d \mathbf{y}_\ell)_k &= d \cdot \left( (\mathbf{x} \circ S_{-\ell} \mathbf{m}) *_d (\bar{\mathbf{x}} \circ S_\ell \tilde{\tilde{\mathbf{m}}}) \right)_{-k} + (\mathbf{F}_d \mathbf{N}_\ell)_k \\ &= d \cdot \left( (\mathbf{x} \circ S_k \bar{\mathbf{x}}) *_d (\tilde{\mathbf{m}} \circ S_{-k} \tilde{\tilde{\mathbf{m}}}) \right)_\ell + (\mathbf{F}_d \mathbf{N}_\ell)_k, \end{aligned} \quad (10)$$

by previous lemma. For fixed  $\ell \in [d]_0$ , the vector  $\mathbf{F}_d \mathbf{y}_\ell \in \mathbb{C}^d$  is the  $\ell^{\text{th}}$  column of the matrix  $\mathbf{F}_d \mathbf{Y}$ , thus its transpose  $\mathbf{y}_\ell^T \mathbf{F}_d^T \in \mathbb{C}^d$  is the  $\ell^{\text{th}}$  row of the matrix  $(\mathbf{F}_d \mathbf{Y})^T$ . Similarly,  $(\mathbf{N})_\ell^T \mathbf{F}_d^T \in \mathbb{C}^d$  is the  $\ell^{\text{th}}$  row of  $(\mathbf{F}_d \mathbf{N})^T$ . Thus we have that

$$\left( \left( \mathbf{Y}^T \mathbf{F}_d^T \right)_k \right)_\ell = d \cdot \left( (\mathbf{x} \circ S_k \bar{\mathbf{x}}) *_d (\tilde{\mathbf{m}} \circ S_{-k} \tilde{\tilde{\mathbf{m}}}) \right)_\ell + \left( \mathbf{N}^T \mathbf{F}_d^T \right)_{k, \ell}. \quad (11)$$

Thus we have that

$$\left( \mathbf{Y}^T \mathbf{F}_d^T \right)_k = d \cdot (\mathbf{x} \circ S_k \bar{\mathbf{x}}) *_d (\tilde{\mathbf{m}} \circ S_{-k} \tilde{\tilde{\mathbf{m}}}) + \left( \mathbf{N}^T \mathbf{F}_d^T \right)_k. \quad (12)$$

□

We note that  $\mathbf{x} \circ S_k \bar{\mathbf{x}}$  is a diagonal of  $\mathbf{x} \mathbf{x}^*$ .

## 2.3 Wigner Distribution Deconvolution Algorithm

We suppose that the mask is known and the specimen is unknown. By taking an additional Fourier transform and using the discretized convolution theorem, we have these variances of the previous lemmas.

**Theorem 2.3. (Discretized Wigner Distribution Deconvolution)** Let  $\mathbf{x}, \mathbf{m} \in \mathbb{C}^d$  denote the unknown specimen and known mask, respectively. Suppose we have  $d^2$  noisy spectrogram measurements of the form

$$(\mathbf{y}_\ell)_k = \left| \sum_{n=0}^{d-1} x_n m_{n-\ell} e^{-2\pi i n k / d} \right|^2 + (\mathbf{N})_{\ell, k}, \quad (\ell, k) \in [d]_0 \times [d]_0. \quad (13)$$

Let  $\mathbf{Y} \in \mathbb{R}^{d \times d}$  be the matrix whose  $\ell^{\text{th}}$  column is  $\mathbf{y}_\ell$ . Then for any  $k \in [d]_0$

$$\mathbf{F}_d \left( \mathbf{Y}^T \mathbf{F}_d^T \right)_k = d \cdot \mathbf{F}_d(\mathbf{x} \circ S_k \bar{\mathbf{x}}) \circ \mathbf{F}_d(\bar{\mathbf{m}} \circ S_{-k} \tilde{\mathbf{m}}) + \mathbf{F}_d \left( \mathbf{N}^T \mathbf{F}_d^T \right)_k. \quad (14)$$

We also have a similar result based on the work in Appendix ??.

**Lemma 2.4. (Sub-Sampling In Frequency)** Suppose that the spectrogram measurements are collected on a subset  $\mathcal{K} \subseteq [d]_0$  of  $K$  equally spaced Fourier modes. Then for any  $\omega \in [K]_0$

$$\mathbf{F}_d \left( (\mathbf{Y}_{K,d})^T \mathbf{F}_K^T \right)_\omega = K \sum_{r=0}^{\frac{d}{K}-1} \mathbf{F}_d(\mathbf{x} \circ S_{\ell L - \alpha} \bar{\mathbf{x}}) \circ \mathbf{F}_d(\bar{\mathbf{m}} \circ S_{\alpha - \ell L} \tilde{\mathbf{m}}) + \mathbf{F}_d \left( (\mathbf{N}_{K,d})^T \mathbf{F}_K^T \right)_\omega$$

where  $\mathbf{Y}_{K,d} \in \mathbb{C}^{K \times d}$  is the matrix of sub-sampled noiseless  $K \cdot d$  measurements.

**Lemma 2.5. (Sub-Sampling In Frequency And Space)** Suppose we have spectrogram measurements collected on a subset  $\mathcal{K} \subseteq [d]_0$  of  $K$  equally spaced frequencies and a subset  $\mathcal{L} \subseteq [d]_0$  of  $L$  equally spaced physical shifts. Then for any  $\omega \in [K]_0, \alpha \in [L]_0$

$$\begin{aligned} \left( \mathbf{F}_L(\mathbf{Y}_{K,L})^T (\mathbf{F}_K^T)_\omega \right)_\alpha &= \frac{KL}{d^3} \sum_{r=0}^{\frac{d}{K}-1} \sum_{\ell=0}^{\frac{d}{L}-1} \left( \mathbf{F}_d(\hat{\mathbf{x}} \circ S_{\ell L - \alpha} \bar{\hat{\mathbf{x}}}) \right)_{\omega - rK} \left( F_d(\hat{\mathbf{m}} \circ S_{\alpha - \ell L} \tilde{\hat{\mathbf{m}}}) \right)_{\omega - rK} \\ &\quad + \left( \mathbf{F}_L(\mathbf{N}_{K,L})^T (\mathbf{F}_K^T)_\omega \right)_\alpha, \end{aligned}$$

where  $\mathbf{Y}_{K,L} \in \mathbb{C}^{K \times L}$  is the matrix of sub-sampled noiseless  $K \cdot L$  measurements.

Assume that  $\mathbf{m}$  is band-limited with  $\text{supp}(\hat{\mathbf{m}}) = [\delta]_0$  for some  $\delta \ll d$ . Then the algorithm below allows for the recovery of an estimate of  $\hat{\mathbf{x}}$  from spectrogram measurements via Wigner distribution deconvolution and angular synchronization.

**Algorithm 1** (Algorithm 1, [21]) Wigner Distribution Deconvolution Algorithm**Input:** 1)  $Y_{d,L} \in \mathbb{C}^{d \times L}$ , matrix of noisy measurements.2) Mask  $\mathbf{m} \in \mathbb{C}^d$  with  $\text{supp}(\hat{\mathbf{m}}) = [\delta]$ .3) Integer  $\kappa \leq \delta$ , so that  $2\kappa - 1$  diagonals of  $\hat{\mathbf{x}}\hat{\mathbf{x}}^*$  are estimated, and  $L = \delta + \kappa - 1$ .**Output:** An estimate  $\mathbf{x}_{est}$  of  $\mathbf{x}$  up to a global phase.

1) Perform pointwise division to compute

$$\frac{1}{d} \frac{\mathbf{F}_d \left( \mathbf{Y}^T \mathbf{F}_d^T \right)_k}{\mathbf{F}_d^{-1}(\tilde{\mathbf{m}} \circ S_{-k} \tilde{\mathbf{m}})}. \quad (15)$$

2) Invert the  $(2\kappa - 1)$  Fourier transforms above.3) Organize values from step 2 to form the diagonals of a banded matrix  $Y_{2\kappa-1}$ .4) Perform angular synchronization on  $Y_{2\kappa-1}$  to obtain  $\hat{\mathbf{x}}_{est}$ .5) Let  $\mathbf{x}_{est} = \mathbf{F}_d^{-1} \hat{\mathbf{x}}_{est}$ .

When the mask is known with  $\text{supp}(\hat{\mathbf{m}}) = [\delta]_0$ ,  $\delta \ll d$ , maximum error guarantees (Theorem 2.1.1., [21]) are given depending on  $\mathbf{x}$ ,  $d$ ,  $\kappa$ ,  $L$ ,  $\|N_{d,L}\|_F$  (the matrix formed by the noise) and the mask dependent constant  $\mu > 0$ ,

$$\mu := \min_{|p| \leq \kappa-1, q \in [d]_0} \left| (F_d(\hat{\mathbf{m}} \circ S_p \tilde{\mathbf{m}}))_q \right|. \quad (16)$$

In the next section, we look at the situation in which both the specimen and mask are unknown. Since we have already shown that we can rewrite the Fourier squared magnitude measurements as convolutions between the shifted autocorrelations, then the next obvious step is when both the specimen and mask are unknown. This is the topic of **blind deconvolution**, which seeks to recover vectors from their deconvolution. In particular, we will look at a couple of approaches which involve making assumptions based on real world applications.

## 3 Blind Deconvolution

### 3.1 Introduction

Blind Deconvolution is a problem that has been mathematically considered for decades, from more past work ([22], [23], [24], [25][26], [27]) to more recent work ([28], [29], [30], [31] [32]), and summarized in [33]. The goal is to recover a sharp image from an initial blurry image. The first application to compressive sensing was considered in [30].

We consider one-dimensional, discrete, noisy measurements of the form  $\mathbf{y} = \mathbf{f} * \mathbf{g} + \mathbf{n}$ , where the  $\mathbf{f}$  is considered to be an object, signal, or image of consideration.

$\mathbf{g}$  is considered to be a blurring, masking, or point-spread function.  $\mathbf{n}$  is considered to be the noise vector.  $*$  refers to circulant convolution<sup>1</sup>. We consider situations when both  $\mathbf{f}$  and  $\mathbf{g}$  are unknown. The process of recovering the object and blurring function can be generalized to two-dimensional measurements.

The problem of estimating the unknown blurring function and unknown object simultaneously is known as **blind image restoration** ([34], [35], [36], [37]). Although strictly speaking, **blind deconvolution** refers to the noiseless model of recovering  $\mathbf{f}$  and  $\mathbf{g}$  from  $\mathbf{y} = \mathbf{f} * \mathbf{g}$ , the noisy model is commonly referred to as blind deconvolution itself, and this notation will be continued in this chapter.

As we will show later, the problem is ill-posed and ambiguities lead to no unique solution to the pair being viable from any approach.

In Section 3.2, we consider the underlying measurements and assumptions that we will consider. We then show how through manipulation, that we can re-write the original problem as a minimization of a non-convex linear function. In Section 3.3, we demonstrate an iterative approach to the minimization problem, in particular, by applying Wirtinger Gradient Descent. In Section 3.5, we outline the initial estimate used for this gradient descent and fully layout the algorithm that will apply in our numerical simulations. In Section 3.6, we will look at the recovery guarantees that currently exist for this approach. Finally, in Section 3.7, we consider the key conditions used to generate the main recovery theorem, and where further work could be done to generalize these conditions and ultimately, allow more guarantees of recovery.

## 3.2 Blind Deconvolution Model

We now want to approach the blind ptychography problem in which both the mask and specimen are unknown. Using the lemmas in the previous section, we can see that this would reduce to solving a blind deconvolution problem.

**Definition 3.1.** *We consider the blind deconvolution model*

$$\mathbf{y}' = \mathbf{f} * \mathbf{g} + \mathbf{n}, \quad \mathbf{y}, \mathbf{f}, \mathbf{g}, \mathbf{n} \in \mathbb{C}^d,$$

where  $\mathbf{y}$  are blind deconvolutional measurements,  $\mathbf{f}$  is the unknown blurring function (which serves a similar role as to our phase retrieval masks),  $\mathbf{n}$  is the noise, and  $\mathbf{g}$  is the signal (which serves a similar role as to our phase retrieval object). Here  $*$  denotes circular convolution.

We will base our work on the algorithm suggested in [32], considering the assumptions used. In [32], the authors impose general conditions on  $\mathbf{f}$  and  $\mathbf{g}$  that are not restricted to any particular application but allows for flexibility. They also assume that  $\mathbf{f}$  and  $\mathbf{g}$  belong to known linear subspaces.

For the blurring function, it is assumed that  $\mathbf{f}$  is either compactly supported, or

---

<sup>1</sup>\* should refer to ordinary convolution, but for  $\mathbf{g}$  that will be considered in this chapter, circulant convolution will be sufficient.



that  $\mathbf{f}$  decays sufficiently fast so that it can be well approximated by a compactly supported function. Therefore, we make the assumption that  $\mathbf{f} \in \mathbb{C}^d$  satisfies

$$\mathbf{f} := \begin{bmatrix} \mathbf{h} \\ \mathbf{0}_{d-K} \end{bmatrix},$$

for some  $k \ll d$ ,  $\mathbf{h} \in \mathbb{C}^K$ . This again reinforces the notion that the blurring function is analogous to our masking function since both are compactly supported.



**Fig. 2** [31] An example of an image deblurring by solving the deconvolution

For the signal, it is assumed that  $\mathbf{g}$  belongs to a linear subspace spanned by the columns of a known matrix  $\mathbf{C}$ , i.e.,  $\mathbf{g} = \mathbf{C}\tilde{\mathbf{x}}$  for some matrix  $\mathbf{C} \in \mathbb{C}^{d \times N}$ ,  $N \ll d$ . This will lead to an additional restriction we have to place on our blind ptychography but one for which there are real world applications for which this assumption makes reasonable sense.

In [32], the authors use that  $\mathbf{C}$  is a Gaussian random matrix for theoretical guarantees although they demonstrated in numerical simulations that this assumption is not necessary to gain results. In particular, they found good results for when  $\mathbf{C}$  represents a wavelet subspace (suitable for images) or when  $\mathbf{C}$  is a Hadamard-type matrix (suitable for communications).

We assume the noise is complex Gaussian, i.e.  $\mathbf{n} \sim \mathcal{N}(0, \frac{\sigma^2 L_0^2}{2} I_d) + i\mathcal{N}(0, \frac{\sigma^2 L_0^2}{2} I_d)$  is a complex Gaussian noise vector, where  $L_0 = \|\mathbf{h}_0\| \cdot \|\mathbf{x}_0\|$ , and  $\mathbf{h}_0, \mathbf{x}_0$  are the true blurring function and signal.  $\sigma^{-2}$  represents the SNR.

The goal is convert the problem into one which can be algorithmically solvable via gradient descent.

**Proposition 3.1.** [32] Let  $\mathbf{F}_d \in \mathbb{C}^{d \times d}$  be DFT matrix. Let  $\mathbf{B} \in \mathbb{C}^{d \times K}$  denote the first  $K$  columns of  $\mathbf{F}_d$ . Then we have that

$$\mathbf{y} = \mathbf{B}\mathbf{h} \circ \overline{\mathbf{A}\mathbf{x}} + \mathbf{e}, \quad (17)$$

where  $\mathbf{y} = \frac{1}{\sqrt{d}}\widehat{\mathbf{y}}'$ ,  $\bar{\mathbf{A}} = \mathbf{F}\mathbf{C} \in \mathbb{C}^{d \times N}$ , and  $\mathbf{e} = \frac{1}{\sqrt{d}}\mathbf{F}_d\mathbf{n}$  represents noise.

*Proof* By the unitary property of  $\mathbf{F}_d$ , we have that  $\mathbf{B}^*\mathbf{B} = \mathbf{I}_K$ . By applying the scaled DFT matrix  $\sqrt{L}\mathbf{F}_d$  to both sides of the convolution, we have that

$$\sqrt{L}\mathbf{F}_d\mathbf{y} = (\sqrt{L}\mathbf{F}_d\mathbf{f}) \circ (\sqrt{L}\mathbf{F}_d\mathbf{g}) + \sqrt{L}\mathbf{F}_d\mathbf{n}.$$

Additionally, we have that

$$\mathbf{F}_d\mathbf{f} = [\mathbf{B} \ \mathbf{M}] \begin{bmatrix} \mathbf{h} \\ \mathbf{0}_{d-K} \end{bmatrix} = \mathbf{B}\mathbf{h},$$

and we let  $\bar{\mathbf{A}} = \mathbf{F}\mathbf{C} \in \mathbb{C}^{L \times N}$ . Since  $\mathbf{C}$  is Gaussian, then  $\bar{\mathbf{A}} = \mathbf{F}\mathbf{C}$  is also Gaussian. In particular,

$$\bar{\mathbf{A}}_{ij} \sim \mathcal{N}(0, \frac{1}{2}) + i\mathcal{N}(0, \frac{1}{2}).$$

Thus by dividing by  $d$ , the problem converts to

$$\frac{1}{\sqrt{d}}\widehat{\mathbf{y}}' = \mathbf{B}\mathbf{h} \circ \overline{\mathbf{A}\mathbf{x}} + \mathbf{e},$$

where  $\mathbf{e} = \frac{1}{\sqrt{d}}\mathbf{F}_d\mathbf{n} \sim \mathcal{N}(0, \frac{\sigma^2 L_0^2}{2L}\mathbf{I}_d) + i\mathcal{N}(0, \frac{\sigma^2 L_0^2}{2L}\mathbf{I}_d)$  serves as complex Gaussian noise.

Hence by letting  $\mathbf{y} = \frac{1}{\sqrt{d}}\widehat{\mathbf{y}}'$ , we arrive at

$$\mathbf{y} = \mathbf{B}\mathbf{h} \circ \overline{\mathbf{A}\mathbf{x}} + \mathbf{e}.$$

□

We have thus transformed the original blind deconvolution model as a hadamard product. This form of the problem is used in the rest of the section, where  $\mathbf{y} \in \mathbb{C}^d$ ,  $\mathbf{B} \in \mathbb{C}^{d \times K}$ ,  $\mathbf{A} \in \mathbb{C}^{d \times N}$  are given. Our goal is to recover  $\mathbf{h}_0$  and  $\mathbf{x}_0$ .

There are inherent ambiguities to the problem however. If  $(\mathbf{h}_0, \mathbf{x}_0)$  is a solution to the blind deconvolution problem, then so is  $(\alpha\mathbf{h}_0, \alpha^{-1}\mathbf{x}_0)$  for any non-zero constant  $\alpha$ . For most real world applications, this is not an issue. Thus for uniformity, it is assumed that  $\|\mathbf{h}_0\| = \|\mathbf{x}_0\| = \sqrt{L_0}$ .

**Definition 3.2.** We define the matrix-valued linear operator  $\mathcal{A} : \mathbb{C}^{K \times N} \rightarrow \mathbb{C}^d$  by

$$\mathcal{A}(\mathbf{Z}) := \{\mathbf{b}_\ell^* \mathbf{Z} \mathbf{a}_\ell\}_{\ell=1}^d,$$

where  $\mathbf{b}_k$  denotes the  $k$ -th column of  $\mathbf{B}^*$ , and  $\mathbf{a}_k$  is the  $k$ -th column of  $\mathbf{A}^*$ . We also define the corresponding adjoint operator  $\mathcal{A}^* : \mathbb{C}^d \rightarrow \mathbb{C}^{K \times N}$ , given by

$$\mathcal{A}^*(\mathbf{z}) := \sum_{k=1}^d \mathbf{z}_k \mathbf{b}_k \mathbf{a}_k^*.$$

We see that this translates to a lifting problem, where

$$\sum_{k=1}^d \mathbf{b}_\ell \mathbf{b}_\ell^* = \mathbf{B}^* \mathbf{B} = \mathbf{I}_K, \quad \|\mathbf{b}_\ell\| = \frac{K}{d}, \quad \mathbb{E}(\mathbf{a}_\ell \mathbf{a}_\ell^*) = \mathbf{I}_N, \quad \forall \ell \in [d].$$

**Lemma 3.1.** *Let  $\mathbf{y}$  be defined as in Proposition 3.1. Then*

$$\mathbf{y} = \mathcal{A}(\mathbf{h}_0 \mathbf{x}_0^*) + \mathbf{e}. \quad (18)$$

This equivalent model to Proposition 3.1 will be the model worked with for the rest of the chapter.

We aim to recover  $(\mathbf{h}_0, \mathbf{x}_0)$  by solving the minimization problem

$$\min_{(\mathbf{h}, \mathbf{x})} F(\mathbf{h}, \mathbf{x}), \quad F(\mathbf{h}, \mathbf{x}) := \|\mathcal{A}(\mathbf{h} \mathbf{x}^*) - \mathbf{y}\|^2 = \|\mathcal{A}(\mathbf{h} \mathbf{x}^* - \mathbf{h}_0 \mathbf{x}_0^*) - \mathbf{e}\|^2.$$

We also define

$$F_0(\mathbf{h}, \mathbf{x}) := \|\mathcal{A}(\mathbf{h} \mathbf{x}^* - \mathbf{h}_0 \mathbf{x}_0^*)\|^2, \quad \delta = \delta(\mathbf{h}, \mathbf{x}) = \frac{\|\mathbf{h} \mathbf{x}^* - \mathbf{h}_0 \mathbf{x}_0^*\|_F}{d_0}.$$

$F(\mathbf{h}, \mathbf{x})$  is highly non-convex and thus attempts at minimization such as alternating minimization and gradient descent, can get easily trapped in some local minima.

### 3.2.1 Main Theorems

**Theorem 3.1. (Existence Of Unique Solution)** ([30], Theorem 1) *Fix  $\alpha \geq 1$ . Then there exists a constant  $C_\alpha = O(\alpha)$ , such that if*

$$\max(K \cdot \mu_{\max}^2, N \cdot \mu_h^2) \leq \frac{d}{C_\alpha \log^3 d},$$

*then  $\mathbf{X}_0 = \mathbf{h}_0 \mathbf{x}_0^*$  is the unique solution to our minimization problem with probability  $1 - O(d^{-\alpha+1})$ , thus we can separate  $\mathbf{y} = \mathbf{f} * \mathbf{g}$  up to a scalar multiple. When coherence is low, this is tight within a logarithmic factor, as we always have  $\max(K, N) \leq d$ .*

**Theorem 3.2. (Stability From Noise)** ([30], Theorem 2) *Let  $\mathbf{X}_0 = \mathbf{h}_0 \mathbf{m}_0^*$  and suppose the condition of previous theorem holds. We observe that*

$$\mathbf{y} = \mathcal{A}(\mathbf{X}_0) + \mathbf{e},$$

where  $\mathbf{e} \in \mathbb{R}^d$  is an unknown noise vector with  $\|\mathbf{z}\|_2 \leq \delta$ , and estimate  $\mathbf{X}_0$  by solving

$$\min \quad \|\mathbf{X}\|_*, \quad \text{subject to} \quad \|\hat{\mathbf{y}} - \mathcal{A}(\mathbf{X})\|_2 \leq \delta.$$

Let  $\lambda_{\min}, \lambda_{\max}$  be the smallest/largest non-zero eigenvalue of  $\mathcal{A}\mathcal{A}^*$ . Then with probability  $1 - d^{-\alpha+1}$ , the solution  $\mathbf{X}$  will obey

$$\|\mathbf{X} - \mathbf{X}_0\|_F \leq C \frac{\lambda_{\max}}{\lambda_{\min}} \sqrt{\min(K, N)} \delta,$$

for a fixed constant  $C$ .

### 3.3 Wirtinger Gradient Descent

In [32], the approach is to solve the minimization problem (Equation 19) using Wirtinger gradient descent. In this subsection, the algorithm is introduced as well as the main theorems which establish convergence of the proposed algorithm to the true solution.

The algorithm consists of two parts: first an initial guess, and secondly, a variation of gradient descent, starting at the initial guess to converge to the true solution. Theoretical results are established for avoiding getting stuck in local minima.

This is ensured by determining that the iterates are inside some properly chosen basin of attraction of the true solution.

### 3.4 Basin of Attraction

**Proposition 3.2. Basin of Attraction:** (Section 3.1, [32]) *Three neighbourhoods are introduced whose intersection will form the basis of attraction of the solution:*

(i) **Non-uniqueness:** *Due to the scale ambiguity, for numerical stability we introduce the following neighbourhood*

$$N_{L_0} := \{(\mathbf{h}, \mathbf{x}) \mid \|\mathbf{h}\| \leq 2\sqrt{L_0}, \|\mathbf{x}\| \leq 2\sqrt{L_0}\}, \quad L_0 = \|\mathbf{h}_0\| \cdot \|\mathbf{x}_0\|.$$

(ii) **Incoherence:** *The number of measurements required for solving the blind deconvolution problem depends on how much  $\mathbf{h}_0$  is correlated with the rows of the matrix  $\mathbf{B}$ , with the hopes of minimizing the correlation. We define the incoherence between the rows of  $\mathbf{B}$  and  $\mathbf{h}_0$ , via*

$$\mu_{\mathbf{h}}^2 = \frac{d \|\mathbf{B}\mathbf{h}_0\|_{\infty}^2}{\|\mathbf{h}_0\|^2}.$$

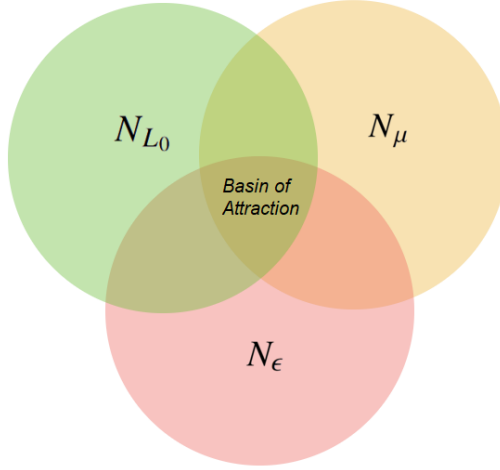
To ensure that the incoherence of the solution is under control, we introduce the neighborhood

$$N_\mu := \{\mathbf{h} \mid \sqrt{d}\|\mathbf{B}\mathbf{h}\|_\infty \leq 4\sqrt{L_0\mu}\}, \quad \mu_{\mathbf{h}} \leq \mu. \quad (19)$$

(iii) **Initial guess:** A carefully chosen initial guess is required due to the non-convexity of the function we wish to minimize. The distance to the true solution is defined via the following neighborhood

$$N_\epsilon := \{(\mathbf{h}, \mathbf{x}) \mid \|\mathbf{h}\mathbf{x}^* - \mathbf{h}_0\mathbf{x}_0^*\|_F \leq \epsilon L_0\}, \quad 0 < \epsilon \leq \frac{1}{15}. \quad (20)$$

Thus the basin of attraction is chosen as  $N_{d_0} \cap N_\mu \cap N_\epsilon$ , where the true solution lies.



**Fig. 3** Basin of Attraction:  $N_{L_0} \cap N_\mu \cap N_\epsilon$

Our approach consists of two parts: We first construct an initial guess that is inside the basin of attraction  $N_{L_0} \cap N_\mu \cap N_\epsilon$ . We then apply a regularized Wirtinger gradient descent algorithm that will ensure that all the iterates remain inside  $N_{L_0} \cap N_\mu \cap N_\epsilon$ . To achieve that, we add a regularizing function  $G(\mathbf{h}, \mathbf{x})$  to the objective function  $F(\mathbf{h}, \mathbf{x})$  to enforce that the iterates remain inside  $N_{L_0} \cap N_\mu$ . Hence we aim to minimize the following regularized objective function, in order to solve the blind deconvolution problem:

$$\tilde{F}(\mathbf{h}, \mathbf{x}) := F(\mathbf{h}, \mathbf{x}) + G(\mathbf{h}, \mathbf{x}),$$

where  $F(\mathbf{h}, \mathbf{x}) := \|\mathcal{A}(\mathbf{h}\mathbf{x}^* - \mathbf{h}_0\mathbf{x}_0^*) - \mathbf{e}\|^2$  is defined as before and  $G(\mathbf{h}, \mathbf{x})$  is the penalty function, of the form

$$G(\mathbf{h}, \mathbf{x}) := \rho \left[ G_0 \left( \frac{\|\mathbf{h}\|^2}{2L} \right) + G_0 \left( \frac{\|\mathbf{x}\|^2}{2L} \right) + \sum_{\ell=1}^d G_0 \left( \frac{d|\mathbf{b}_\ell^* \mathbf{h}|^2}{8L\mu^2} \right) \right],$$

where

$$G_0(z) := \max\{z - 1, 0\}^2, \quad \rho \geq L^2 + 2\|\mathbf{e}\|^2.$$

It is assumed  $\frac{9}{10}L_0 \leq L \leq \frac{11}{10}L_0$  and  $\mu \geq \mu_h$ .

**Remark 3.1.** The matrix  $\mathcal{A}^*(\mathbf{e}) = \sum_{k=1}^d \mathbf{e}_k \mathbf{b}_k \mathbf{a}_k^*$  as a sum of  $d$  rank-1 random matrices, has nice concentration of measure properties. Asymptotically,  $\|\mathcal{A}^*(\mathbf{e})\|$  converges to 0 with rate  $O(d^{1/2})$ . Note that

$$F(\mathbf{h}, \mathbf{x}) = \|\mathbf{e}\|^2 + \|\mathcal{A}(\mathbf{h}\mathbf{x}^* - \mathbf{h}_0\mathbf{x}_0^*)\|_F^2 - 2\text{Re}(\langle \mathcal{A}^*(\mathbf{e}), \mathbf{h}\mathbf{x}^* - \mathbf{h}_0\mathbf{x}_0^* \rangle).$$

If one lets  $d \rightarrow \infty$ , then  $\|\mathbf{e}\|^2 \sim \frac{\sigma^2 d_0^2}{2d} \chi_{2d}^2$  will converge almost surely to  $\sigma^2 d_0^2$ <sup>2</sup> and the cross term  $\text{Re}(\langle \mathbf{h}\mathbf{x}^* - \mathbf{h}_0\mathbf{x}_0^*, \mathcal{A}^*(\mathbf{e}) \rangle)$  will converge to 0. In other words, asymptotically

$$\lim_{d \rightarrow \infty} F(\mathbf{h}, \mathbf{x}) = F_0(\mathbf{h}, \mathbf{x}) + \sigma^2 L_0^2,$$

for all fixed  $(\mathbf{h}, \mathbf{x})$ . This implies that if the number of measurements is large, then  $F(\mathbf{h}, \mathbf{x})$  behaves "almost like"  $F_0(\mathbf{h}, \mathbf{x}) = \|\mathcal{A}(\mathbf{h}\mathbf{x}^* - \mathbf{h}_0\mathbf{x}_0^*)\|^2$ , the noiseless version of  $F(\mathbf{h}, \mathbf{x})$ . So for large  $d$ , we effectively ignore the noise.

**Theorem 3.3.** For any given  $Z \in \mathbb{C}^{K \times N}$ , we have that

$$\mathbb{E}(\mathcal{A}^*(\mathcal{A}(Z))) = Z.$$

*Proof* By linearity, sufficient to prove for  $Z \in \mathbb{C}^{K \times N}$  where  $Z_{i,j} = 1, 0$  otherwise. Then we have that

$$\mathbb{E}((\mathcal{A}^* \mathcal{A}(Z))) = \mathbb{E} \left( \sum_{k=1}^L b_{ki}^* a_{kj} \mathbf{b}_k \mathbf{a}_k^* \right) = \sum_{k=1}^d (b_{ki}^* \mathbf{b}_k \mathbb{E}(a_{kj} \mathbf{a}_k^*)) = \sum_{k=1}^d (b_{ki}^* \mathbf{b}_k) z_j^* = \mathbf{e}_i \mathbf{e}_j^* = Z,$$

Thus we have that

$$\mathbb{E}(\mathcal{A}^*(\mathbf{y})) = \mathbb{E}(\mathcal{A}^*(\mathcal{A}(\mathbf{h}_0\mathbf{x}_0^*) + \mathbf{e})) = \mathbb{E}(\mathcal{A}^*(\mathcal{A}(\mathbf{h}_0\mathbf{x}_0^*))) + \mathbb{E}(\mathcal{A}^*(\mathbf{e})) = \mathbf{h}_0\mathbf{x}_0^*,$$

---

<sup>2</sup>This is due to the law of large numbers

since  $\mathbb{E}(\mathcal{A}^*(\mathbf{e}))$  by the definition of  $\mathbf{e}$ . □

Hence it makes logical sense that the leading singular value and vectors of  $\mathcal{A}^*(\mathbf{y})$  would be a good approximation of  $L_0$  and  $(\mathbf{h}_0, \mathbf{x}_0)$  respectively.

### 3.5 Algorithms

We can now state the algorithm for generating an initial estimate.

---

#### Algorithm 2 Blind Deconvolution Initial Estimate

---

**Input:** Blind Deconvolutional measurements  $\mathbf{y}$ ,  $K = \text{supp}(f)$ .

**Output:** Estimate underlying signal and blurring function.

- 1) Compute  $\mathcal{A}^*(\mathbf{y})$  and find the leading singular value, left and right singular vectors of  $\mathcal{A}^*(\mathbf{y})$ , denoted by  $d$ ,  $\tilde{\mathbf{h}}_0$ , and  $\tilde{\mathbf{x}}_0$  respectively.
- 2) Solve the following optimization problem

$$\mathbf{u}_0 := \underset{\mathbf{z}}{\operatorname{argmin}} \|\mathbf{z} - \sqrt{d}\tilde{\mathbf{h}}_0\|^2, \quad \text{subject to } \sqrt{d}\|\mathbf{B}\mathbf{z}\|_\infty \leq 2\sqrt{L}\mu,$$

and  $\mathbf{x}_0 = \sqrt{L}\tilde{\mathbf{x}}_0$ .

---

Since we are dealing with complex variables, for the gradient descent, Wirtinger derivatives are utilized. Since  $\tilde{F}$  is a real-valued function, we only need to consider the derivative of  $\tilde{F}$ , with respect to  $\tilde{\mathbf{h}}$  and  $\tilde{\mathbf{x}}$ , and the corresponding updates of  $\mathbf{h}$  and  $\mathbf{x}$  since

$$\frac{\partial \tilde{F}}{\partial \tilde{\mathbf{h}}} = \overline{\frac{\partial \tilde{F}}{\partial \mathbf{h}}}, \quad \frac{\partial \tilde{F}}{\partial \tilde{\mathbf{x}}} = \overline{\frac{\partial \tilde{F}}{\partial \mathbf{x}}}.$$

In particular, we denote

$$\nabla \tilde{F}_{\tilde{\mathbf{h}}} := \frac{\partial \tilde{F}}{\partial \tilde{\mathbf{h}}}, \quad \nabla \tilde{F}_{\tilde{\mathbf{x}}} := \frac{\partial \tilde{F}}{\partial \tilde{\mathbf{x}}}. \quad (21)$$

We can now state the full algorithm.

**Algorithm 3** Wirtinger Gradient Descent Blind Deconvolution Algorithm**Input:** Blind Deconvolutional measurements  $\mathbf{y}$ ,  $K = \text{supp}(f)$ .**Output:** Estimate underlying signal and blurring function.

- 1) Compute  $\mathcal{A}^*(\mathbf{y})$  and find the leading singular value, left and right singular vectors of  $\mathcal{A}^*(\mathbf{y})$ , denoted by  $d$ ,  $\tilde{\mathbf{h}}_0$ , and  $\tilde{\mathbf{x}}_0$  respectively.
- 2) Solve the following optimization problem

$$\mathbf{u}_0 := \underset{\mathbf{z}}{\operatorname{argmin}} \|\mathbf{z} - \sqrt{L}\tilde{\mathbf{h}}_0\|^2, \quad \text{subject to } \sqrt{d}\|B\mathbf{z}\|_\infty \leq 2\sqrt{L}\mu,$$

and  $\mathbf{x}_0 = \sqrt{L}\tilde{\mathbf{x}}_0$ .

- 3) Compute Wirtinger Gradient Descent

**while** halting criterion false **do**

$$\mathbf{u}_t = \mathbf{u}_{t-1} - \eta \nabla \tilde{F}_{\mathbf{h}}(\mathbf{u}_{t-1}, \mathbf{u}_{t-1})$$

$$\mathbf{v}_t = \mathbf{v}_{t-1} - \eta \nabla \tilde{F}_{\mathbf{x}}(\mathbf{v}_{t-1}, \mathbf{v}_{t-1})$$

**end while**

- 4) Set  $(\mathbf{h}, \mathbf{x}) = (\mathbf{u}_t, \mathbf{v}_t)$

In [32], the authors show that with a carefully chosen initial guess  $(\mathbf{u}_0, \mathbf{v}_0)$ , running Wirtinger gradient descent to minimize  $\tilde{\mathbf{F}}(h, x)$  will guarantee linear convergence of the sequence  $(\mathbf{u}_t, \mathbf{v}_t)$  to the global minimum  $(\mathbf{h}_0, \mathbf{x}_0)$  in the noiseless case, and also provide robust recovery in the presence of noise. The results are summarized in the following two theorems.

### 3.6 Main Theorems

**Theorem 3.4. (Main Theorem 1)** ([32], Theorem 3.1) *The initialization obtained via Algorithm 2 satisfies*

$$(\mathbf{u}_0, \mathbf{v}_0) \in \frac{1}{\sqrt{3}}N_{L_0} \cap \frac{1}{\sqrt{3}}N_\mu \cap N_{\frac{2}{5}\epsilon}, \quad \frac{9}{10}L_0 \leq L \leq \frac{11}{10}L_0,$$

with probability at least  $1 - d^{-\gamma}$  if the number of measurements is sufficient large, that is

$$d \geq C_\gamma(\mu_h^2 + \sigma^2) \max\{K, N\} \log^2 d / \epsilon^2,$$

where  $\epsilon$  is a predetermined constant on  $(0, \frac{1}{15}]$ , and  $C_\gamma$  is a constant only linearly depending on  $\gamma$  with  $\gamma \geq 1$ .



The following theorem establishes that as long as the initial guess lies inside the basin of attraction of the true solution, regularized gradient descent will converge to this solution (or to a nearby solution in case of noisy data).

**Theorem 3.5. (Main Theorem 2)** ([32], Theorem 3.2) *Assume that the initialization  $(\mathbf{u}_0, \mathbf{v}_0) \in \frac{1}{\sqrt{3}}N_{L_0} \cap \frac{1}{\sqrt{3}}N_\mu \cap N_{\frac{2}{5}\epsilon}$ , and that  $d \geq C_\gamma(\mu^2 + \sigma^2) \max\{K, N\} \log^2(L)/\epsilon^2$ . Then Algorithm 3 will create a sequence  $(\mathbf{u}_t, \mathbf{v}_t) \in N_{d_0} \cap N_\mu \cap N_\epsilon$  which converges geometrically to  $(\mathbf{h}_0, \mathbf{x}_0)$  in the sense that with probability at least  $1 - 4d^{-\gamma} - \frac{1}{\gamma}e^{-(K+N)}$ , we have that*

$$\max\{\sin\angle(\mathbf{u}_t, \mathbf{h}_0), \sin\angle(\mathbf{v}_t, \mathbf{x}_0)\} \leq \frac{1}{L_t} \left( \frac{2}{3}(1 - \eta\omega)^{t/2} \eta L_0 + 50\|\mathcal{A}^*(\mathbf{e})\| \right),$$

and

$$|L_t - L_0| \leq \frac{2}{3}(1 - \eta\omega)^{t/2} \epsilon L_0 + 50\|\mathcal{A}^*(\mathbf{e})\|,$$

where  $L_t := \|\mathbf{u}_t\| \cdot \|\mathbf{v}_t\|$ ,  $\omega > 0$ ,  $\eta$  is the fixed stepsize. Here

$$\|\mathcal{A}^*(\mathbf{e})\| \leq C_0 \sigma d_0 \max \left\{ \sqrt{\frac{(\gamma + 1) \max\{K, N\} \log d}{d}}, \frac{(\gamma + d) \sqrt{KN} \log^2 d}{d} \right\},$$

holds with probability  $1 - d^{-\gamma}$ .

It has been shown with high probability that as long as the initial guess lies inside the basin of attraction of the true solution, Wirtinger gradient descent will converge towards the solution.

### 3.7 Key Conditions

**Theorem 3.6. Four Key Conditions:**

(i) **(Local RIP Condition)** ([32], Condition 5.1) *The following local Restricted Isometry Property (RIP) for  $A$  holds uniformly for all  $(\mathbf{h}, \mathbf{x})$  in the basin of attraction  $(N_{L_0} \cap N_\mu \cap N_\epsilon)$*

$$\frac{3}{4} \|\mathbf{h}\mathbf{x}^* - \mathbf{h}_0\mathbf{x}_0^*\|_F^2 \leq \|\mathcal{A}(\mathbf{h}\mathbf{x}^* - \mathbf{h}_0\mathbf{x}_0^*)\|^2 \leq \frac{5}{4} \|\mathbf{h}\mathbf{x}^* - \mathbf{h}_0\mathbf{x}_0^*\|_F^2.$$

(ii) **(Robustness Condition)** ([32], Condition 5.2) *For the complex Gaussian noise  $\mathbf{e}$ , with high probability*

$$\|\mathcal{A}^*(\mathbf{e})\| \leq \frac{\epsilon L_0}{10\sqrt{2}},$$

*for  $d$  sufficiently large, that is,  $d \geq C_\gamma(\frac{\sigma^2}{\epsilon^2} + \frac{\sigma}{\epsilon}) \max\{K, N\} \log d$ ;*

(iii) **(Local Regularity Condition)** ([32], Condition 5.3) *There exists a regularity constant  $\omega = \frac{d_0}{5000} > 0$  such that*

$$\|\nabla \tilde{F}(\mathbf{h}, \mathbf{x})\|^2 \geq \omega[\tilde{F}(\mathbf{h}, \mathbf{x}) - c]_+, \quad c = \|\mathbf{e}\|^2 + 1700\|\mathcal{A}^*(\mathbf{e})\|^2,$$

*for all  $(\mathbf{h}, \mathbf{x}) \in N_{L_0} \cap N_\mu \cap N_\epsilon$ ;*

(iv) **(Local Smoothness Condition)** ([32], Condition 5.4) *Denote  $\mathbf{z} := (\mathbf{h}, \mathbf{x})$ . There exists a constant  $C_d$  such that*

$$\|\nabla f(\mathbf{z} + t\Delta\mathbf{z}) - \nabla f(\mathbf{z})\| \leq C_d t \|\Delta\mathbf{z}\|, \quad 0 \leq t \leq 1,$$

*for all  $\{(\mathbf{z}, \Delta\mathbf{z}) \mid \mathbf{z} + t\Delta\mathbf{z} \in N_\epsilon \cap N_{\tilde{F}}, \forall 0 \leq t \leq 1\}$ , i.e., the whole segment connecting  $(\mathbf{h}, \mathbf{x})$  and  $\nabla(\mathbf{h}, \mathbf{x})$  belongs to the non-convex set  $N_\epsilon \cap N_{\tilde{F}}$ .*

## 4 Blind Ptychography

### 4.1 Introduction

A more recent area of study is blind ptychography, in which both the object and the mask are considered unknown, up to reasonable assumptions. The first successful recovery was given in ([38],[39]), further study into the sufficient overlap ([40], [41], [42]), and summarized in [43].

Let  $\mathbf{x}, \mathbf{m} \in \mathbb{C}^d$  denote the unknown sample and mask, respectively. We suppose

that we have  $d^2$  noisy ptychographic measurements of the form

$$(\mathbf{Y})_{\ell,k} = |(\mathbf{F}(\mathbf{x} \circ S_k \mathbf{m}))_{\ell}|^2 + (\mathbf{N})_{\ell,k}, \quad (\ell, k) \in [d]_0 \times [d]_0, \quad (22)$$

where  $S_k, \circ, \mathbf{F}$  denote  $k^{\text{th}}$  circular shift, Hadamard product, and  $d$ -dimensional discrete Fourier transform, and  $\mathbf{N}$  is the matrix of additive noise. By Theorem 2.2, we have shown we can rewrite the measurements as

$$\left( \mathbf{Y}^T \mathbf{F}^T \right)_k = d \cdot (\mathbf{x} \circ S_k \bar{\mathbf{x}}) * (\tilde{\mathbf{m}} \circ S_{-k} \tilde{\tilde{\mathbf{m}}}) + \left( \mathbf{N}^T \mathbf{F}^T \right)_k, \quad (23)$$

where  $*$  denotes the  $d$ -dimensional discrete convolution, and  $\tilde{\mathbf{m}}$  denotes the reversal of  $\mathbf{m}$  about its first entry. This is now a scaled blind deconvolution problem which has been studied in [30],[32].

## 4.2 Main Results

### 4.2.1 Recovering the Sample

To recover the sample, we will need to assume that  $\mathbf{x}$  belongs to a known subspace. Initially we solve algorithmically for the zero shift case ( $k = 0$ ) and then generalize the method to solve for the estimate which utilizes all the obtained shifts.

Our assumptions are as follows:  $\mathbf{x} \in \mathbb{C}^d$  unknown,  $\mathbf{x} = C\mathbf{x}', C \in \mathbb{C}^{d \times N}, N \ll L$  known,  $\mathbf{x}' \in \mathbb{C}^N$  or  $\mathbb{R}^N$  unknown  $\mathbf{m} \in \mathbb{C}^d$  unknown,  $\text{supp}(\mathbf{m}) \subseteq [\delta]_0, K$  known,  $\|\mathbf{m}\|_2$  known. Known noisy measurements  $\mathbf{Y}$ .

Our first goal is to compute an estimate  $\mathbf{x}_{est}$  of  $\mathbf{x}$ , true up to a global phase. We will use this estimate to then produce an estimate  $\mathbf{m}_{est}$  of  $\mathbf{m}$ , again true up to a global phase.

Firstly, we let  $\mathbf{y}$  be the first column of  $\frac{1}{\sqrt{d}} \cdot \overline{\mathbf{F}((\mathbf{F}\mathbf{Y})^T)}$ ,  $\mathbf{f} = \tilde{\mathbf{m}} \circ \overline{\tilde{\tilde{\mathbf{m}}}}$  (so  $\|\mathbf{f}\|_2$  known). We next set  $\mathbf{g} = \mathbf{x} \circ \bar{\mathbf{x}}$  but to fully utilize the blind deconvolution algorithm, we will need a lemma concerning hadamard products of products of matrices. Firstly, we need define some products between matrices.

**Definition 4.1.** Let  $\mathbf{A} = (A_{i,j}) \in \mathbb{C}^{m \times n}$  and  $\mathbf{B} = (B_{i,j}) \in \mathbb{C}^{p \times q}$ . Then the **Kronecker product**  $\mathbf{A} \otimes \mathbf{B} \in \mathbb{C}^{mp \times nq}$  is defined by

$$(\mathbf{A} \otimes \mathbf{B})_{n(i-1)+k, q(j-1)+\ell} = A_{i,j} B_{k,\ell}.$$

**Definition 4.2.** Let  $\mathbf{A} \in \mathbb{C}^{m \times n}$  and  $\mathbf{B} \in \mathbb{C}^{p \times n}$  with columns  $\mathbf{a}_i, \mathbf{b}_i$  for  $i \in [n]_0$ . Then the **Khatri-Rao product**  $\mathbf{A} \bullet \mathbf{B} \in \mathbb{C}^{mp \times n}$  is defined by

$$\mathbf{A} \bullet \mathbf{B} = [\mathbf{a}_0 \otimes \mathbf{b}_0 \ \mathbf{a}_1 \otimes \mathbf{b}_1 \ \dots \ \mathbf{a}_{n-1} \otimes \mathbf{b}_{n-1}]. \quad (24)$$

**Definition 4.3.** Let  $\mathbf{A} \in \mathbb{C}^{m \times n}$  and  $\mathbf{B} \in \mathbb{C}^{m \times p}$  be matrices with rows  $\mathbf{a}_i, \mathbf{b}_i$  for  $i \in [m]_0$ . Then the **transposed Khatri-Rao product** (or **face-splitting product**), denoted  $\odot$ , is the matrix whose rows are Kronecker products of the columns of  $\mathbf{A}$  and  $\mathbf{B}$  i.e. the rows of  $\mathbf{A} \odot \mathbf{B} \in \mathbb{C}^{m \times np}$  are given by

$$(\mathbf{A} \odot \mathbf{B})_i = \mathbf{a}_i \otimes \mathbf{b}_i, \quad i \in [m]_0.$$

We then utilize the following lemma concerning the transposed Khatri-Rao product.

**Lemma 4.1** (Theorem 1, [44]). Let  $\mathbf{A} \in \mathbb{C}^{m \times n}, \mathbf{B} \in \mathbb{C}^{n \times p}, \mathbf{C} \in \mathbb{C}^{m \times q}, \mathbf{D} \in \mathbb{C}^{q \times p}$ . Then we have that

$$(\mathbf{AB}) \circ (\mathbf{CD}) = (\mathbf{A} \bullet \mathbf{C})(\mathbf{B} \odot \mathbf{D}),$$

where  $\circ$  is the Hadamard product,  $\bullet$  is the standard Khatri-Rao product, and  $\odot$  is the transposed Khatri-Rao product.

Thus by Lemma 4.1 we have that for  $\mathbf{g} = \mathbf{x} \circ \bar{\mathbf{x}} = \mathbf{C}\mathbf{x}' \circ \bar{\mathbf{C}}\bar{\mathbf{x}}'$ . Then  $\mathbf{g} = \mathbf{C}'\mathbf{x}''$  where  $\mathbf{C}' \in \mathbb{C}^{L \times N^2}, \mathbf{x}'' \in \mathbb{C}^{N^2}$  are given by

$$\mathbf{C}' = \mathbf{C} \bullet \bar{\mathbf{C}}, \quad \mathbf{x}'' = \mathbf{x}' \odot \bar{\mathbf{x}}'.$$

We now compute RRR Blind Deconvolution (Algorithm 3) with  $\mathbf{y}, \mathbf{f}, \mathbf{g}, \mathbf{C}, K = \delta$  as above ( $\mathbf{B}$  last  $K$  columns of DFT matrix) to obtain estimate for  $\mathbf{x}' \odot \bar{\mathbf{x}}'$ . Use angular synchronisation to solve for  $\mathbf{x}'$ , and thus solve for  $\mathbf{x}$ .

---

**Algorithm 4** Blind Ptychography (Zero Shift)

---

**Input:**

- 1)  $\mathbf{x} \in \mathbb{C}^d$  unknown,  $\mathbf{x} = \mathbf{C}\mathbf{x}', \mathbf{C} \in \mathbb{C}^{d \times N}, N \ll L$  known,  $\mathbf{x}' \in \mathbb{C}^N$  or  $\mathbb{R}^N$  unknown.
- 2)  $\mathbf{m} \in \mathbb{C}^d$  unknown,  $\text{supp}(\mathbf{m}) \subseteq [\delta]_0, K$  known,  $\|\mathbf{m}\|_2$  known Known noisy measurements  $\mathbf{Y}$ .
- 3) Known noisy measurements  $\mathbf{Y}$ .

**Output:** Estimate  $\mathbf{x}_{est}$  of  $\mathbf{x}$  true up to a global phase.

- 1) Let  $\mathbf{y}$  be the first column of  $\frac{1}{\sqrt{d}} \cdot \overline{\mathbf{F}(\mathbf{F}\mathbf{Y})^T}$ ,  $\mathbf{f} = \bar{\mathbf{m}} \circ \bar{\mathbf{m}}$  (so  $\|\mathbf{f}\|_2$  known)
- 2) Let  $\mathbf{g} = \mathbf{x} \circ \bar{\mathbf{x}} = \mathbf{C}\mathbf{x}' \circ \bar{\mathbf{C}}\bar{\mathbf{x}}'$ . Then  $\mathbf{g} = \mathbf{C}'\mathbf{x}''$  where  $\mathbf{C}' \in \mathbb{C}^{d \times N^2}, \mathbf{x}'' \in \mathbb{C}^{N^2}$  are given by

$$\mathbf{C}' = \mathbf{C} \bullet \bar{\mathbf{C}}, \quad \mathbf{x}'' = \mathbf{x}' \odot \bar{\mathbf{x}}'.$$

- 3) Compute RRR Blind Deconvolution (Algorithm 1 & 2, [32]) with  $\mathbf{y}, \mathbf{f}, \mathbf{g}, \mathbf{C}, K = \delta$  as above ( $\mathbf{B}$  last  $K$  columns of DFT matrix) to obtain estimate for  $\mathbf{x}' \odot \bar{\mathbf{x}}'$ .
  - 4) Use angular synchronisation to solve for  $\mathbf{x}'$ , and thus compute  $\mathbf{x}_{est}$ .
-

## 4.2.2 Recovering the Mask

Once the estimate of  $\mathbf{x}$  has been found, denoted  $\mathbf{x}_{est}$ , we use this estimate to find  $\mathbf{m}_{est}$ . We first compute  $\mathbf{g}_{est} = \mathbf{x}_{est} \circ \bar{\mathbf{x}}_{est}$ , and then we use point-wise division to find

$$\mathbf{F}(\tilde{\mathbf{m}} \circ \bar{\tilde{\mathbf{m}}}) = \frac{\overline{\mathbf{F}^{-1}((\mathbf{F}\mathbf{Y})^T)}}{\mathbf{F}(\mathbf{x}_{est} \circ \bar{\mathbf{x}}_{est})}. \quad (25)$$

Then use an inverse Fourier transform, a reversal and then angular synchronization, similar to obtaining  $\mathbf{x}_{est}$ .

---

### Algorithm 5 Recovering The Mask

---

**Input:** 1)  $\mathbf{x}_{est}$  generated by Algorithm 4.

2) Known noisy measurements  $\mathbf{Y}$ .

3)  $\text{supp}(\mathbf{m}) \subseteq [\delta]_0, K$  known,  $\|\mathbf{m}\|_2$  known.

**Output:** Estimate  $\mathbf{m}_{est}$  of  $\mathbf{m}$  true up to a global phase.

1) Compute  $\mathbf{g}_{est} = \mathbf{x}_{est} \circ \bar{\mathbf{x}}_{est}$  and  $2\delta - 1$  perform point-wise divisions to obtain

$$\mathbf{F}(\tilde{\mathbf{m}}_{est} \circ S_{-k}\bar{\tilde{\mathbf{m}}}_{est}) = \frac{\overline{\mathbf{F}^{-1}((\mathbf{F}\mathbf{Y})^T)_k}}{\mathbf{F}(\mathbf{x}_{est} \circ S_k\bar{\mathbf{x}}_{est})}. \quad (26)$$

2) Compute inverse Fourier transform to obtain  $\tilde{\mathbf{m}}_{est} \circ S_{-k}\bar{\tilde{\mathbf{m}}}_{est}$  and use these to form the diagonals of a banded matrix.

3) Use angular synchronisation to solve for  $\tilde{\mathbf{m}}_{est}$ , and thus perform a reversal to compute  $\mathbf{m}_{est}$ .

4) Let  $\alpha = \frac{\|\mathbf{m}_{est}\|_2}{\|\mathbf{m}\|_2}$ . Finally, let  $\mathbf{x}_{est} = \alpha\mathbf{x}_{est}$ ,  $\mathbf{m}_{est} = \alpha^{-1}\mathbf{m}_{est}$

---

## 4.2.3 Multiple Shifts

To generalize the setup, we let  $\mathbf{y}_{(k)}$  denote the  $k^{\text{th}}$  column of  $\frac{1}{\sqrt{d}} \cdot \overline{\mathbf{F}((\mathbf{F}\mathbf{Y})^T)}$ ,  $\mathbf{f}_{(k)} = \tilde{\mathbf{m}} \circ S_{-k}\bar{\tilde{\mathbf{m}}}$ . Let  $\mathbf{g}_{(k)} = \mathbf{x} \circ S_k\bar{\mathbf{x}} = \mathbf{C}\mathbf{x}' \circ S_k\bar{\mathbf{C}}\bar{\mathbf{x}}'$ . Then again by another application of Lemma 4.1,  $\mathbf{g}_{(k)} = \mathbf{C}'_{(k)}\mathbf{x}''$  where  $\mathbf{C}' \in \mathbb{C}^{d \times N^2}$ ,  $\mathbf{x}'' \in \mathbb{C}^{N^2}$  are given by

$$\mathbf{C}'_{(k)} = \mathbf{C} \bullet S_k\bar{\mathbf{C}}, \quad 0 \leq k \leq K, d - K + 1 \leq k \leq d, \quad \mathbf{x}'' = \mathbf{x}' \odot \bar{\mathbf{x}}' = \text{vec}(\mathbf{x}'(\mathbf{x}')^*).$$

We then perform  $2\delta - 1$  blind deconvolutions to obtain  $2\delta - 1$  estimates of  $\mathbf{x}$  and  $\mathbf{m}$ , labelled  $\mathbf{x}_{est}^i, \mathbf{m}_{est}^j$  respectively for  $i, j \in [2\delta - 1]_0$ .

Ideally, we would want to select the estimates which generates the minimum error for each  $\mathbf{x}$  and  $\mathbf{m}$  but that implies prior knowledge of  $\mathbf{x}$  and  $\mathbf{m}$ . Instead, compute

$(2\delta - 1)^2$  estimates of the Fourier measurements by

$$(\mathbf{Y}_{est}^{i,j})_{\ell,k} = |(\mathbf{F}(\mathbf{x}_{est}^i \circ S_k \mathbf{m}_{est}^j))_{\ell}|^2, \quad i, j \in [2\delta - 1]_0. \quad (27)$$

We then compute the associated error

$$(i', j') = \underset{(i,j)}{\operatorname{argmin}} \frac{\|\mathbf{Y}_{est}^{i,j} - \mathbf{Y}\|_F^2}{\|\mathbf{Y}\|_F^2}, \quad i, j \in [2\delta - 1]_0. \quad (28)$$

Then let  $\mathbf{x}_{est} = \mathbf{x}_{est}^{i'}$ ,  $\mathbf{m}_{est} = \mathbf{m}_{est}^{j'}$ .

---

**Algorithm 6** Blind Ptychography (Multiple Shifts)

---

**Input:**

- 1)  $\mathbf{x} \in \mathbb{C}^d$  unknown,  $\mathbf{x} = C\mathbf{x}'$ ,  $C \in \mathbb{C}^{d \times N}$ ,  $N \ll d$  known,  $\mathbf{x}' \in \mathbb{C}^N$  or  $\mathbb{R}^N$  unknown.
- 2)  $\mathbf{m} \in \mathbb{C}^d$  unknown,  $\operatorname{supp}(\mathbf{m}) \subseteq [\delta]_0$ ,  $K$  known,  $\|\mathbf{m}\|_2$ .
- 3) Known noisy measurements  $\mathbf{Y}$ .

**Output:** Estimate  $\mathbf{x}_{est}$  of  $\mathbf{x}$ , true up to a global phase

- 1) Let  $\mathbf{y}_{(k)}$  denote the  $k^{\text{th}}$  column of  $\frac{1}{\sqrt{d}} \cdot \mathbf{F}((\mathbf{F}\mathbf{Y})^T)$ ,  $\mathbf{f}_{(k)} = \tilde{\mathbf{m}} \circ S_{-k} \tilde{\mathbf{m}}$  (so  $\|\mathbf{f}_{(k)}\|_2$  known).
- 2) Let  $\mathbf{g}_{(k)} = \mathbf{x} \circ S_k \bar{\mathbf{x}} = C\mathbf{x}' \circ S_k \bar{C}\bar{\mathbf{x}}'$ . Then  $\mathbf{g}_{(k)} = C'\mathbf{x}''$  where  $C' \in \mathbb{C}^{d \times N^2}$ ,  $\mathbf{x}'' \in \mathbb{C}^{N^2}$  are given by

$$C'_{(k)} = C \bullet S_k \bar{C}, \quad 0 \leq k \leq K, d - K + 1 \leq K \leq d, \quad \mathbf{x}'' = \mathbf{x}' \odot \bar{\mathbf{x}}'.$$

- 3) Perform  $2\delta - 1$  RRR Blind Deconvolutions (Algorithm 1 & 2, [32]) with  $\mathbf{y}_{(k)}$ ,  $\mathbf{f}_{(k)}$ ,  $\mathbf{g}_{(k)}$ ,  $C$ , as above to obtain  $2\delta - 1$  estimates for  $\mathbf{x}' \odot \bar{\mathbf{x}}'$ .
- 4) Use angular synchronisation to solve for  $2\delta - 1$  estimates  $\mathbf{x}'_{est}$ , and thus solve for  $2\delta - 1$  estimates  $\mathbf{x}_{est}^i = C\mathbf{x}'_{est}$ ,  $i \in [2\delta - 1]_0$ .
- 5) Use these estimates  $\mathbf{x}_{est}^i$  to compute  $2\delta - 1$  estimates  $\mathbf{m}_{est}^j$ ,  $j \in [2\delta - 1]_0$ .
- 6) Let  $\alpha^i = \frac{\|\mathbf{m}_{est}^i\|_2}{\|\mathbf{m}\|_2}$ , and for  $i \in [2\delta - 1]_0$ , let  $\mathbf{x}_{est}^i = \alpha^i \mathbf{x}'_{est}$ ,  $\mathbf{m}_{est}^i = \frac{1}{\alpha^i} \mathbf{m}_{est}$ .
- 7) Compute  $(2\delta - 1)^2$  estimates of the Fourier measurements by

$$(\mathbf{Y}_{est}^{i,j})_{\ell,k} = |(\mathbf{F}(\mathbf{x}_{est}^i \circ S_k \mathbf{m}_{est}^j))_{\ell}|^2, \quad i, j \in [2\delta - 1]_0. \quad (29)$$

We then compute the associated error  $(i', j') = \underset{(i,j)}{\operatorname{argmin}} \frac{\|\mathbf{Y}_{est}^{i,j} - \mathbf{Y}\|_F^2}{\|\mathbf{Y}\|_F^2}$ ,  $i, j \in [2\delta - 1]_0$ .

- 8) Let  $\mathbf{x}_{est} = \mathbf{x}_{est}^{i'}$ ,  $\mathbf{m}_{est} = \mathbf{m}_{est}^{j'}$ .
-

### 4.3 Numerical Simulations

All simulations were performed using MATLAB R2021b on an Intel desktop with a 2.60GHz i7-10750H CPU and 16GB DDR4 2933MHz memory. All code used to generate the figures below is publicly available at <https://github.com/MarkPhilipRoach/BlindPtychography>.

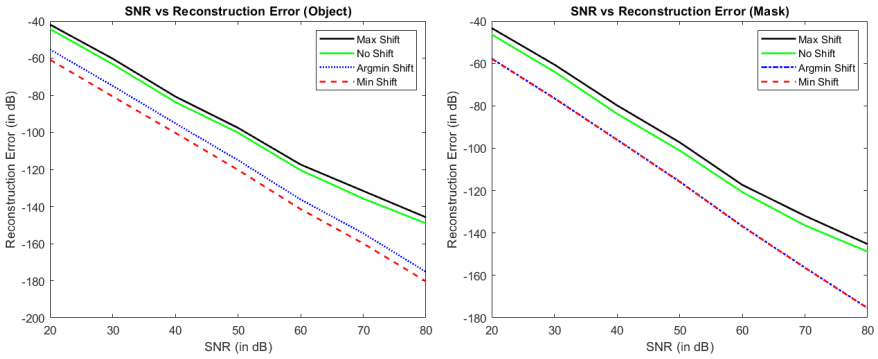
To be more precise, we have defined the immeasurable (in practice since  $\mathbf{x}$  and  $\mathbf{m}$  are both unknown) estimates

$$\begin{aligned} \text{Max Shift}^{(x)} &= \underset{x_{est}^i}{\text{argmax}} \|\mathbf{x} - \mathbf{x}_{est}^i\|_2^2, \\ \text{Max Shift}^{(m)} &= \underset{m_{est}^j}{\text{argmax}} \|\mathbf{m} - \mathbf{m}_{est}^j\|_2^2, \quad i, j \in [2\delta - 1]_0, \\ \text{Min Shift}^{(x)} &= \underset{x_{est}^i}{\text{argmin}} \|\mathbf{x} - \mathbf{x}_{est}^i\|_2^2, \\ \text{Min Shift}^{(m)} &= \underset{m_{est}^j}{\text{argmin}} \|\mathbf{m} - \mathbf{m}_{est}^j\|_2^2, \quad i, j \in [2\delta - 1]_0. \end{aligned}$$

and the measurable estimates. First, No Shift<sup>(x)</sup> and No Shift<sup>(m)</sup> refer to the zero shift estimates outlined in Algorithm 4. Secondly, the estimates achieved in Algorithm 6

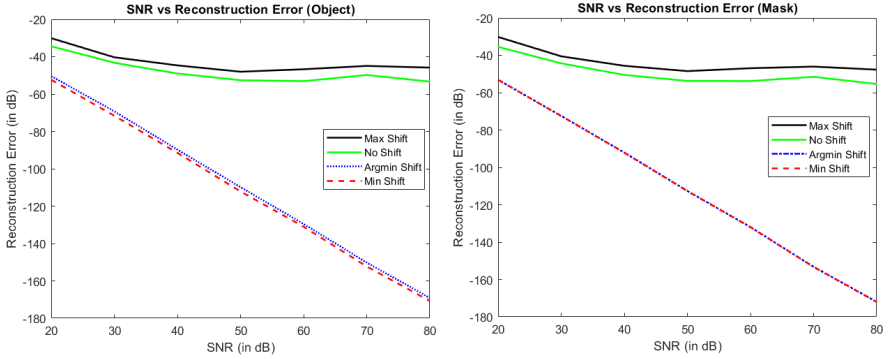
$$(\text{Argmin Shift}^{(x)}, \text{Argmin Shift}^{(m)}) = (\mathbf{x}^{i'}, \mathbf{m}^{j'}),$$

$$(i', j') = \underset{(i, j)}{\text{argmin}} \frac{\|\mathbf{Y}_{est}^{i, j} - \mathbf{Y}\|_F^2}{\|\mathbf{Y}\|_F^2}, \quad i, j \in [2\delta - 1]_0.$$

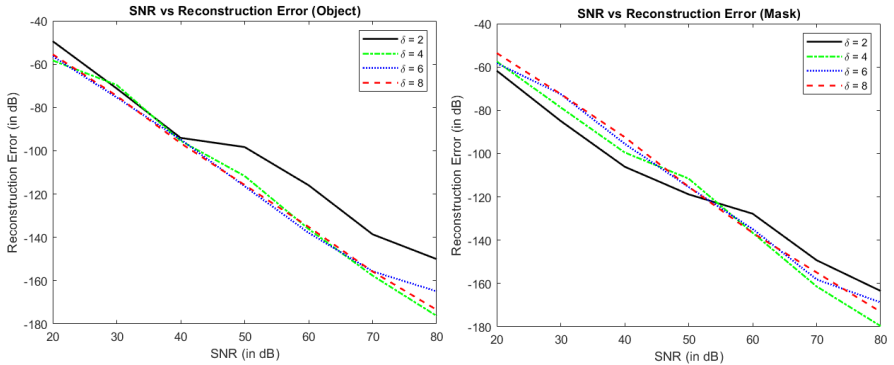


**Fig. 4**  $d = 2^6$ ,  $K = \delta = \log_2 d$ ,  $N = 4$ ,  $\mathbf{C}$  complex Gaussian. Max shift refers to the maximum error achieved from a blind deconvolution of a particular shift. Min shift refers to the maximum error achieved from a blind deconvolution of a particular shift. Argmin Shift refers to the choice of object and mask chosen in Step 6 of Algorithm 6. Averaged over 100 simulations. 1000 iterations.

Figure 4 demonstrates robust recovery under noise. It also demonstrates the impact of performing the  $2\delta - 1$  blind deconvolutions and taking the Argmin Shift, versus simply taking the non-shifted object and mask. It also demonstrates how closely the reconstructions error from Argmin Shift and Min Shift are, in particular for the mask. Figure 5 demonstrates the impact even more, showing that with a higher value for the known subspace, the more accurate the Argmin Shift and Min Shift are, as well as demonstrating the large difference between the Max Shift and Min Shift.

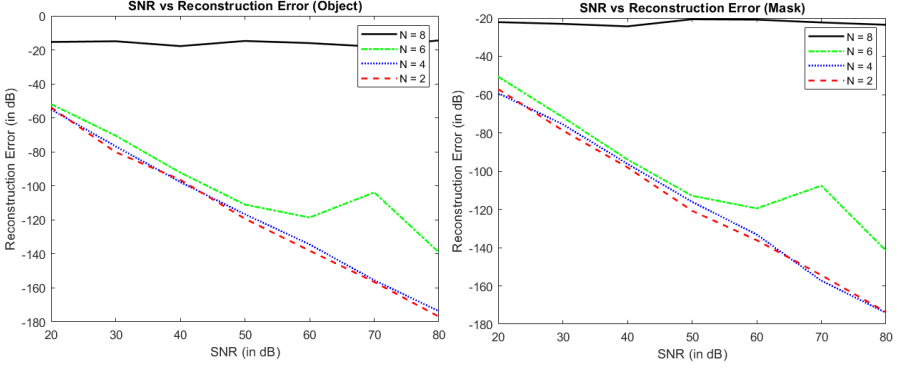


**Fig. 5**  $d = 2^6$ ,  $K = \delta = \log_2 d$ ,  $N = 6$ ,  $\mathbf{C}$  complex Gaussian. Max shift refers to the maximum error achieved from a blind deconvolution of a particular shift. Min shift refers to the maximum error achieved from a blind deconvolution of a particular shift. Argmin Shift refers to the choice of object and mask chosen in Step 6 of Algorithm 6. Averaged over 100 simulations. 1000 iterations.



**Fig. 6**  $d = 2^6$ ,  $N = 4$ ,  $\mathbf{C}$  complex Gaussian. Application of Algorithm 6 with varying  $K = \delta$ .

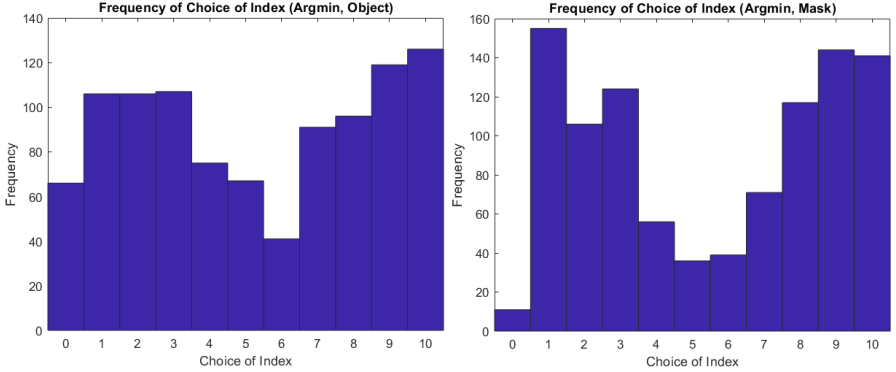




**Fig. 7**  $d = 2^6$ ,  $K = \delta = 6$ ,  $\mathbf{C}$  complex Gaussian. Application of Algorithm 6 with varying  $N$ .

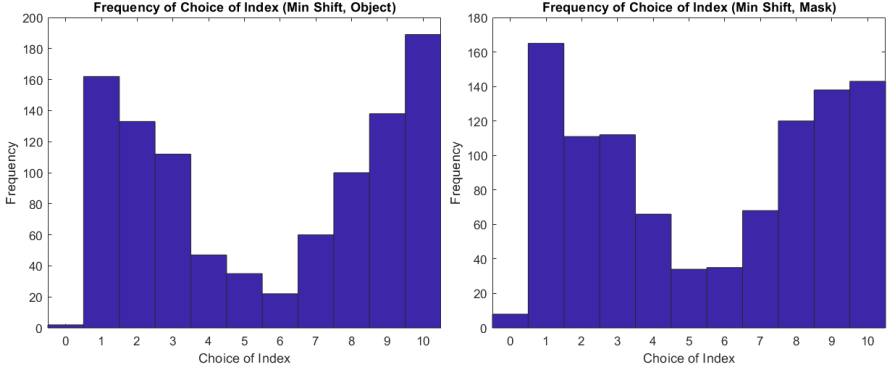
Next, we consider the frequency of the chosen index from performing the argmin function (step 6 of Algorithm 6) compared to the true minimizing indices for the object and mask separately.

Firstly, we have the frequency of the argmin indices.



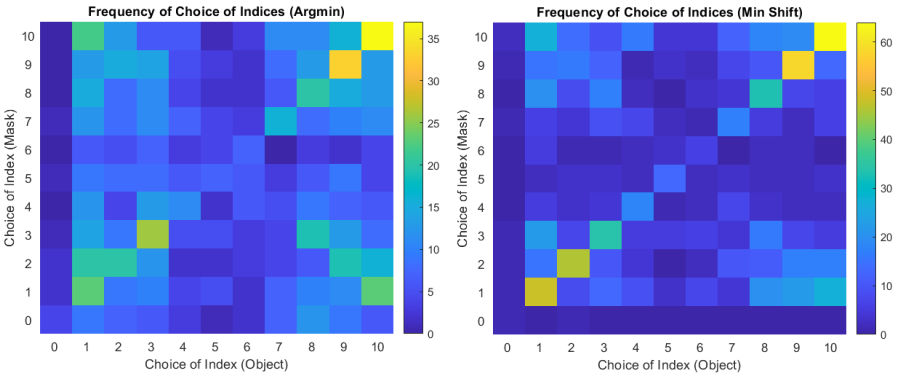
**Fig. 8**  $d = 2^6$ ,  $\delta = 6$ ,  $N = 4$ ,  $\mathbf{C}$  complex Gaussian. 1000 simulations. Frequency of index being chosen to compute Argmin Shift<sup>(x)</sup> and Argmin Shift<sup>(m)</sup>.

Secondly, we have the frequency of the min shift for both of the object and mask. Both of Figure 8 and Figure 9 were computed on the same 1000 tests.



**Fig. 9**  $d = 2^6$ ,  $\delta = 6$ ,  $N = 4$ ,  $\mathbf{C}$  complex Gaussian. 1000 simulations. Frequency of index being chosen to compute  $\text{Min Shift}^{(x)}$  and  $\text{Min Shift}^{(m)}$ .

Finally, we plot these choice of indices for both the Argmin Shift and Min Shift onto a two dimensional plot.



**Fig. 10**  $d = 2^6$ ,  $\delta = 6$ ,  $N = 4$ ,  $\mathbf{C}$  complex Gaussian. 1000 simulations. Frequency of indices being chosen to compute  $(\text{Argmin Shift}^{(x)}, \text{Argmin Shift}^{(m)})$  and  $(\text{Min Shift}^{(x)}, \text{Min Shift}^{(m)})$ .

## 5 Conclusions and Future Work

We have introduced an algorithm for recovering a specimen of interest from blind far-field ptychographic measurements. This algorithm relies on reformulating the measurements so that they resemble widely-studied blind deconvolutional measurements. This leads to transposed Khatri-Rao product estimates of our specimen which are then able to be recovered by angular synchronization. We then use these estimates in applying inverse Fourier transforms, point-wise division, and angular synchronization to recover estimates for the mask. Finally, we use a best error estimate sorting algorithm to find the final estimate of both the specimen and mask. As shown in numerical results, Algorithm 6 recovers both the sample and mask within a good margin

of error. It also provides stability under noise. A further goal for this research would be to adapt the existing recovery guarantee theorems for the selected blind deconvolutional recovery algorithm, in which the assumed Gaussian matrix  $\mathbf{C}$  is replaced with Khatri-Rao matrix  $\mathbf{C}'_{(k)} = \mathbf{C} \bullet S_k \tilde{\mathbf{C}}$ . In particular, this would mean providing alternate inequalities for the four key conditions laid out in Theorem 3.6.

## References

- [1] Hoppe, W.: Diffraction in inhomogeneous primary wave fields. *Acta Crystallogr. A* **25**, 495–501 (1969)
- [2] Hegerl, R., Hoppe, W.: Dynamische theorie der kristallstrukturanalyse durch elektronenbeugung im inhomogenen primärstrahlwellenfeld. *Berichte der Bunsengesellschaft für physikalische Chemie* **74**(11), 1148–1154 (1970)
- [3] Hegerl, R., Hoppe, W.: Phase evaluation in generalized diffraction (ptychography). *Proc. Fifth Eur. Cong. Electron Microscopy*, 628–629 (1972)
- [4] Clark, J.N., Huang, X., Harder, R.J., Robinson, I.K.: Continuous scanning mode for ptychography. *Optics letters* **39**(20), 6066–6069 (2014)
- [5] Huang, X., Lauer, K., Clark, J.N., Xu, W., Nazaretski, E., Harder, R., Robinson, I.K., Chu, Y.S.: Fly-scan ptychography. *Scientific reports* **5**(1), 1–5 (2015)
- [6] Odstrvcil, M., Holler, M., Guizar-Sicairos, M.: Arbitrary-path fly-scan ptychography. *Optics express* **26**(10), 12585–12593 (2018)
- [7] Edo, T., Batey, D., Maiden, A., Rau, C., Wagner, U., Pevsić, Z., Waigh, T., Rodenburg, J.: Sampling in x-ray ptychography. *Physical Review A* **87**(5), 053850 (2013)
- [8] Tsai, E.H., Usov, I., Diaz, A., Menzel, A., Guizar-Sicairos, M.: X-ray ptychography with extended depth of field. *Optics express* **24**(25), 29089–29108 (2016)
- [9] Pfeiffer, F.: X-ray ptychography. *Nature Photonics* **12**(1), 9–17 (2018)
- [10] Yang, H., Rutte, R., Jones, L., Simson, M., Sagawa, R., Ryll, H., Huth, M., Pennycook, T., Green, M., Soltau, H., *et al.*: Simultaneous atomic-resolution electron ptychography and z-contrast imaging of light and heavy elements in complex nanostructures. *Nature Communications* **7**(1), 1–8 (2016)
- [11] Gao, S., Wang, P., Zhang, F., Martinez, G.T., Nellist, P.D., Pan, X., Kirkland, A.I.: Electron ptychographic microscopy for three-dimensional imaging. *Nature communications* **8**(1), 1–8 (2017)
- [12] Jiang, Y., Chen, Z., Han, Y., Deb, P., Gao, H., Xie, S., Purohit, P., Tate, M.W.,

- Park, J., Gruner, S.M., *et al.*: Electron ptychography of 2d materials to deep sub-ångström resolution. *Nature* **559**(7714), 343–349 (2018)
- [13] Godard, P., Allain, M., Chamard, V.: Imaging of highly inhomogeneous strain field in nanocrystals using x-ray bragg ptychography: A numerical study. *Physical Review B* **84**(14), 144109 (2011)
- [14] Takahashi, Y., Suzuki, A., Furutaku, S., Yamauchi, K., Kohmura, Y., Ishikawa, T.: Bragg x-ray ptychography of a silicon crystal: Visualization of the dislocation strain field and the production of a vortex beam. *Physical Review B* **87**(12), 121201 (2013)
- [15] Hruszkewycz, S., Allain, M., Holt, M., Murray, C., Holt, J., Fuoss, P., Chamard, V.: High-resolution three-dimensional structural microscopy by single-angle bragg ptychography. *Nature materials* **16**(2), 244–251 (2017)
- [16] Li, P., Phillips, N.W., Leake, S., Allain, M., Hofmann, F., Chamard, V.: Revealing nano-scale lattice distortions in implanted material with 3d bragg ptychography. *Nature communications* **12**(1), 1–13 (2021)
- [17] Zheng, G., Horstmeyer, R., Yang, C.: Wide-field, high-resolution fourier ptychographic microscopy. *Nature photonics* **7**(9), 739–745 (2013)
- [18] Tian, L., Li, X., Ramchandran, K., Waller, L.: Multiplexed coded illumination for fourier ptychography with an led array microscope. *Biomedical optics express* **5**(7), 2376–2389 (2014)
- [19] Ou, X., Horstmeyer, R., Zheng, G., Yang, C.: High numerical aperture fourier ptychography: principle, implementation and characterization. *Optics express* **23**(3), 3472–3491 (2015)
- [20] Zheng, G., Shen, C., Jiang, S., Song, P., Yang, C.: Concept, implementations and applications of fourier ptychography. *Nature Reviews Physics* **3**(3), 207–223 (2021)
- [21] Merhi, S.E.: Phase Retrieval from Continuous and Discrete Ptychographic Measurements. Michigan State University, ??? (2019)
- [22] Ayers, G., Dainty, J.C.: Iterative blind deconvolution method and its applications. *Optics letters* **13**(7), 547–549 (1988)
- [23] Katsaggelos, A.K., Lay, K.-T.: Maximum likelihood blur identification and image restoration using the em algorithm. *IEEE Transactions on Signal Processing* **39**(3), 729–733 (1991)
- [24] Thiébaud, E., Conan, J.-M.: Strict a priori constraints for maximum-likelihood blind deconvolution. *JOSA A* **12**(3), 485–492 (1995)

- [25] Fish, D., Brinicombe, A., Pike, E., Walker, J.: Blind deconvolution by means of the richardson–lucy algorithm. *JOSA A* **12**(1), 58–65 (1995)
- [26] Molina, R., Katsaggelos, A.K., Abad, J., Mateos, J.: A bayesian approach to blind deconvolution based on dirichlet distributions. In: 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 4, pp. 2809–2812 (1997). IEEE
- [27] Kundur, D., Hatzinakos, D.: A novel blind deconvolution scheme for image restoration using recursive filtering. *IEEE transactions on signal processing* **46**(2), 375–390 (1998)
- [28] Carasso, A.S.: Direct blind deconvolution. *SIAM Journal on Applied Mathematics* **61**(6), 1980–2007 (2001)
- [29] Likas, A.C., Galatsanos, N.P.: A variational approach for bayesian blind image deconvolution. *IEEE transactions on signal processing* **52**(8), 2222–2233 (2004)
- [30] Ahmed, A., Recht, B., Romberg, J.: Blind deconvolution using convex programming. *IEEE Transactions on Information Theory* **60**(3), 1711–1732 (2013)
- [31] Fortunato, H.E., Oliveira, M.M.: Fast high-quality non-blind deconvolution using sparse adaptive priors. *The Visual Computer* **30**(6), 661–671 (2014)
- [32] Li, X., Ling, S., Strohmaier, T., Wei, K.: Rapid, robust, and reliable blind deconvolution via nonconvex optimization. *Applied and computational harmonic analysis* **47**(3), 893–934 (2019)
- [33] Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Understanding blind deconvolution algorithms. *IEEE transactions on pattern analysis and machine intelligence* **33**(12), 2354–2367 (2011)
- [34] You, Y.-L., Kaveh, M.: Blind image restoration by anisotropic regularization. *IEEE Transactions on Image Processing* **8**(3), 396–407 (1999)
- [35] Wu, S., Dong, C., Qiao, Y.: Blind image restoration based on cycle-consistent network. *IEEE Transactions on Multimedia* (2022)
- [36] Kundur, D., Hatzinakos, D.: Blind image restoration via recursive filtering using deterministic constraints. In: 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings, vol. 4, pp. 2283–2286 (1996). IEEE
- [37] Sroubek, F., Flusser, J.: Multichannel blind iterative image restoration. *IEEE Transactions on Image Processing* **12**(9), 1094–1106 (2003)

- [38] Thibault, P., Dierolf, M., Menzel, A., Bunk, O., David, C., Pfeiffer, F.: High-resolution scanning x-ray diffraction microscopy. *Science* **321**(5887), 379–382 (2008)
- [39] Thibault, P., Dierolf, M., Bunk, O., Menzel, A., Pfeiffer, F.: Probe retrieval in ptychographic coherent diffractive imaging. *Ultramicroscopy* **109**(4), 338–343 (2009)
- [40] Bunk, O., Dierolf, M., Kynde, S., Johnson, I., Marti, O., Pfeiffer, F.: Influence of the overlap parameter on the convergence of the ptychographical iterative engine. *Ultramicroscopy* **108**(5), 481–487 (2008)
- [41] Maiden, A.M., Rodenburg, J.M.: An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy* **109**(10), 1256–1262 (2009)
- [42] Maiden, A., Johnson, D., Li, P.: Further improvements to the ptychographical iterative engine. *Optica* **4**(7), 736–745 (2017)
- [43] Fannjiang, A., Chen, P.: Blind ptychography: uniqueness and ambiguities. *Inverse Problems* **36**(4), 045005 (2020)
- [44] Slyusar, V.: A family of face products of matrices and its properties. *Cybernetics and systems analysis* **35**(3), 379–384 (1999)

## Appendix A Sub-Sampling

In this section, we discuss sub-sampling lemmas that can be used in conjunction with Algorithm 1. In many cases, an illumination of the sample can cause damage to the sample, and applying the illumination beam (which can be highly irradiative) repeatedly at a single point can destroy it. Considering the risks to the sample and the costs of operating the measurement equipment, there are strong incentives to reduce the number of illuminations applied to any object.

**Definition A.1.** Let  $s \in \mathbb{N}$  such that  $s \mid d$ . We define the **sub-sampling operator**  $Z_s : \mathbb{C}^d \rightarrow \mathbb{C}^{\frac{d}{s}}$  defined component-wise via

$$(Z_s \mathbf{x})_n := x_{n \cdot s}, \quad \forall n \in [d/s]_0 \quad (\text{A1})$$

We now have an aliasing lemma which allows us to see the impact of performing the Fourier transform on a sub-sampled specimen.

**Lemma A.1. (Aliasing)** ([21], Lemma 2.0.1.) Let  $s \in \mathbb{N}$  such that  $s \mid d$ ,  $\mathbf{x} \in \mathbb{C}^d$ ,  $\omega \in [\frac{d}{s}]_0$ . Then we have that

$$\left(F_{\frac{d}{s}}(Z_s \mathbf{x})\right)_\omega = \frac{1}{s} \sum_{r=0}^{s-1} \hat{\mathbf{x}}_{\omega - r \frac{d}{s}} \quad (\text{A2})$$

*Proof* Let  $d \in \mathbb{N}$  and suppose  $s \in \mathbb{N}$  divides  $d$ . Let  $\mathbf{x} \in \mathbb{C}^d$  and  $\omega \in [\frac{d}{s}]_0$  be arbitrary. By the definition of the discrete Fourier transform and sub-sampling operator, we have that

$$\left(F_{\frac{d}{s}}(Z_s \mathbf{x})\right)_\omega = \sum_{n=0}^{\frac{d}{s}-1} (Z_s \mathbf{x})_n e^{-\frac{2\pi i n \omega}{d/s}} = \sum_{n=0}^{\frac{d}{s}-1} x_{ns} e^{-\frac{2\pi i n \omega}{d}} \quad (\text{A3})$$

By the inverse DFT and by collecting terms, we have that

$$\sum_{n=0}^{\frac{d}{s}-1} x_{ns} e^{-\frac{2\pi i n \omega}{d}} = \frac{1}{d} \sum_{n=0}^{\frac{d}{s}-1} \left( \sum_{r=0}^{d-1} \hat{x}_r e^{\frac{2\pi i r n s}{d}} \right) e^{-\frac{2\pi i n (r-\omega)s}{d}} \quad (\text{A4})$$

By treating this as a sum of DFTs, we then have that

$$\frac{1}{d} \sum_{n=0}^{\frac{d}{s}-1} \left( \sum_{r=0}^{d-1} \hat{x}_r e^{\frac{2\pi i r n s}{d}} \right) e^{-\frac{2\pi i n (r-\omega)s}{d}} = \frac{1}{s} \sum_{r=0}^{s-1} \hat{x}_{\omega + r \frac{d}{s}} = \frac{1}{s} \sum_{r=0}^{s-1} \hat{x}_{\omega - r \frac{d}{s}} \quad (\text{A5})$$

□

Before we start looking at aliased WDD, we need to introduce a lemma which will show the effect of taking a Fourier transform of an autocorrelation.

**Lemma A.2. (Fourier Transform Of Autocorrelation)** ([21], Lemma 2.0.2.) Let  $\mathbf{x} \in \mathbb{C}^d$  and  $\alpha, \omega \in [d]_0$ . Then

$$\left(\mathbf{F}_d(\mathbf{x} \circ S_\omega \bar{\mathbf{x}})\right)_\alpha = \frac{1}{d} e^{2\pi i \omega \alpha / d} \left(\mathbf{F}_d(\hat{\mathbf{x}} \circ S_{-\alpha} \tilde{\mathbf{x}})\right)_\omega \quad (\text{A6})$$

*Proof* Let  $\mathbf{x} \in \mathbb{C}^d$  and let  $\alpha, \omega \in [d]_0$  be arbitrary. By the convolution theorem, we have that

$$\left(\mathbf{F}_d(\mathbf{x} \circ S_\omega \bar{\mathbf{x}})\right)_\alpha = \frac{1}{d} (\hat{\mathbf{x}} *_d \mathbf{F}_d(S_\omega \bar{\mathbf{x}}))_\alpha \quad (\text{A7})$$

By technical equality (iii), we can revert the Fourier transform of the shift operator to the modulation operator of the Fourier transform

$$\frac{1}{d} (\hat{\mathbf{x}} *_d \mathbf{F}_d(S_\omega \bar{\mathbf{x}}))_\alpha = \frac{1}{d} (\hat{\mathbf{x}} *_d (W_\omega \hat{\mathbf{x}}))_\alpha = \frac{1}{d} \sum_{n=0}^{d-1} \hat{x}_n (W_\omega \hat{\mathbf{x}})_{\alpha-n} = \frac{1}{d} \sum_{n=0}^{d-1} \hat{x}_n \hat{x}_{\alpha-n} e^{\frac{2\pi i \omega (\alpha-n)}{d}} \quad (\text{A8})$$

with the latter equalities being the definition of the convolution and modulation. By applying reversals and using that  $\tilde{\mathbf{x}} = \mathbf{x}$ , we have that

$$\frac{1}{d} \sum_{n=0}^{d-1} \hat{x}_n \hat{x}_{\alpha-n} e^{\frac{2\pi i \omega (\alpha-n)}{d}} = \frac{1}{d} e^{\frac{2\pi i \omega \alpha}{d}} \sum_{n=0}^{d-1} \hat{x}_n \tilde{\mathbf{x}}_{n-\alpha} e^{-\frac{2\pi i \omega \alpha}{d}} \quad (\text{A9})$$

Finally by applying technical equality (vi) and using the definition of the shift operator and Hadamard product, we have that

$$\frac{1}{d} e^{\frac{2\pi i \omega \alpha}{d}} \sum_{n=0}^{d-1} \hat{x}_n \tilde{\mathbf{x}}_{n-\alpha} e^{-\frac{2\pi i \omega \alpha}{d}} = \frac{1}{d} e^{\frac{2\pi i \omega \alpha}{d}} \sum_{n=0}^{d-1} \hat{x}_n \tilde{\mathbf{x}}_{n-\alpha} e^{-\frac{2\pi i \omega \alpha}{d}} = \frac{1}{d} e^{2\pi i \omega \alpha / d} \left(\mathbf{F}_d(\hat{\mathbf{x}} \circ S_{-\alpha} \tilde{\mathbf{x}})\right)_\omega \quad (\text{A10})$$

□

## A.1 Sub-Sampling In Frequency

We will first look at sub-sampling in frequency.

**Definition A.2.** Let  $K$  be a positive factor of  $d$ , and assume that the data is measured at  $K$  equally spaced Fourier modes. We denote the set of Fourier modes of step-size  $\frac{d}{K}$  by

$$\mathcal{K} = \frac{d}{K} [K]_0 = \left\{0, \frac{d}{K}, \frac{2d}{K}, \dots, d - \frac{d}{K}\right\} \quad (\text{A11})$$

**Definition A.3.** Let  $\mathbf{A} \in \mathbb{C}^{d \times d}$  with columns  $\mathbf{a}_j$ ,  $K \mid d$ . We denote by  $\mathbf{A}_{K,d} \in \mathbb{C}^{K \times d}$  the sub-matrix of  $\mathbf{A}$  whose  $\ell^{\text{th}}$  column is equal to  $Z_{\frac{d}{K}}(\mathbf{a}_\ell)$ .

With these definitions, we will now convert the sub-sampled measurements into a more solvable form.



**Lemma A.3.** ([21], Lemma 2.1.1.) Suppose that the noisy spectrogram measurements are collected on a subset  $\mathcal{K} \subseteq [d]_0$  of  $K$  equally spaced Fourier modes. Then for any  $\omega \in [K]_0$

$$\left( (\mathbf{F}_K \mathbf{Y}_{K,d})^T \right)_\omega = K \sum_{r=0}^{\frac{d}{K}-1} (\mathbf{x} \circ S_{\omega-rK} \bar{\mathbf{x}}) *_d (\tilde{\mathbf{m}} \circ S_{rK-\omega} \tilde{\mathbf{m}}) + \left( (\mathbf{F}_K \mathbf{N}_{K,d})^T \right)_\omega$$

where  $\mathbf{Y}_{K,d} \in \mathbb{C}^{K \times d}$  –  $\mathbf{N}_{K,d} \in \mathbb{C}^{K \times d}$  is the matrix of sub-sampled noiseless  $K \cdot d$  measurements.

*Proof* For  $\ell \in [d]_0$ , the  $\ell^{\text{th}}$  column of the matrix  $\mathbf{Y}$  is

$$\mathbf{y}_\ell = \mathbf{F}_d \left( (\mathbf{x} \circ S_{-\ell} \mathbf{m}) *_d (\tilde{\mathbf{x}} \circ S_\ell \tilde{\mathbf{m}}) \right) + \eta_\ell \quad (\text{A12})$$

and thus for any  $\alpha \in [K]_0$

$$\left( Z_{\frac{d}{K}}(\mathbf{y}_\ell) \right)_\alpha = \left( \mathbf{F}_d \left( (\mathbf{x} \circ S_{-\ell} \mathbf{m}) *_d (\tilde{\mathbf{x}} \circ S_\ell \tilde{\mathbf{m}}) \right) \right)_{\alpha \frac{d}{K}} + \left( Z_{\frac{d}{K}}(\mathbf{J}_\ell) \right)_\alpha \quad (\text{A13})$$

and by aliasing lemma (with  $s = \frac{d}{K}$ )

$$\left( \mathbf{F}_K \left( Z_{\frac{d}{K}}(\mathbf{y}_\ell) \right)_\alpha \right)_\omega = \frac{K}{d} \sum_{r=0}^{\frac{d}{K}-1} (\hat{\mathbf{y}}_\ell)_{\omega-rK} \quad (\text{A14})$$

$$= d \cdot \frac{K}{d} \sum_{r=0}^{\frac{d}{K}-1} \left( (\mathbf{x} \circ S_{\omega-rK} \bar{\mathbf{x}}) *_d (\tilde{\mathbf{m}} \circ S_{rK-\omega} \tilde{\mathbf{m}}) \right)_\ell + \left( \mathbf{F}_K \left( Z_{\frac{d}{K}}(\mathbf{J}_\ell) \right)_\alpha \right)_\omega \quad (\text{A15})$$

The  $\ell^{\text{th}}$  column of  $\mathbf{Y}_{K,d} \in \mathbb{C}^{K \times d}$  is equal to  $Z_{\frac{d}{K}}(\mathbf{y}_\ell)$ . Then for any  $\omega \in [K]_0$ , the  $\omega^{\text{th}}$  column of  $(\mathbf{F}_K \mathbf{Y}_{K,d})^T \in \mathbb{C}^{d \times K}$  may be computed as

$$\left( (\mathbf{F}_K \mathbf{Y}_{K,d})^T \right)_\omega = K \sum_{r=0}^{\frac{d}{K}-1} (\mathbf{x} \circ S_{\omega-rK} \bar{\mathbf{x}}) *_d (\tilde{\mathbf{m}} \circ S_{rK-\omega} \tilde{\mathbf{m}}) + \left( (\mathbf{F}_K \mathbf{N}_{K,d})^T \right)_\omega \in \mathbb{C}^d \quad (\text{A16})$$

□

## A.2 Sub-Sampling In Frequency And Space

We will now look at sub-sampling in both frequency and space.

**Definition A.4.** Let  $L$  be a positive factor of  $d$ . Suppose measurements are collected at  $L$  equally spaced physical shifts of step-size  $\frac{d}{L}$ . We denote the set of shifts by  $\mathcal{L}$ , that is

$$\mathcal{L} = \frac{d}{L} [L]_0 = \left\{ 0, \frac{d}{L}, \frac{2d}{L}, \dots, d - \frac{d}{L} \right\} \quad (\text{A17})$$

**Definition A.5.** Let  $\mathbf{A} \in \mathbb{C}^{d \times d}$ ,  $L \mid d$ . We denote by  $\mathbf{A}_{d,L} \in \mathbb{C}^{d \times L}$  the sub-matrix of  $\mathbf{A}$  whose rows are those of  $\mathbf{A}$ , sub-sampled in step-size  $\frac{d}{L}$ .

We will now prove a similar lemma as before, but now we will sub-sample in both frequency and space.

**Lemma A.4.** ([21], Lemma 2.1.2.) *Suppose we have noisy spectrogram measurements collected on a subset  $\mathcal{K} \subseteq [d]_0$  of  $K$  equally spaced frequencies and a subset  $\mathcal{L} \subseteq [d]_0$  of  $L$  equally spaced physical shifts. Then for any  $\omega \in [K]_0, \alpha \in [L]_0$*

$$\begin{aligned} \left( \mathbf{F}_L \left( \mathbf{Y}_{K,L}^T (\mathbf{F}_K^T)_\omega \right) \right)_\alpha &= \frac{KL}{d} \sum_{r=0}^{\frac{d}{K}-1} \sum_{\ell=0}^{\frac{d}{L}-1} \left( \mathbf{F}_d((\mathbf{x} \circ S_{\omega-rK} \bar{\mathbf{x}}) *_d (\tilde{\mathbf{m}} \circ S_{rk-\omega} \tilde{\tilde{\mathbf{m}}})) \right)_{\alpha-\ell L} \\ &\quad + \left( \mathbf{F}_L \left( \mathbf{N}_{K,L}^T (\mathbf{F}_K^T)_\omega \right) \right)_\alpha \end{aligned}$$

where  $\mathbf{Y}_{K,L} - \mathbf{N}_{K,L} \in \mathbb{C}^{K \times L}$  is the matrix of sub-sampled noiseless  $K \cdot L$  measurements.

*Proof* For fixed  $\ell \in [d]_0, \omega \in [K]_0$ , we have computed

$$\left( \mathbf{F}_K(Z_{\frac{d}{K}}(\mathbf{y}_\ell)) \right)_\omega = K \sum_{r=0}^{\frac{d}{K}-1} \left( (\mathbf{x} \circ S_{\omega-rK} \bar{\mathbf{x}}) *_d \mathbf{F}_d(\tilde{\mathbf{m}} \circ S_{rk_\omega} \tilde{\tilde{\mathbf{m}}}) \right)_\ell$$

Fix  $\omega \in [K]_0$ , and define the vector  $\mathbf{p}_\omega \in \mathbb{C}^L$  by

$$(\mathbf{p}_\omega)_\ell := \left( \mathbf{F}_K(Z_{\frac{d}{K}}(\mathbf{y}_{\ell \frac{d}{L}})) \right)_\omega + \left( \mathbf{F}_K(Z_{\frac{d}{K}}(\mathbf{j}_{\ell \frac{d}{L}})) \right)_\omega, \quad \forall \ell \in [L]_0$$

Note that the rows of  $\mathbf{Y}_{K,L}, \mathbf{N}_{K,L} \in \mathbb{C}^{K \times L}$  are those of  $\mathbf{Y}_{K,L}, \mathbf{N}_{K,L} \in \mathbb{C}^{K \times d}$ , sub-sampled in step-size of  $\frac{d}{L}$ . Thus

$$(\mathbf{p}_\omega)_\ell = \left( (\mathbf{Y}_{K,L})^T (\mathbf{F}_K^T)_\omega \right)_\ell + \left( (\mathbf{N}_{K,L})^T (\mathbf{F}_K^T)_\omega \right)_\ell$$

where  $(\mathbf{F}_K^T)_\omega \in \mathbb{C}^K$  is the  $\omega^{\text{th}}$  column of  $\mathbf{F}_K^T$ . Therefore

$$\mathbf{p}_\omega = \mathbf{Y}_{K,L}^T (\mathbf{F}_K^T)_\omega + \mathbf{N}_{K,L}^T (\mathbf{F}_K^T)_\omega \in \mathbb{C}^L, \quad \forall \omega \in [K]_0$$

For any  $\ell \in [L]_0$ , we have

$$\begin{aligned} (\mathbf{p}_\omega)_\ell &= K \sum_{r=0}^{\frac{d}{K}-1} \left( (\mathbf{x} \circ S_{\omega-rK} \bar{\mathbf{x}}) *_d \mathbf{F}_d(\tilde{\mathbf{m}} \circ S_{rk_\omega} \tilde{\tilde{\mathbf{m}}}) \right)_{\ell \frac{d}{L}} + \mathbf{N}_{K,L}^T (\mathbf{F}_K^T)_\omega \\ &= K \cdot \left( Z_{\frac{d}{L}} \left( \sum_{r=0}^{\frac{d}{K}-1} (\mathbf{x} \circ S_{\omega-rK} \bar{\mathbf{x}}) *_d \mathbf{F}_d(\tilde{\mathbf{m}} \circ S_{rk_\omega} \tilde{\tilde{\mathbf{m}}}) \right) \right)_\ell + \mathbf{N}_{K,L}^T (\mathbf{F}_K^T)_\omega \end{aligned}$$

and for any  $\alpha \in [L]_0$ , by aliasing, we have that

$$(\mathbf{F}_L \mathbf{p}_\omega)_\alpha = \frac{KL}{d} \sum_{r=0}^{\frac{d}{K}-1} \sum_{\ell=0}^{\frac{d}{L}-1} \left( \mathbf{F}_d((\mathbf{x} \circ S_{\omega-rK} \bar{\mathbf{x}}) *_d (\tilde{\mathbf{m}} \circ S_{rk-\omega} \tilde{\tilde{\mathbf{m}}})) \right)_{\alpha-\ell L} + \left( \mathbf{F}_L (\mathbf{N}_{K,L}^T (\mathbf{F}_K^T)_\omega) \right)_\alpha$$

□

## Appendix B Alternative Approach

In this section we discuss the convex relaxation approach studied in [30].

### B.1 Convex Relaxation

In [30], the approach is to solve a convex version of the problem. Given  $\mathbf{y} \in \mathbb{C}^L$ , their goal is to find  $\mathbf{h} \in \mathbb{R}^k$ ,  $\mathbf{x} \in \mathbb{R}^N$  that are consistent with the observations. Making no additional assumptions other than the dimensions, the way to choose between multiple feasible points is by solving using least-squares. That is,

$$\text{minimize}_{\mathbf{u}, \mathbf{v}} \|\mathbf{u}\|_2^2 + \|\mathbf{v}\|_2^2 \quad \text{subject to} \quad \mathbf{y}(\ell) = \langle \mathbf{c}_\ell, \mathbf{u} \rangle \langle \mathbf{v}, \mathbf{b}_\ell \rangle, \quad 1 \leq \ell \leq L$$

This is a non-convex quadratic optimization problem. The cost function is convex, but the quadratic equality constraints mean that the feasible set is non-convex. The dual of this minimization problem is the SDP and taking the dual again will give us a convex program

$$\min_{\mathbf{W}_1, \mathbf{W}_2, \mathbf{X}} \frac{1}{2} \text{tr}(\mathbf{W}_1) + \frac{1}{2} \text{tr}(\mathbf{W}_2) \quad \text{subject to} \quad \begin{bmatrix} \mathbf{W}_1 & \mathbf{X} \\ \mathbf{X}^* & \mathbf{W}_2 \end{bmatrix} \succeq 0, \mathbf{y} = \mathcal{A}(\mathbf{X})$$

which is equivalent to

$$\min \|\mathbf{X}\|_* \quad \text{subject to} \quad \mathbf{y} = \mathcal{A}(\mathbf{X})$$

where  $\|\mathbf{X}\|_* = \text{tr}(\sqrt{\mathbf{X}^* \mathbf{X}})$  denotes the nuclear norm. In [30], they achieved guarantees for relatively large  $K$  and  $N$ , when  $B$  is incoherent in the Fourier domain, and when  $C$  is generic. We can now outline the algorithm from [30].

---

#### Algorithm 7 Convex Relaxed Blind Deconvolution Algorithm

---

**Input:** Normalized Fourier measurement  $\mathbf{y}$ ,

**Output:** Estimate underlying signal and blurring function

- 1) Compute  $\mathcal{A}^*(\mathbf{y})$
- 2) Find the leading singular value, left and right singular vectors of  $\mathcal{A}^*(\mathbf{y})$ , denoted by  $d$ ,  $\tilde{\mathbf{h}}_0$ , and  $\tilde{\mathbf{x}}_0$  respectively
- 3) Let  $\mathbf{X}_0 = \tilde{\mathbf{h}}_0 \tilde{\mathbf{x}}_0^*$  denote the initial estimate and solve the following optimization problem

$$\begin{aligned} \min \quad & \|\mathbf{X}\|_* \\ \text{subject to} \quad & \|\mathbf{y} - \mathcal{A}(\mathbf{X})\| \leq \delta \end{aligned} \tag{B18}$$

where  $\|\cdot\|_*$  denotes the nuclear norm and  $\|\mathbf{e}\|_2 \leq \delta$

Return  $(\mathbf{h}, \mathbf{x})$  for  $\mathbf{X} = \mathbf{h} \mathbf{x}^*$

---