DEVELOPMENT OF NOVEL ANALYSIS AND RECONSTRUCTION
TECHNIQUES FOR COHERENT OPTICAL IMAGING SYSTEMS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ÇAĞATAY IŞIL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

JULY 2019

Approval of the thesis:

## DEVELOPMENT OF NOVEL ANALYSIS AND RECONSTRUCTION TECHNIQUES FOR COHERENT OPTICAL IMAGING SYSTEMS

submitted by **ÇAĞATAY IŞIL** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences**  ——————

Prof. Dr. İlkay Ulusoy
Head of Department, **Electrical and Electronics Engineering**  ——————

Assist. Prof. Dr. Sevinç Figen Öktem
Supervisor, **Electrical and Electronics Engineering, METU**  ——————

Dr. Aykut Koç
Co-supervisor, **Aselsan Research Center**  ——————


**Examining Committee Members:**

Prof. Dr. A. Aydın Alatan
Electrical and Electronics Engineering, METU  ——————

Assist. Prof. Dr. Sevinç Figen Öktem
Electrical and Electronics Engineering, METU  ——————

Assoc. Prof. Dr. Tolga Çukur
Electrical and Electronics Engineering, Bilkent University  ——————

Assoc. Prof. Dr. Yeşim Serinağaoğlu
Electrical and Electronics Engineering, METU  ——————

Assist. Prof. Dr. Elif Vural
Electrical and Electronics Engineering, METU  ——————

Date: 09.07.2019

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

Name, Surname:   Çağatay Işıl

Signature       :

**ABSTRACT**

**DEVELOPMENT OF NOVEL ANALYSIS AND RECONSTRUCTION
TECHNIQUES FOR COHERENT OPTICAL IMAGING SYSTEMS**

Işıl, Çağatay

M.S., Department of Electrical and Electronics Engineering

Supervisor: Assist. Prof. Dr. Sevinç Figen Öktem

Co-Supervisor: Dr. Aykut Koç

July 2019, 74 pages

We develop novel analysis and reconstruction techniques for coherent optical imaging systems. Firstly, we present a phase-space approach to analyze coherent imaging systems with multiple diffracting apertures. The degrees of freedom of a coherent imaging system can be computed from its phase-space window, which takes into account diffraction effects from all apertures. We show how the phase-space window is linked to important imaging parameters of the system such as diffraction-limited resolution. A single-lens system and a microscope objective design are considered as examples to illustrate the utility of the approach. Secondly, we focus on the classical phase retrieval problem, which is a fundamental problem in coherent imaging. Although there are several well-known phase retrieval algorithms, the reconstruction performance is generally sensitive to initialization and measurement noise. We develop two different novel phase retrieval algorithms by jointly exploiting deep neural networks (DNNs) and traditional model-based inversion methods. The used model-based inversion approach is the well-known hybrid-input-output (HIO) method for phase retrieval. In the first approach, the main idea is to use a DNN in an iterative

manner with the HIO method to improve the HIO reconstructions. Numerical results demonstrate the effectiveness of this approach, which also has little additional computational cost compared to the HIO method. In the second approach, the main idea is to incorporate a learning-based prior to the HIO method through plug-and-play regularization. The developed method is flexible such that it can also be used with different image priors. The performance of the second approach is illustrated with numerical simulations. Both of the developed phase retrieval methods not only achieve state-of-the-art reconstruction performance but also are more robust to different initialization and noise levels.

# ÖZ

## FAZ UYUMLU OPTİK GÖRÜNTÜLEME SİSTEMLERİ İÇİN YENİLİKÇİ ANALİZ VE GERİÇATIM YAKLAŞIMLARININ GELİŞTİRİLMESİ

Işıl, Çağatay

Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü

Tez Yöneticisi: Dr. Öğr. Üyesi. Sevinç Figen Öktem

Ortak Tez Yöneticisi: Dr. Aykut Koç

Temmuz 2019 , 74 sayfa

Bu tezde faz uyumlu optik görüntüleme sistemleri için yenilikçi analiz ve geriçatım metotları geliştirilmektedir. İlk olarak kırınıma neden olan birden fazla açıklıklı faz uyumlu görüntüleme sistemlerini analiz etmek için bir faz-uzay yaklaşımı geliştirilmektedir. Bir faz uyumlu görüntüleme sisteminin serbestlik derecesi, onun birden fazla açıklıktan gelen kırınım etkilerini içeren faz-uzay penceresini kullanılarak hesaplanabilir. Bu çalışmada, faz-uzay penceresi kullanılarak bir sistemin önemli görüntüleme parametrelerinin nasıl elde edildiği gösterilmektedir. Geliştirilen metodun sağladığı fayda, tek mercekli ideal bir sistem ve optik tasarımı bilinen bir mikroskop merceği için gösterilmektedir. İkinci olarak faz uyumlu görüntüleme sistemlerindeki temel problemlerden biri olan faz geri kazanımı problemine odaklanılmaktadır. Bu problem için sıkça kullanılan birçok algoritma olmasına rağmen, bunların geriçatım performansı genellikle ilklendirmeye ve ölçüm gürültüsüne bağlıdır. Bu çalışmada, model tabanlı geleneksel metotlar ve derin sinir ağları kullanılarak iki tane yenilikçi faz geri kazanımı algoritması geliştirilmektedir. Kullanılan model tabanlı model, sıkça tercih edilen karma girdi-çıktı yöntemidir. İlk yaklaşımdaki ana fikir, derin sinir

ağlarıyla karma girdi-çıktı yöntemiyle özyinelemeli bir şekilde kullanarak geriçatım sonuçlarını geliştirmektir. Sayısal sonuçlar, karma girdi-çıktı yöntemine göre hesaplama masrafı çok az daha fazla olan yaklaşımın faydalarını göstermektedir. İkinci yaklaşımdaki ana fikir, karma girdi-çıktı yöntemi ve derin öğrenme tabanlı önsel bilgileri tak ve çalıştır düzenlileştirmesi yardımıyla birleştirmektir. Geliştirilen metot farklı görüntü önsel bilgilerle de kullanılabilme esnekliğine sahiptir. Geliştirilen yaklaşımın faydaları sayısal benzetimlerle gösterilmektedir. Geliştirilen faz gerikazanımı yaklaşımları, hem en gelişkin geriçatım performansı göstermekte hem de farklı ilklendirme ve gürültü seviyelerine karşı daha gürbüz davranmaktadır.

Anahtar Kelimeler: faz uyumlu optik görüntüleme sistemleri, faz-uzay optiği, Fourier optiği, ters problemler, faz gerikazanımı, görüntü geriçatımı, derin öğrenme

to my family

# ACKNOWLEDGMENTS

# TABLE OF CONTENTS

# LIST OF TABLES

TABLES

xiv

# CHAPTER 1

# INTRODUCTION

Fourier optics is an important branch of optics that provides great utility in the modeling and analysis of optical imaging systems [3]. If the illumination used in an imaging system owns a property called spatial coherence, then these systems are referred as coherent imaging systems and the light through the system is described as a spatial distribution of complex-valued field amplitude. In fact, spatial coherence is defined as a measure of the correlation between the phase of a light wave observed at different spatial locations [4]. In this thesis, the focus is on coherent imaging systems.

Coherent imaging is important for several imaging applications including microscopy and holography [3]. The reason is that using coherent light sources such as lasers for illumination enables to measure the phase information, which has physical significance in the study of the structure of objects. For example, phase-contrast microscopy is used to see the transparent objects, which are not seen directly. This imaging method, which is invented by Zernike, uses coherent illumination. It contains a filter that converts phase modulation to amplitude modulation. Another example is holography, which is invented by Gabor. In holography, the phase information is converted to amplitude modulation of another wavefront and the intensity of this wavefront is measured [5]. Coherent illumination is predominantly used for holography [3]. Thanks to these inventions, Zernike and Gabor were awarded Nobel prizes in physics in 1953 and 1971, respectively.

Coherent imaging systems often contain apertures that cause diffraction. The term diffraction was defined by Sommerfeld as "any deviation of light rays from rectilinear paths which cannot be interpreted as reflection or refraction." [3]. Diffraction is generally studied to relate the complex amplitude distribution of a propagating wave at one

plane to a subsequent plane. An aberration-free optical system is called diffraction-limited because diffraction from finite-size apertures limits the spatial frequency response of the system and consequently sets a limit for resolution. Diffraction limit to resolution is often determined through the cutoff of the computed frequency response, or using the shape of the point spread function (PSF), which corresponds to the impulse response of the system,

In this thesis, firstly, a novel phase-space approach is developed to analyze diffraction-limited coherent imaging systems with multiple diffracting apertures. The approach is based on the concept of phase-space (space-frequency) window [6], which characterizes the degrees of freedom of a coherent imaging system in the space-frequency plane. We show how the space-frequency window is linked to important imaging parameters of the system including cut-off frequency, diffraction-limited resolution, effective focal length, and magnification. Different than the widely used geometrical-optics methods for analysis, the proposed approach takes into account diffraction effects from all apertures, and also requires a simple computation. To illustrate the utility of the approach, a single-lens system and an objective lens of a microscope are analyzed, and the results are confirmed with the known properties and specifications of the inspected systems.

Coherent imaging systems consisting of lenses and other optical elements suffer from physical limitations that result from aberrations and diffraction. Even if an optical system is free of aberrations, there is still an ultimate limit for resolution, due to diffraction resulting from the wave nature of the light. To reach resolution beyond the diffraction-limit, several different techniques have been developed [3]. One such popular approach is coherent diffractive imaging (CDI). This is a lensless imaging technique that forms images using computational phase retrieval algorithms.

In CDI, a coherent light source is used to illuminate an object and the diffraction pattern that corresponds to the Fourier transform of this object in the far-field is measured. However, only the intensity of the diffraction pattern can be measured with practical light detectors such as charge-coupled devices (CCDs) [7]. Since only Fourier intensity measurements are available, a phase retrieval problem needs to be solved to recover the object distribution from these measurements. CDI with an X-ray

source is illustrated in Fig. 1.1.
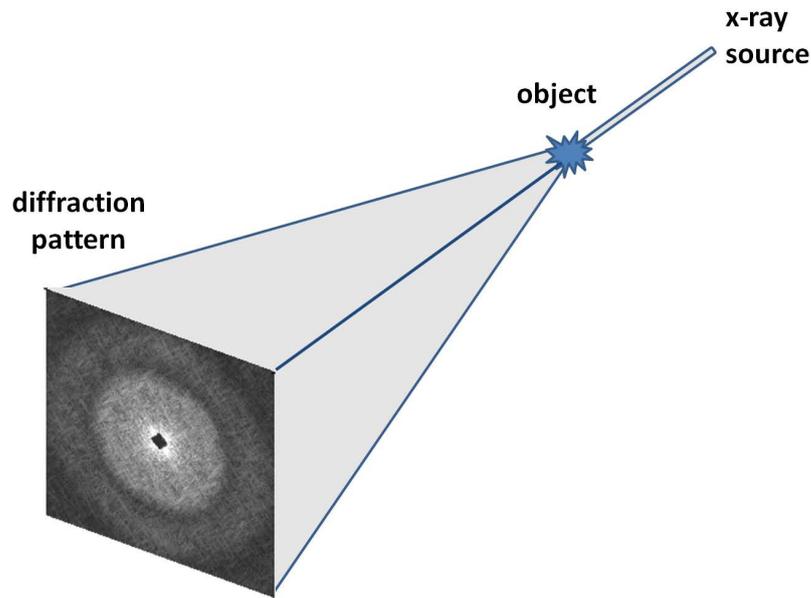


Figure 1.1: The illustration of coherent diffraction imaging with an X-ray source

The classical phase retrieval problem is the recovery of a constrained signal from the magnitude of its Fourier transform, or equivalently from its autocorrelation. This problem is encountered in a variety of applications in science and engineering such as crystallography [8], microscopy [9, 10], astronomy [11], optical imaging [12, 13], and speech processing [14]. Throughout history, many scientists were awarded Nobel prizes for their works involving phase retrieval. For example, James Watson and Francis Crick discovered the structure of the DNA molecule by solving this problem from the diffraction images obtained by Rosalind Franklin. For this work, they were awarded the 1962 Nobel prize in medicine. In 1953, Herbert Hauptman and Jerome Karle demonstrated how to find a solution for this problem in X-ray crystallography by using prior information about the molecules composing the crystal and then they received the 1985 Nobel prize in chemistry.

In the second part of this thesis, the classical phase retrieval problem is studied and two novel methods are developed for the solution of this problem. Existing phase retrieval algorithms including the popular alternating projection-based methods [7, 15] generally lead to artifacts due to stuck in local minima or amplification of noise in the solution. The developed methods in this thesis aim to overcome the limitations

of these methods.

The first method is a novel deep learning-based approach that utilizes DNNs with a model-based inversion approach. Here, the used model-based inversion approach is the well-known hybrid input-output (HIO) method, which incorporates the physical model and the constraints into the solution, but may lead to artifacts. The main idea in the developed method is to use a DNN in an iterative manner with the HIO method to improve the HIO reconstructions. The developed approach consists of two main stages: the iterative DNN-HIO stage and the final DNN stage. For the iterative DNN-HIO stage, a DNN is trained to remove the HIO artifacts. This trained DNN is then used iteratively with the HIO method to generate an intermediate reconstruction. In the final stage, the intermediate reconstructions are used to train a second DNN to remove the remaining artifacts. The performance of the developed approach is compared with classical and state-of-the-art methods through numerical simulations. The results demonstrate the effectiveness of our approach, which has relatively little additional computational cost compared to HIO.

In the second approach, the main idea is to incorporate a learning-based prior to the HIO method through plug-and-play regularization [16, 17]. Half quadratic splitting (HQS) is used to decouple the inverse problem into two separate sub-problems containing the data fidelity and prior terms [18, 19]. The sub-problem containing the data fidelity term is solved in an ad hoc manner with the HIO method for a small number of iterations. The other sub-problem containing the prior term is solved with a deep learning-based denoiser. The iterative solution of these sub-problems with a varying split parameter provides the final reconstruction. The developed method is flexible such that it can be used with different image priors coming from different denoisers. Moreover, this method can be used for different Fourier magnitude measurements including coded diffraction patterns and Fourier magnitude measurements. Its reconstruction performance is compared with classical and state-of-the-art methods through numerical experiments. Both of the developed approaches achieve state-of-the-art reconstruction performance, and also more robust to different initialization and noise levels.

The thesis is organized as follows. In Chapter 2, the analysis of coherent optical

systems with a phase-space approach is presented. The first phase retrieval approach corresponding to a hybrid method utilizing DNNs with the HIO method is presented in Chapter 3. The second phase retrieval approach utilizing plug-and-play priors is developed in Chapter 4. Finally, we summarize the thesis and conclude in Chapter 5.

**CHAPTER 2**

**A PHASE-SPACE APPROACH TO ANALYZE COHERENT IMAGING SYSTEMS**

## 2.1 Introduction

Coherent imaging systems are widely analyzed using conventional approaches based on geometrical optics [22, 3]. Although there are also many works on determining the object-image relationship of an imaging system based on Fourier optics [22, 23, 3], to the best of our knowledge, there has been no development of a simple and a straightforward method for quantifying the important imaging parameters such as cut-off frequency, resolution, and effective focal length.

In this chapter, we present a simple phase-space approach to analyze coherent imaging systems with multiple diffracting apertures. The approach is based on the concept of *phase-space (space-frequency) window* [6], which characterizes the degrees of freedom of a coherent imaging system in the space-frequency plane. We show how the space-frequency window is linked to important imaging parameters of the system including cut-off frequency, diffraction-limited resolution, effective focal length, and magnification. Different than the widely used geometrical-optics methods for analysis, the proposed approach takes into account diffraction effects from *all* apertures, and also requires a simple computation. To illustrate the utility of the approach, a single-lens system and an objective lens of a microscope are analyzed, and the results are confirmed with the known properties and specifications of the inspected systems. Preliminary version of this approach was presented in [20].

The rest of this chapter is organized as follows. The space-frequency window is

described in Section 2.2. The analysis of coherent optical systems with the phase-space approach is discussed in Section 2.3. Also, in this section, a single-lens system and a microscope objective are analyzed. Finally, we summarize conclude in Section 2.4.

## 2.2 Phase-space (Space-frequency) Window

Our approach is developed for a first-order coherent imaging system (quadratic-phase system) with the arbitrary number of apertures. Each aperture is assumed to be centered around the origin and free of obscuration. Light propagation from one aperture to the next one through this first-order system can be modeled by a linear canonical transform (LCT). For simplicity, we consider one-dimensional signals and systems. In the one-dimensional case, LCT is a three-parameter family of linear integral transforms, whose definition is as follows [24, 25]:

$$
f_{\mathbf{T}}(x) \equiv (C_{\mathbf{T}}f)(x) \equiv \int_{-\infty}^{\infty} C_{\mathbf{T}}(x, x')f(x')dx',
$$

$$
C_{\mathbf{T}}(x, x') \equiv \sqrt{\frac{1}{B}}e^{-i\pi/4}e^{i\pi(\frac{D}{B}x^2 - 2\frac{1}{B}xx' + \frac{A}{B}x'^2)},
$$

(21)

for $B \neq 0$, where $f_{\mathbf{T}}(x)$ denotes the output of a first order optical system, $f(x)$ represents the input of the system, $C_{\mathbf{T}}$ is the unitary LCT operator with system parameter matrix $\mathbf{T} = [AB; CD]$ and $AD - BC = 1$. In the trivial case $B = 0$, the LCT is defined as $f_{\mathbf{T}}(x) \equiv \sqrt{D}\exp(i\pi CDx^2)f(Dx)$. The parameters of the LCT operator is commonly arranged in a $2 \times 2$ unitary matrix $\mathbf{T} = [A\,B;\,C\,D]$, which corresponds to the ray (ABCD) matrix in geometrical optics.

Examples of LCT parameter matrices frequently encountered in imaging are the scaling matrix $\mathbf{M}_M$, chirp multiplication matrix $\mathbf{Q}_q$, and chirp convolution matrix $\mathbf{R}_r$, whose general forms are given as follows [26]:

$$
\mathbf{M}_M = \begin{bmatrix} M & 0 \\ 0 & 1/M \end{bmatrix}, \quad \mathbf{Q}_q = \begin{bmatrix} 1 & 0 \\ -q & 1 \end{bmatrix}, \quad \mathbf{R}_r = \begin{bmatrix} 1 & r \\ 0 & 1 \end{bmatrix}.
$$

(22)

Here the matrix $\mathbf{M}_M$ corresponds to the scaling operation, which maps a function $f(x)$ into $\sqrt{1/M}f(x/M)$ with the scaling parameter $M > 0$. The matrix $\mathbf{Q}_q$ performs multiplication with a chirp function of the form $\exp[-i\pi qx^2]$, which corre-

sponds to refraction. For example, $\mathbf{Q}_q$ with $q = 1/\lambda f$ models passage through a thin lens of the focal length of $f$, and $\mathbf{Q}_q$ with $q = (n_2 - n_1)/\lambda R$ represents refraction from a surface with spherical curvature of radius $R$ when the refractive indices of the input and output media are $n_1$ and $n_2$, respectively. Moreover, the matrix $\mathbf{R}_r$ corresponds to the general convolution operation with a chirp function of the form $e^{-i\pi/4}\sqrt{1/r}\exp[i\pi x^2/r]$. This models propagation through a medium of length $d$ and refractive index $n$ in the Fresnel approximation when $r = \lambda d/n$. An important special case is free-space propagation with $n = 1$. These parameter matrices are useful in the analysis of imaging systems because when several subsystems are cascaded, the overall system can be characterized by a single LCT integral whose parameters can be obtained by the multiplication of the corresponding matrices for each subsystem.

To analyze first-order coherent imaging systems, our approach is developed based on phase-space optics. Phase-space distributions provide the distribution of the signal energy over space and frequency. A commonly used phase-space (space-frequency) distribution is the Wigner distribution [24]. The space-frequency region that contains the substantial portion of the energy of a signal is commonly referred as the phase-space support of the signal. This is also related to the widely used concept of *space-bandwidth product*, or equivalently the number of degrees of freedom [27]. Recently, a simple procedure has been developed [6] to compute the *phase-space (space-frequency) window* of a system, which corresponds to the largest signal support in the space-frequency plane that can pass through an optical system without any information loss. That is, when the space-frequency support of an input signal does not lie inside the system window, the parts of the signal within the window pass while the parts lying outside are blocked, a result that is approximately valid for many systems of practical interest. Hence we can view the system space-frequency window as an equivalent aperture that combines the effects of all individual apertures in different planes into a single space-frequency aperture. Moreover, the area of the space-frequency window corresponds to the degrees of freedom of the imaging system.

The phase-space (space-frequency) window is computed using the aperture sizes of the optical elements in the system together with the LCT parameter matrices that

model the propagation from each aperture to the next one. This computation relies on the fact that each aperture corresponds to a truncation of the input signal in an oblique axis in the space-frequency plane [26]. Thus each aperture introduces a restriction on the signal, which can be geometrically visualized in the form of a corridor in the space-frequency plane. Intersecting all the corridors corresponding to all the apertures gives us the phase-space window of the system [6]. This simple procedure is illustrated in the space-frequency plane in Fig. 2.1 for a system with four apertures. Here, $x$ represents the spatial coordinate, and $\sigma_x$ denotes the spatial frequency coordinate.
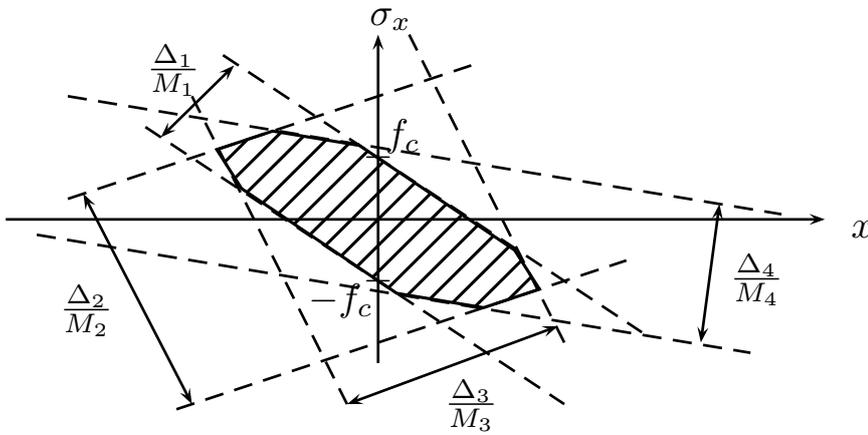


Figure 2.1: Illustration of the space-frequency window for a system with four limiting apertures.

To clearly describe how the space-frequency window is computed, let us define the input and output planes in the imaging system as $z = 0$ and $z = d$, where $d$ is the length of the system and $z$ denotes the optical axis. Let $L$ represent the total number of apertures in the system, $z_j$ and $\Delta_j$ denote the location and extent of the $j^{th}$ aperture in the system, $j = 1, 2, ..., L$. The matrix $\mathbf{T}_j$ is used to denote the parameter matrix of the system from the input to the position of the $j$th aperture, that is the system lying between 0 and $z_j$ excluding the apertures. The matrix $\mathbf{T}_j$ can be readily computed using the respective parameter matrices for each subsystem corresponding to different optical processes such as refraction and propagation through a medium, as given in Eqn. 22, together with the concatenation property of these matrices.

Let us denote the elements of the matrix $\mathbf{T}_j$ by $A_j, B_j, C_j$, and $D_j$. Alternative to

10

the ABCD matrix, an LCT can be equivalently characterized in terms of its Iwasawa decomposition parameters, which are required for the computation of the system window. In the Iwasawa decomposition of interest, an arbitrary LCT is decomposed into a fractional Fourier transform (FRT) of order $a$, followed by scaling with magnification $M$, and followed lastly by chirp multiplication with parameter $q$. The associated Iwasawa decomposition parameters, $a$, $M$, and $q$, can be computed from the matrix entries $A, B, C$, and $D$ using the following well-known formulas [25, 26]:

$$
a = \begin{cases} \frac{2}{\pi}\arctan(\frac{B}{A}), & \text{if } A \geq 0 \\ \frac{2}{\pi}\arctan(\frac{B}{A}) + 2, & \text{if } A < 0 \end{cases} \tag{23}
$$

$$
M = \sqrt{A^2 + B^2} \tag{24}
$$

$$
q = \begin{cases} -\frac{C}{A} - \frac{B/A}{A^2+B^2}, & \text{if } A \neq 0 \\ -\frac{D}{B}, & \text{if } A = 0 \end{cases} \tag{25}
$$

Here the range of the arctangent lies in $(-\pi/2, \pi/2]$, and the dimensional to dimensionless conversion parameter is chosen as $s = 1$ in our measurement unit (i.e. meters). The FRT order $a$ in the Iwasawa decomposition begins from $0$ at the input of the system, and then monotonically increases as a function of distance [25, 26].

By denoting the Iwasawa decomposition parameters for each matrix $\mathbf{T}_j$ by $a_j, M_j$ and $q_j$, the steps for computing the phase-space (space-frequency) window at the input plane ($z = 0$) are summarized as follows [6]:

1. Compute the LCT parameter matrix $\mathbf{T}_j$ for each aperture $j = 1, 2, ..., L$ by multiplying the respective parameter matrices for each subsystem corresponding to different optical processes such as refraction and propagation through a medium.

2. Compute the corresponding Iwasawa decomposition parameters $a_j$ and $M_j$ (the fractional order and the magnification) using the formulas in Eqn. 23, 24 and 25.

3. For each aperture $j$, draw a corridor of width $\Delta_j/M_j$ making angle $(a_j + 1)\pi/2$ with the $x$-axis in the phase-space (space-frequency) plane. (The cor-

ridor is explicitly defined by the following two lines: $y = -\cot(a_j\pi/2)x \pm \frac{\Delta_j}{2M_j}\csc(a_j\pi/2)$.)

4. Intersect the corridors from all apertures to find the phase-space (space-frequency) window.

## 2.3 Analysis of Coherent Imaging Systems

We now show how this space-frequency window can be used to analyze coherent imaging systems. The important imaging parameters such as cut-off frequency, diffraction-limited resolution, effective focal length, and magnification are obtained utilizing this system window and its parameters.

Conventionally, the cut-off frequency, $f_c$, is obtained from the optical transfer function (OTF) of the imaging system. OTF is the Fourier transform of system's point spread function (PSF), which corresponds to the response of the system to a point source (namely, an impulse). In the space-frequency plane, the support of a point source containing all spatial frequencies (i.e. impulse) is the line corresponding to the frequency axis. As a result, the intersection of this line with the system space-frequency window approximately provides the support of the OTF. Hence the boundary of the space-frequency window on the spatial frequency axis provides the cut-off frequency $f_c$, as illustrated in Fig. 2.1.

The diffraction-limited resolution can be obtained from the cut-off frequency $f_c$. In fact, the product of diffraction-limited resolution and $f_c$ is a constant [28]. For example, if Rayleigh criterion is used for defining the resolution $\Delta l$, then $\Delta l$ is equal to $0.61/f_c$ for circular apertures and $0.5/f_c$ for rectangular apertures in the coherent case. Similarly, the related parameter of numerical aperture $NA$ can be obtained from $f_c$ as [29, 28]

$$NA = \lambda f_c = \frac{0.61\lambda}{\Delta l},\qquad(26)$$

where the last equality holds for the circular aperture case.

As part of the space-frequency window computation, the overall LCT parameter matrix, $\mathbf{T}_L$, of the system is obtained. Let $\mathbf{T}_L = [A_L\,B_L;\,C_L\,D_L]$. The effective focal

12

length, $EFL$, and overall magnification, $M_L$, of the system can be obtained from the elements of the matrix $\mathbf{T}_L$. Because $EFL$ parameter can be extracted from the overall ray-transfer matrix of the imaging system [30], by using the relation between the ray-transfer matrix and the LCT matrix, the parameter $EFL$ can be obtained as follows:

$$EFL = -\frac{1}{\lambda C_L}. \tag{27}$$

The magnification parameter also follows from the overall system matrix $\mathbf{T}_L$. In fact, the overall magnification from the input plane to any plane within the system is given by the scaling parameter, $M$, of the Iwasawa decomposition of the corresponding LCT matrix up to that plane, excluding the apertures. Hence the magnification, $M_L$, of the imaging system can be obtained from the scaling parameter of the matrix $\mathbf{T}_L$ as follows:

$$M_L = \sqrt{A_L^2 + B_L^2}. \tag{28}$$

Using the obtained space-frequency window, one can also compute the number of degrees of freedom of the system from the area of the system window. This corresponds to the minimum number of samples required to faithfully represent the output of this imaging system. Hence, this number is also closely related to the concept of resolution. For example, if a detector is used to record the output image, the pixel size of the detector can be chosen based on this. Increasing the number of degrees of freedom of the imaging system can be one of the ultimate goals in design, which will consequently enhance the resolution.

Phase-space window computation also reveals which apertures are limiting in the system. In the conventional approaches, the analysis of the system is performed based on the most limiting aperture. However, in this phase-space approach, all the limiting apertures can be determined. The limiting apertures are the ones that determine the space-frequency window; that is, the space-frequency window is the intersection of the corridors defined by the limiting apertures. The apertures that do not contribute to the final space-frequency window are non-limiting, which means that decreasing their sizes to a certain point will not affect the imaging performance of the system. This may provide flexibility during the design phase of the imaging system.

13

### 2.3.1 Analysis of a Single Lens System

To illustrate the utility of the developed approach, we first analyze a simple single-lens system and confirm the consistency of the analysis results with the well-known properties of the system. The single-lens system involving a thin lens is illustrated in Fig. 2.2. Here $d_s$ and $d_i$ respectively denote the distances from the source and image planes to the thin lens, and $f$ is the focal length of the lens. The object is assumed to have finite circular support with a diameter of $D_o$. The aperture of the thin lens is also assumed to be symmetric, in particular, circular-shaped with a diameter $D$.
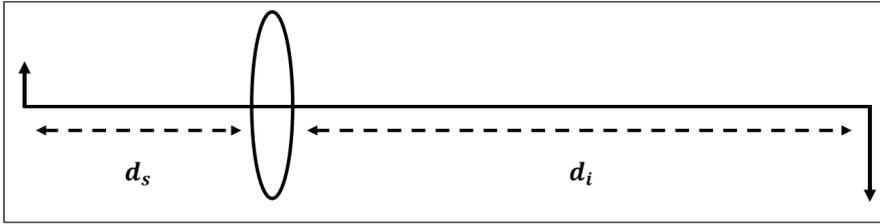


Figure 2.2: Illustration of the thin lens system

To compute the space-frequency window of the thin lens system, the LCT parameter matrices up to each aperture location $z_j$ (source plane, lens plane, and image plane) are obtained by using the chirp multiplication matrix $\mathbf{Q}_q$ with $q = 1/\lambda f$ corresponding to refraction from a thin lens and the chirp convolution matrix $\mathbf{R}_r$ with $r = \lambda d_s$ and $r = \lambda d_i$ corresponding to free space propagation. The resulting matrices $T_j$ for each $j$ are given by

$$\mathbf{T}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \mathbf{T}_2 = \begin{bmatrix} 1 & \lambda d_s \\ 0 & 1 \end{bmatrix}, \ \mathbf{T}_3 = \begin{bmatrix} -d_i/d_s & 0 \\ -1/\lambda f & -d_s/d_i \end{bmatrix}. \tag{29}$$

Then, the Iwasawa decomposition parameters of each $\mathbf{T}_j$ are computed as given in Table 2.1. When the corridors corresponding to the two apertures (i.e. the support of the object and the lens aperture) are intersected by using these decomposition parameters together with the size of apertures, we obtain the space-frequency window shown in Fig. 2.3.

The imaging parameters of the thin lens system are obtained from the overall system matrix $\mathbf{T}_3$ and the computed system window using Eqn. 26, 27 and 28. First of all,
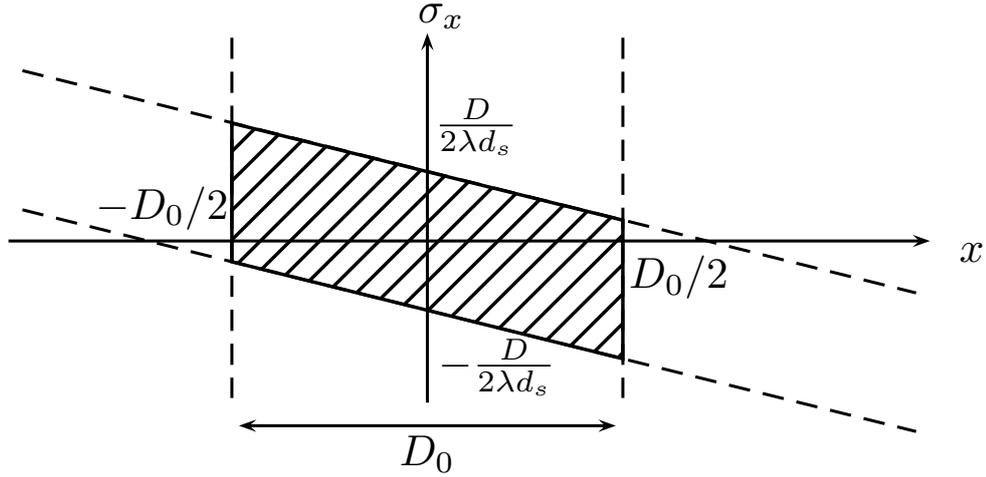
Figure 2.3: The input space-frequency window of the thin lens imaging system (the effects of the thin lens imaging system on $s(x, y)$ in the space-frequency plane).

Table 2.1: The Iwasawa decomposition parameters $a_j$,$M_j$ and $q_j$ of each LCT system matrix $\mathbf{T}_j$ for the thin lens imaging system

| $\mathbf{T}_j$ | $a_j$ | $M_j$ | $q_j$ |
|---|---|---|---|
| $\mathbf{T}_1$ for input plane | $0$ | $1$ | $0$ |
| $\mathbf{T}_2$ for thin lens plane | $\frac{2}{\pi}arctan(\lambda d_s)$ | $\sqrt{1 + (\lambda d_s)^2}$ | $\frac{-\lambda d_s}{1+(\lambda d_s)^2}$ |
| $\mathbf{T}_3$ for output plane | $2$ | $\frac{d_i}{d_s}$ | $\frac{d_i+d_s}{\lambda d_i^2}$ |

the cut-off frequency $f_c$ is obtained as $\frac{D}{2\lambda d_s}$ from the boundary of the space-frequency window on the frequency axis. The parameter $NA$ of the system is also computed as $\frac{D}{2d_s}$ by inserting $f_c$ in Eqn. 26. The effective focal length $EFL$ follows as $f$ by using $\mathbf{T}_3$ and Eqn. 27. The magnification $M_L$ is computed as $d_i/d_s$ using Eqn. 28. Hence the true imaging parameters of the thin lens system can be obtained using the developed approach. Moreover, the number of degrees of freedom of this system is given by $DD_o/\lambda d_s$ from the area of the window, and clearly all the apertures in the system are limiting.

As another remark, note that the space-frequency window of the thin lens imaging system is consistent with the well-known imaging formula in terms of its PSF. To

clarify this, let us remember that the relation between the object distribution $s(x, y)$ on the source (input) plane and the formed image $i(x, y)$ is given, within the Fresnel approximation, by

$$i(x, y) = -\frac{d_s}{d_i} \exp(i\pi \frac{x^2 + y^2}{\lambda d_i}) \left( \tilde{s}(x, y) * g_c(x, y) \right) \tag{210}$$

where

$$\tilde{s}(x, y) = s(-\frac{d_s}{d_i}x, -\frac{d_s}{d_i}y) \exp(i\pi d_s \frac{x^2 + y^2}{\lambda d_i^2}), \tag{211}$$

$$g_c(x, y) = A(\frac{x}{\lambda d_i}, \frac{y}{\lambda d_i}) \tag{212}$$

under the imaging condition:

$$\frac{1}{f} = \frac{1}{d_i} + \frac{1}{d_s}. \tag{213}$$

Here $g_c(x, y)$ denotes the coherent optical PSF of the system, and $A(x, y)$ represents the Fourier transform of the aperture function of the thin lens.

To find the input system window directly from the PSF formula (Eqn. 210), we visualize the effect of the thin lens imaging system on the input $s(x, y)$ in the phase-space plane. The combination of these effects will provide the overall input space-frequency window of the system. The idea is to start with the effects of the system on $\tilde{s}(x, y)$ and then obtain the effect on the original input $s(x, y)$ by using the relation in Eqn. 211 as follows:

$$s(x, y) = \tilde{s}(-\frac{d_i}{d_s}x, -\frac{d_i}{d_s}y) \exp(-i\pi \frac{x^2 + y^2}{\lambda d_s}). \tag{214}$$

In the given imaging formula, $\tilde{s}(x, y)$ is convolved with the PSF $g_c(x, y)$. The effect of the convolution of $\tilde{s}(x, y)$ with $g_c(x, y)$ can be visualized as a horizontal corridor in the phase-space plane as shown in Fig. 2.4a because this is equivalent to low-pass filtering with a cut-off frequency of $\frac{D}{2\lambda d_i}$. The effect of this convolution on $\tilde{s}(-\frac{d_i}{d_s}x, -\frac{d_i}{d_s}y)$ can then be obtained by scaling the corridors in Fig. 2.4a with $d_i/d_s$. These scaled corridors are shown in Fig. 2.4b. As seen from Eqn. 214, the last operation we need to consider is chirp multiplication $\mathbf{Q}_q$ with $q = 1/\lambda d_s$. The chirp multiplication leads to shearing of the space-frequency window parallel to the frequency axis in the phase-space plane [25]. Hence the combined effects of convolution, scaling and chirp multiplication can be seen in Fig. 2.4c. Finally by including the finite support of the object (between $-D_o/2$ and $D_o/2$), the overall effect of the

16

system on $s(x, y)$ can be obtained as the space-frequency window given in Fig. 2.3. This illustrates the consistency of our approach with the imaging relation of the system obtained through diffraction computations (under Fresnel approximation). Hence our approach can compensate the information provided by the imaging formula for a general imaging system when such a formula is not available or difficult to obtain.

(a) The effect of the convolution with $g_c(x, y)$ on $\tilde{s}(x, y)$ in the space-frequency plane.
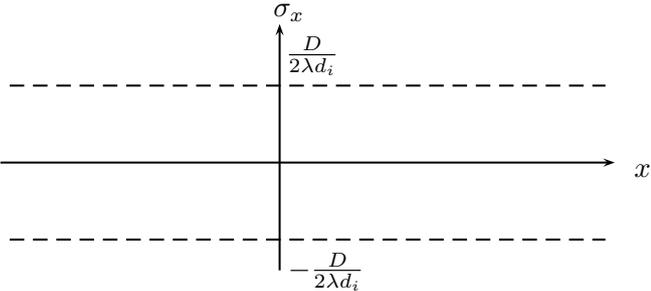


(b) The effects of the convolution with $g_c(x, y)$ and the scaling with $-d_i/d_s$ on $\tilde{s}(-\frac{d_i}{d_s}x, -\frac{d_i}{d_s}y)$ in the space-frequency plane.
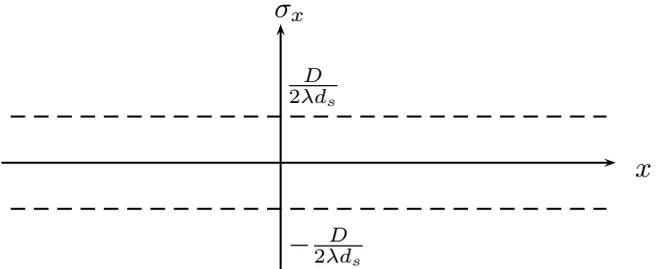


(c) The effects of the convolution, the scaling, and the chirp multiplication $\mathbf{Q}_q$ with $q = 1/\lambda d_s$ on $\tilde{s}(-\frac{d_i}{d_s}x, -\frac{d_i}{d_s}y)\exp(-i\pi\frac{x^2+y^2}{\lambda d_s})$ in the space-frequency plane.
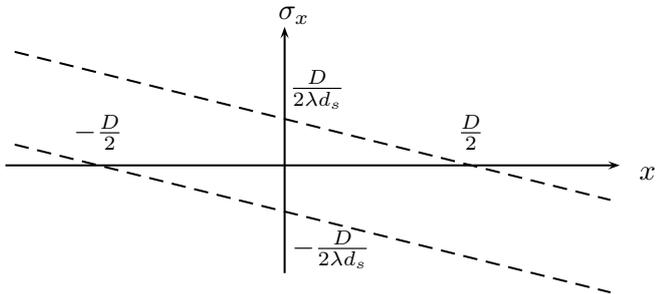


Figure 2.4: The effects of the operations in the PSF formula in the space-frequency plane for a thin lens system

### 2.3.2  Analysis of a Microscope Objective

We now use the space-frequency window approach to also analyze a microscope objective design [31] with $10\times$ magnification and $NA$ of $0.25$ at the wavelength of $560$ nm. The $EFL$ of this system is 15.0028 mm. The microscope objective design layout is given in Fig. 2.5. The design consists of two widely spaced consolidated doublets and is known as Lister-type design [32]. The parameters of each optical element in the design are given in Table 2.2, which lists for each surface, respectively, position, the radius of curvature, the thickness and refractive index of the following medium, and its aperture diameter.
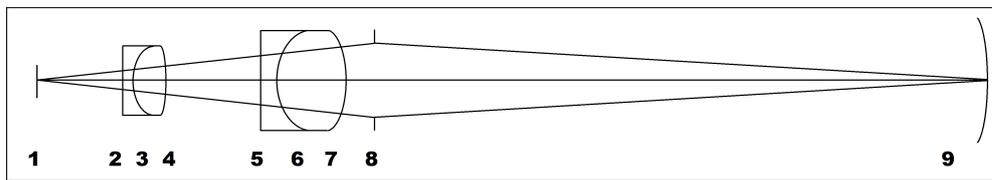


Figure 2.5: The layout of the microscope objective design

Table 2.2: Microscope objective design parameters

| Surface | Position (mm) | Radius (mm) | Thickness (mm) | Refractive index | Diameter (mm) |
|---------|---------------|-------------|----------------|------------------|---------------|
| 1 | 0.0000 | Inf | 7.6600 | 1.0000 | 1.6000 |
| 2 | 7.6600 | 21.2500 | 0.9700 | 1.6226 | 4.2000 |
| 3 | 8.6300 | 6.0000 | 3.3500 | 1.5238 | 8.0000 |
| 4 | 11.9800 | -11.4800 | 8.6400 | 1.0000 | 8.0000 |
| 5 | 20.6200 | 112.2400 | 1.1300 | 1.6226 | 10.0000 |
| 6 | 21.7500 | 9.3800 | 3.5000 | 1.5238 | 10.0000 |
| 7 | 25.2500 | -13.0900 | 0.3000 | 1.0000 | 10.0000 |
| 8 | 25.5500 | Inf | 155.0000 | 1.0000 | 7.6600 |
| 9 | 180.5500 | -18.0000 | – | – | 16.0000 |

To compute the space-frequency window of this microscope objective, the LCT system parameter matrices for each aperture location $z_j$ are obtained using the matrix $\mathbf{Q}_q$ with $q = (n_2 - n_1)/\lambda R$ for refraction from a spherical surface and the matrix $\mathbf{R}_r$ with $r = \lambda d/n$ for propagation through a refractive medium. The resulting space-frequency window is shown in Fig. 2.6. This window defines the set of all input

18

signals that can pass through the system without any information loss.
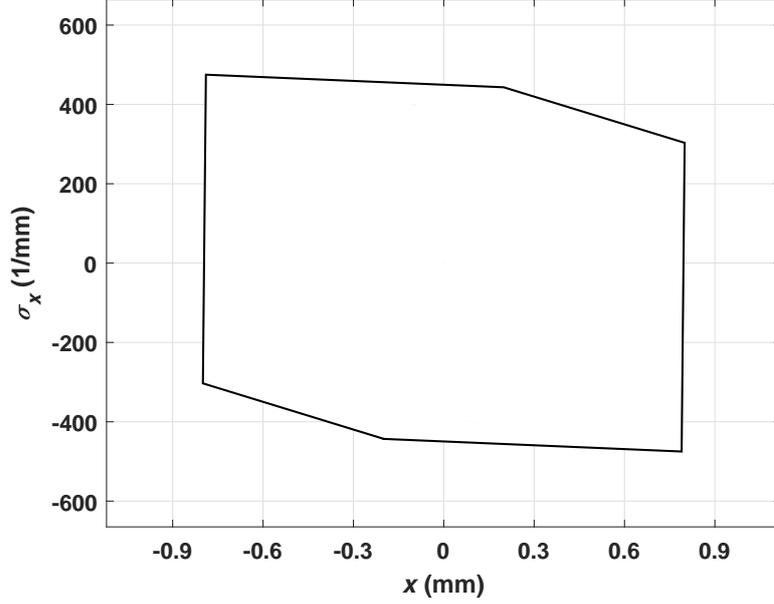


Figure 2.6: The input space-frequency window for the microscope objective design

The cut-off frequency $f_c$ is obtained from the space-frequency window as $449.58$ cyc/mm for the coherent imaging case. As discussed before, this can be used to determine the diffraction-limited resolution and NA. Using Eqn. 26, the $NA$ parameter is found as $0.25$, which is the same as the specified $NA$ value of the design.

As part of the space-frequency window computation, the overall LCT parameter matrix, $\mathbf{T}_{obj}$, of the objective design is obtained as follows:

$$\mathbf{T}_{obj} = \begin{bmatrix} -10.0534 & 1.2955 \times 10^{-10} \\ -1.1902 \times 10^{8} & -0.0979 \end{bmatrix}. \tag{215}$$

By using Eqn. 27 and 215, the $EFL$ of the objective is computed as $15.0038$, which coincides with the specified $EFL$. The overall magnification of the design is also obtained as $10.0534$ using Eqn. 28 and 215. The magnification as a function of $z$ throughout the system is also plotted in Fig. 2.7, which also shows the final magnification of $10\times$.

Using the obtained space-frequency window, one can also compute the number of degrees of freedom of the system as $1,358$ (from the area of the window). This

corresponds to the minimum number of samples required to faithfully represent the output of this imaging component. Hence, if a detector is used to record the output image, the pixel size of the detector can be chosen based on this number. Clearly, if the design is modified to increase the number of degrees of freedom of the system, the resolution will also get improved.
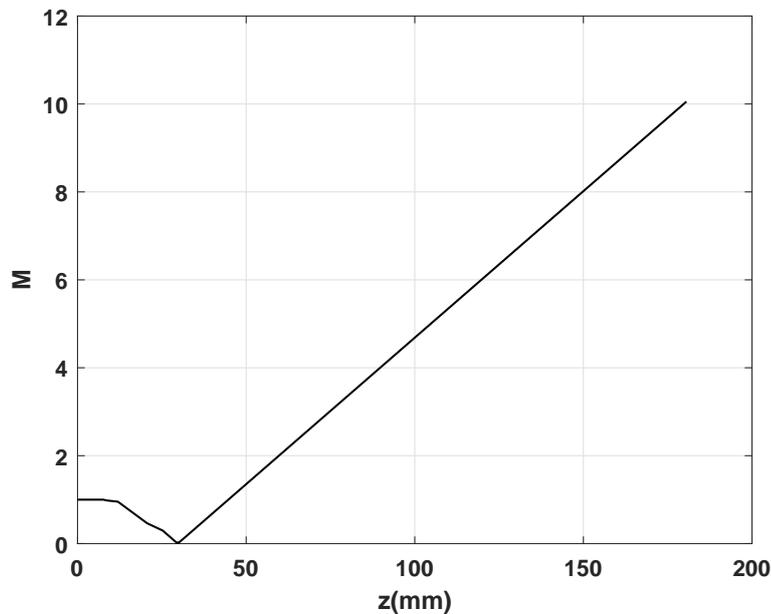


Figure 2.7: The evolution of the magnification $M$ as a function of $z$ for the microscope objective design

Phase-space window computation also reveals that only the apertures of the first, second and eighth surfaces are limiting as the system window is determined only by these three apertures. The other apertures are non-limiting, which means that decreasing their sizes to a certain point will not affect the diffraction-limited resolution of the system.

We also analyze a modified version of this microscope objective design [31] with $10\times$ magnification and $NA$ of $0.245$ at the wavelength of $560$ nm. The $EFL$ of this system is 16.2137 mm. The microscope objective design layout is given in Fig. 2.8. The design consists of two widely spaced consolidated doublets and is known as Lister-type design [32]. The parameters of each optical element in the design are given in Table 2.3, which lists for each surface, respectively, position, the radius of curvature,

20

the thickness and refractive index of the following medium, and its aperture diameter.
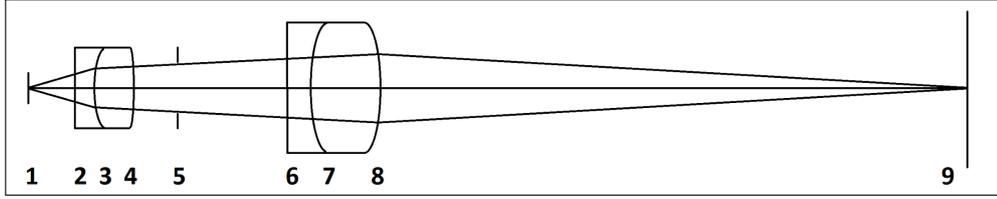


Figure 2.8: The layout of the modified microscope objective design

Table 2.3: Modified microscope objective design parameters

| Surface | Position (mm) | Radius (mm) | Thickness (mm) | Refractive index | Diameter (mm) |
|---------|---------------|-------------|----------------|------------------|---------------|
| 1 | 0.0000 | Inf | 8.2700 | 1.0000 | 1.8000 |
| 2 | 8.2700 | 22.9400 | 1.0000 | 1.6226 | 4.4000 |
| 3 | 9.2700 | 5.9700 | 3.6000 | 1.5238 | 8.0000 |
| 4 | 12.8700 | -11.7100 | 4.6600 | 1.0000 | 8.0000 |
| 5 | 17.5300 | Inf | 4.6600 | 1.0000 | 6.4891 |
| 6 | 22.1900 | 558.0000 | 1.3000 | 1.6226 | 10.0000 |
| 7 | 23.4900 | 10.5100 | 3.8000 | 1.5238 | 10.0000 |
| 8 | 27.2900 | -13.3000 | 167.8000 | 1.0000 | 8.4000 |
| 9 | 195.0900 | -20.0000 | – | – | 18.0000 |

To compute the space-frequency window of this microscope objective, the LCT system parameter matrices for each aperture location $z_j$ are obtained using the matrix $\mathbf{Q}_q$ with $q = (n_2 - n_1)/\lambda R$ for refraction from a spherical surface and the matrix $\mathbf{R}_r$ with $r = \lambda d/n$ for propagation through a refractive medium. The resulting space-frequency window is shown in Fig. 2.9. This window defines the set of all input signals that can pass through the system without any information loss.

The cut-off frequency $f_c$ is obtained from the space-frequency window as $447.95$ cyc/mm for the coherent imaging case. As discussed before, this can be used to determine the diffraction-limited resolution and NA. Using Eqn. 26, the $NA$ parameter is found as $0.25$, which is almost the same as the specified $NA$ value of the design.

As part of the space-frequency window computation, the overall LCT parameter ma-
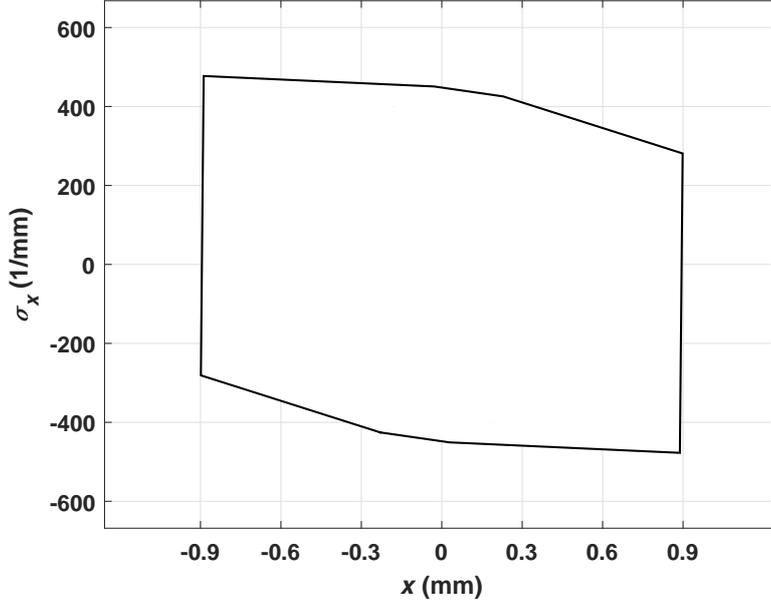
Figure 2.9: The input space-frequency window for the modified microscope objective design

trix, $\mathbf{T}_{obj}$, of the objective design is obtained as follows:

$$\mathbf{T}_{obj} = \begin{bmatrix} -10.0605 & 1.3765 \times 10^{-10} \\ -1.1013 \times 10^8 & -0.0979 \end{bmatrix}. \tag{216}$$

By using Eqn. 27 and 216, the $EFL$ of the objective is computed as $16.2147$, which coincides with the specified $EFL$. The overall magnification of the design is also obtained as $10.0605$ using Eqn. 28 and 216. The magnification as a function of $z$ throughout the system is also plotted in Fig. 2.10, which also shows the final magnification of $10\times$.

Using the obtained space-frequency window, one can also compute the number of degrees of freedom of the system as $1,499$ (from the area of the window). This corresponds to the minimum number of samples required to faithfully represent the output of this imaging component. Hence, if a detector is used to record the output image, the pixel size of the detector can be chosen based on this number. Clearly, if the design is modified to increase the number of degrees of freedom of the system, the resolution will also get improved.

Phase-space window computation also reveals that only the apertures of the first,

22
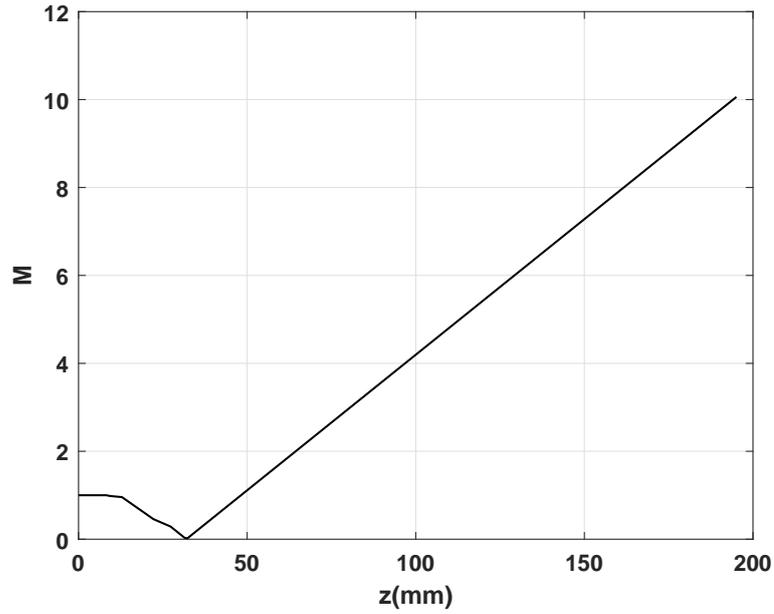
Figure 2.10: The evolution of the magnification $M$ as a function of $z$ for the modified microscope objective design

second, fifth and eighth surfaces are limiting as the system window is determined only by these four apertures. The other apertures are non-limiting, which means that decreasing their sizes to a certain point will not affect the diffraction-limited resolution of the system.

## 2.4 Conclusions

In conclusion, we have demonstrated how to analyze coherent imaging systems with multiple diffracting apertures using the space-frequency window. The approach is validated for a single thin lens system and a practical microscope objective design. The developed approach will be useful for the design and analysis of imaging systems as it is a simple and straightforward method for quantifying the important imaging parameters.

# CHAPTER 3

# DEEP LEARNING-BASED ITERATIVE RECONSTRUCTION FOR PHASE RETRIEVAL

## 3.1 Introduction

The classical phase retrieval problem is one of the important inverse problems in coherent imaging systems. It is the recovery of a constrained signal from the magnitude of its Fourier transform, or equivalently from its autocorrelation. Practically, measuring the phase of a light wave (oscillating at visible spectrum frequencies or higher) is not straightforward. It requires additional processes such as interfering wave with another known field [7]. However, the phase is also obtained by using numerical methods. Under the coherent illumination, the diffraction intensity at the far field corresponds to the Fourier magnitude of the object. This is also called coherent diffraction imaging (CDI). In CDI, the phase of the light can be recovered by the solution of phase retrieval problem when prior information and Fourier magnitude measurements are given.

This problem is encountered in a variety of applications in science and engineering such as crystallography [8], microscopy [9, 10], astronomy [11], optical imaging [12, 13], and speech processing [14]. In this chapter, we consider two types of measurements. There can be either one Fourier magnitude measurement or multiple coded diffraction pattern (CDP) measurements. In CDP measurements, the object is multiplied with different masks or phase plates to obtain multiple diffraction measurements. This approach gives some redundancy, which helps uniqueness issues, and additional information about the object [35]. Although a unique solution almost

---

Some parts of this chapter have been presented in [33], and also published [34].

always exists for most of the practical scenarios [36], solving the phase retrieval problem is generally a difficult task because of the inherent ill-posedness and nonlinearity involved.

Although there are several approaches developed for phase retrieval, each suffers from different limitations. Alternating projection-based methods, including the hybrid input-output (HIO) algorithm, are the most commonly used methods because of their low computational complexity and image generality [7, 15]. This class of methods alternates between the space and frequency domains by imposing the available information in each domain through projections [7, 37, 38, 15]. However, some of these projections involve non-convex sets, and hence convergence to the global solution cannot be guaranteed. The resulting reconstructions may have artifacts and errors mostly due to being stuck in local minima or amplification of noise in the solution. More recent phase retrieval algorithms have been developed to overcome some of these limitations. Examples include semi-definite programming-based approaches [39, 40, 41], regularization-based methods [42, 2, 43], and Wirtinger flow and its variants [44].

Recently, DNNs [45] have been successfully applied for the solution of several inverse problems in imaging [46]. There are two main approaches in exploiting DNNs for the solution of inverse problems. In the first class of approaches, a DNN is used to reconstruct the unknown image directly from an available measurement or from an initial estimate obtained with a simple model-based inversion approach. Hence, these approaches exploit DNNs either to perform direct inversion or to improve a rough estimate that may involve artifacts or errors. For this, a DNN is trained by minimizing a loss function between the ground truth images and the available measurements or estimates. This approach has been utilized to solve several inverse problems [47, 48, 49, 1], including phase retrieval as encountered in holography, lensless imaging and Fourier ptychography [50, 10, 51]. In the second class of approaches, DNNs are utilized for the regularization of model-based inversion methods by using plug-and-play regularization and its variants [16, 17]. In Chapter 4, this class of approaches is discussed in more details and such a novel approach is also developed for phase retrieval.

In this chapter, we develop a hybrid phase retrieval algorithm that utilizes DNNs with a model-based inversion approach. Here, the used model-based inversion approach is the well-known HIO method, which incorporates the physical model and the constraints into the solution, but may lead to artifacts. The main idea in the developed method is to use a DNN in *an iterative manner* with the HIO method to remove the artifacts. The developed approach consists of two main stages: the iterative DNN-HIO stage and the final DNN stage. For the iterative DNN-HIO stage, a DNN is trained to remove the HIO artifacts. This trained DNN is then used iteratively with the HIO method to generate an intermediate reconstruction. In the final stage, the intermediate reconstructions are used to train a second DNN to remove the remaining artifacts. The performance of the developed approach is compared with classical and state-of-the-art methods through numerical simulations. The results demonstrate the effectiveness of our approach, which has relatively little additional computational cost compared to HIO. Our approach not only achieves state-of-the-art reconstruction performance but also is more robust to different initialization and noise levels.

The rest of this chapter is organized as follows. The classical phase retrieval problem is described in Section 3.2. Related work on phase retrieval and DNN-based methods are discussed in Section 3.3. Section 3.4 presents the developed approach. The performance of the approach is compared with classical and state-of-the-art methods in Section 3.5 through simulations. Finally, we summarize the results and conclude in Section 3.6.

## 3.2 Phase Retrieval Problem

In the classical phase retrieval problem, available measurements can be modeled as

$$\mathbf{y^2} = |\mathbf{Fx}|^2 + \mathbf{w}, \qquad \mathbf{w} \sim N(\mathbf{0}, \alpha^2 Diag(|\mathbf{Fx}|^2)) \qquad (31)$$

where $\mathbf{y^2} \in \mathbb{R}^{M^2}$ denotes the noisy Fourier intensity measurements, $\mathbf{F}$ is the $M \times M$-point DFT matrix, and $\mathbf{x} \in \mathbb{R}^{N^2}$ represents the unknown image of interest. $Diag(|\mathbf{Fx}|^2)$ is a square diagonal matrix with the elements of vector $|\mathbf{Fx}|^2$ on the main diagonal. The unknown image $\mathbf{x}$ is assumed to be non-negative, real-valued and have finite support. Moreover, $\mathbf{w} \in \mathbb{R}^{M^2}$ denotes the measurement noise, and $\alpha$ is a scaling pa-

rameter that controls the signal-to-noise ratio (SNR). The noise is generally assumed to be Poisson-distributed, and here its normal approximation [2] is used.

For two or higher dimensional real-valued discrete signals with finite support, Fourier intensity measurements at discrete frequencies, $|\mathbf{Fx}|^{\mathbf{2}}$, can uniquely determine the unknown signal, $\mathbf{x}$. To guarantee uniqueness, for an image with support $N \times N$, the magnitude of its $M \times M$-point oversampled DFT with $M \geq 2N - 1$ should be provided [36]. In this work, $M$ is chosen as $2N$ for simplicity.

## 3.3   Related Work

### 3.3.1   Alternating Projection Methods for Phase Retrieval

Alternating projection-based methods are widely used for phase retrieval. In the classical Gerchberg-Saxton (GS) algorithm [37], magnitude constraints are iteratively imposed in space and Fourier domains to reconstruct the unknown signal. The error reduction (ER) algorithm is a modified version of the GS algorithm, which uses other space domain constraints instead of the magnitude in the space domain [38]. The ER algorithm corresponds to the steepest descent method for the optimization of $\|\mathbf{y} - |\mathbf{Fx}|\|^{2}$ [15]. The most commonly used alternating projection-based method is the HIO algorithm [15], which is developed based on the ER algorithm.

Similar to the ER algorithm, in the HIO method, Fourier magnitude constraint and space domain constraints (such as support, non-negativity, and real valuedness) are iteratively used. However, unlike ER, HIO does not force the iterates to satisfy the constraints exactly, but it uses the iterates to eventually drive the algorithm to a solution that satisfy the constraints [15]. The HIO iterations can be expressed as follows:

$$\mathbf{x}_{k+1}[n] = \begin{cases} \mathbf{x}'_k[n] & \text{for} \quad n \notin \gamma \\ \mathbf{x}_k[n] - \beta \mathbf{x}'_k[n] & \text{for} \quad n \in \gamma \end{cases} \tag{32}$$

where

$$\mathbf{x}'_k = \mathbf{F}^{-1}\left\{\mathbf{y} \odot \frac{\mathbf{Fx}_k}{|\mathbf{Fx}_k|}\right\} \tag{33}$$

Here, $\mathbf{x}_k \in \mathbb{R}^{N^2}$ is the reconstruction at the $k^{th}$ iteration, $\mathbf{F}^{-1}$ denotes the inverse

DFT matrix, $\odot$ represents the element-wise (Hadamard) multiplication operation, $\beta$ is a constant parameter (with a typical value of $0.9$) and $\gamma$ is the set of indices $n$ for which $\mathbf{x}'_k[n]$ violates the space domain constraints [15]. Although the convergence behavior of the HIO method cannot be completely analyzed, it often converges to a reasonably good solution empirically in a wide variety of applications. However, the HIO reconstructions may have artifacts and errors mostly due to being trapped in local minima or amplification of noise in the solution. Variants of the HIO method have also been proposed to improve its performance [7].

### 3.3.2 Deep Neural Network (DNN)-based Methods for Inverse Problems

In the last decade, DNNs have been successfully used for the solution of various inverse problems including denoising, deconvolution, and superresolution [46]. There are two main approaches in utilizing DNNs for solving inverse problems.

In the first class of approaches, a DNN is used to reconstruct the unknown image directly from an available measurement or from an initial estimate obtained with a simple model-based inversion. That is, these approaches exploit DNNs either to solve end-to-end inverse problems or to improve a rough estimate that may have artifacts or errors. For this purpose, a DNN is trained by minimizing a loss function using a dataset containing the ground truth images and the measurements (or the initial estimates). In general, this approach provides a faster reconstruction than a model-based inversion approach since it works in a *non-iterative* feed-forward fashion to solve the problem. However, a DNN usually needs specialized training and dataset for each inverse problem, which reduces its flexibility to handle different inverse problems. More importantly, this approach works successfully only when the measurements or the initial estimates used for reconstruction are similar in appearance to the ground truth images. This approach has been used to solve several inverse problems in imaging applications such as denoising [47], deconvolution [48, 52], superresolution [49, 53], tomography [1], holographic image reconstruction [10], phase retrieval for phase objects [50], and Fourier ptychography [51].

In the second class of approaches, DNNs are utilized for the regularization of model-based inversion methods by using plug-and-play regularization [16, 17]. This ap-

proach is discussed with more details in Chapter 4.

## 3.4 DNN-based Iterative Phase Retrieval Approach

Our deep learning-based hybrid algorithm utilizes DNNs together with the HIO method. The main idea in our approach is to use the HIO method to directly incorporate the physical model and the constraints into the reconstruction, and DNNs to improve the resulting HIO reconstructions. The first DNN, namely DNN-1, is trained to remove the artifacts of the initial HIO reconstructions, and is used iteratively with the HIO method to generate an intermediate reconstruction. Then, a second DNN, namely DNN-2, is trained to remove the remaining artifacts after this iterative stage. The output of DNN-2 is the final reconstruction of our method. The overall approach is illustrated in Fig. 3.1 using representative images for the input and output of each step. A preliminary version of this iterative approach was presented in [33].
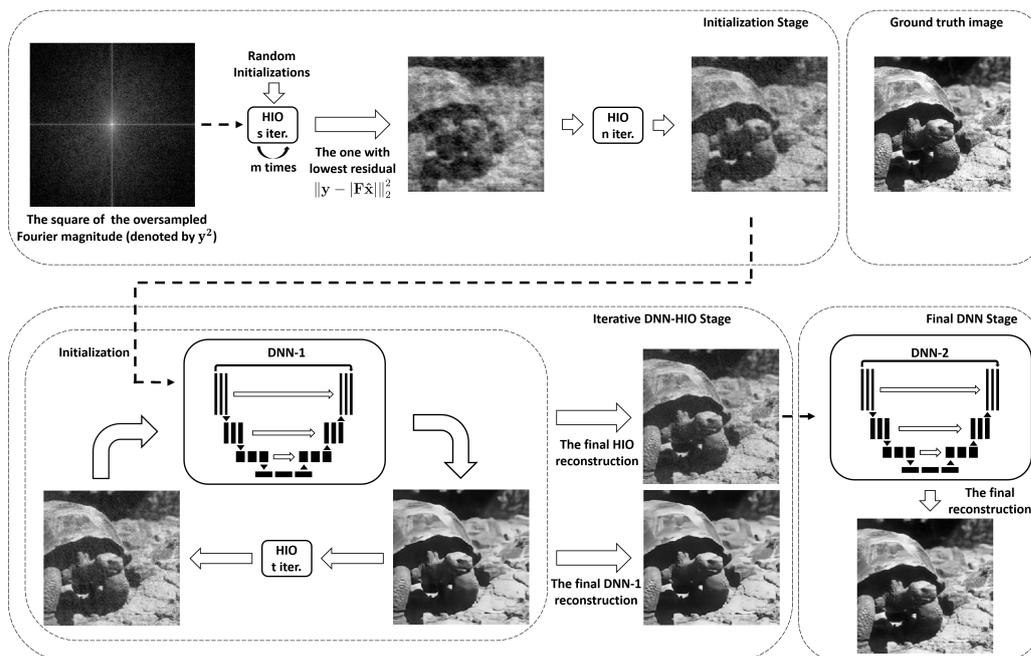


Figure 3.1: The developed method with initialization, iterative DNN-HIO and final DNN stages.

As shown in Fig. 3.1, the approach consists of three stages: the initialization stage, the iterative DNN-HIO stage, and the final DNN stage. The initialization stage helps

to achieve robustness to initialization. For this aim, the HIO reconstructions with different random initialization are obtained and the one that provides Fourier magnitude closest to the given measurement is chosen as the input (initialization) for the iterative stage. In the iterative DNN-HIO stage, a DNN and the HIO method are used iteratively to generate an intermediate reconstruction. This DNN is trained using the HIO reconstruction at the output of the initialization stage with the ground truth images. Hence this training aims to remove the HIO artifacts at the output of the initialization stage, but this can be performed to some extent. After this iterative stage, the intermediate reconstructions have less artifacts than the initial HIO reconstructions. In the final DNN stage, the intermediate reconstructions are used with the ground truth images to train a second DNN to remove the remaining artifacts.

As DNN architectures, the modified U-net architecture developed in [1] is used. This architecture, which is shown in Fig. 3.2, works in a *non-iterative* feed-forward fashion to solve general inverse problems in imaging. In particular, in [1], this is used to obtain reconstructions for computed tomography. Here we use the same architecture in an *iterative* manner with the HIO method to solve the phase retrieval problem.
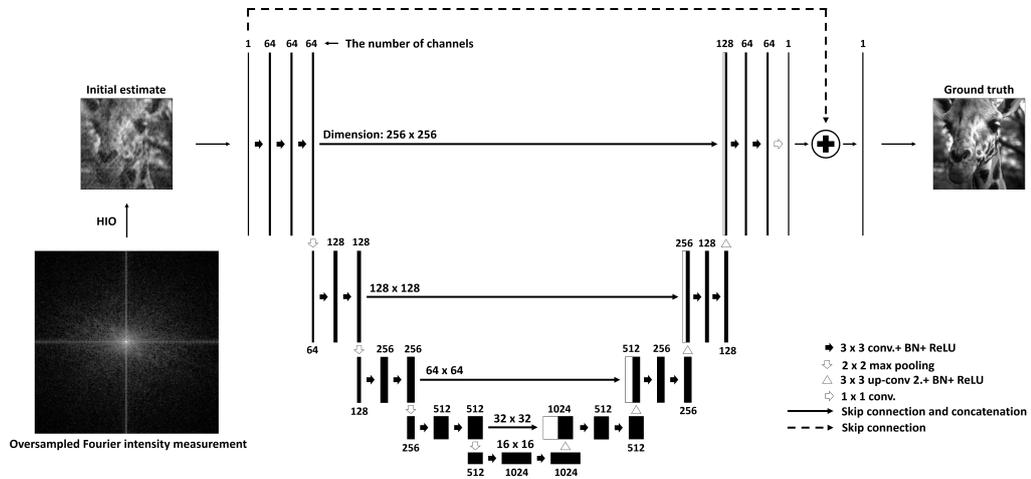


Figure 3.2: The U-net deep neural network architecture (figure adapted from [1]), BN: batch normalization, conv.: convolutional filter, up-conv 2.: transposed convolutional layer with a stride of 2, ReLU: rectifier linear unit.

This architecture is the modified version of the original U-net architecture [54]. The original U-net is developed for biomedical image segmentation and it exploits en-

coding and decoding convolutional layers with skip connections between symmetric downsampling and upsampling convolutional layers [46]. These features were shown to be useful for solving many inverse problems including denoising [55], image inpainting [56], optical flow estimation [57] and computed tomography [1]. In addition to these features, the modified U-net architecture contains batch normalization layers and direct skip connection between the input and output. These modifications help the DNN to better learn the residual between the input and output images [47].

In what follows, we provide the details of each stage in our approach.

### 3.4.1 Initialization Stage

Due to the nonlinearity (and non-convexity) involved in the phase retrieval problem, the reconstruction algorithms are generally sensitive to initialization. Here, to increase the robustness of our approach, a particular initialization procedure described earlier in [2] is used. In this procedure, first, the HIO method is run with $m$ different random initialization for a small number of $k$ iterations. Then, the reconstruction $\hat{x}$ with the lowest residual $\|y - |F\hat{x}|\|_2^2$ is used for another HIO run for a larger number of $n$ iterations. The final reconstruction is used as the input (initialization) for the iterative DNN-HIO stage.

### 3.4.2 Iterative DNN-HIO Stage

As mentioned before, although the HIO method benefits from the physical model and the constraints during the reconstruction process, the results may have artifacts and errors caused mostly by the presence of noise or being stuck in local minima. In this stage, a DNN (namely DNN-1) and the HIO method are used alternately to solve the phase retrieval problem.

DNN-1 is trained to remove the artifacts of the HIO method after the initialization stage. That is, DNN-1 is trained by using a dataset containing true images and their corresponding HIO reconstructions at the output of the initialization stage. Then, the HIO method and the trained DNN are used in an iterative manner until the re-

constructions start to change slightly. This iterative approach aims to improve the reconstructions by escaping from local minima and reducing artifacts.

More specifically, at the $k^{th}$ iteration of this stage, the last HIO reconstruction, $\mathbf{x}_k$, is used as the input for DNN-1. Then, the improved reconstruction, $\mathbf{u}_k$, at the output of DNN-1 is used as the initialization for the HIO method, which is run for a small number of $t$ iterations. This iterative procedure continues until the normalized error between two consecutive DNN-1 reconstructions, i.e. $\|\mathbf{u}_k - \mathbf{u}_{k-1}\|_2 \, / \, \|\mathbf{u}_k\|_2$, is smaller than $10^{-3}$.

As the iterations proceed, both the reconstructions of DNN-1 and HIO are improved. In particular, the HIO method better preserves the high spatial frequencies of the original image, which represent sudden spatial changes in the image, compared to DNN-1, while DNN-1 provides reconstructions with less artifacts. This has two main reasons. First, DNNs generally smooth out the high frequencies during its learning process when they are trained with a mean squared error (MSE) based loss, which is a common problem in DNNs [46]. Moreover, the main task of DNN-1 here is to remove the large artifacts, which inherently comes with the side effect of smoothing (i.e. low-pass filtering). Secondly, unlike DNN-1, the HIO method uses the available measurements together with the forward model, which helps to preserve high frequencies, although it comes with artifacts. The final HIO reconstruction is used as the input for the last stage in order to preserve high frequencies in the final reconstruction.

### 3.4.3 Final DNN Stage

In this last stage, a second DNN (namely DNN-2) is used to improve the reconstruction of the iterative DNN-HIO stage by removing the remaining artifacts. The reason for using a different DNN here is that DNN-1 is trained to remove the HIO artifacts at the output of the initialization stage, but the reconstructions of the iterative DNN-HIO stage has less artifacts than before. Therefore, training another DNN enables to obtain improved reconstructions with better preserved high frequencies and reduced artifacts.

33

DNN-2 is trained to remove the artifacts of the iterative DNN-HIO stage. That is, DNN-2 is trained by using a dataset containing the same ground truth images and the corresponding HIO reconstructions at the output of the iterative DNN-HIO stage. As mentioned before, MSE-based loss function is used for training, but different loss functions could also be utilized to better preserve high frequencies. This trained DNN is used in a non-iterative feed-forward fashion to obtain the final reconstruction of our method.

## 3.5   Numerical Results

Here, we present numerical simulations to illustrate the effectiveness of our approach. In particular, the reconstruction performance of the method is compared with classical and state-of-the-art phase retrieval methods.

To compare the algorithms in terms of noise tolerance, image generality, and computational efficiency, the reconstruction performance of the developed method is investigated using two different kinds of images, which are called natural and unnatural images. The unnatural images have distinct statistics from natural images. Test dataset consists of $236$ images containing $230$ natural and $6$ unnatural images. These include $200$ test images of BSD, $24$ Kodak dataset images [58], $6$ natural and $6$ unnatural images taken from [2]. The unnatural image dataset consists of images acquired by scanning electron microscopes and telescopes. Sample natural and unnatural images from the test dataset are shown in Fig. 3.3 and 3.4 respectively. The pixel values of all images are between $0$ and $255$, and all are of size $256 \times 256$.

For training DNN-1 and DNN-2, only natural images are used. This training dataset consists of $3000$ natural images. These include $200$ training and $100$ validation images of Berkeley segmentation dataset (BSD) [59], $400$ selected images from validation set of ImageNet database [60, 61], and randomly chosen $2300$ images of Waterloo Exploration Database [62].

The noisy Fourier measurements were simulated using Eqn. 31 with $\alpha = 3$, resulting in an average SNR of $31.84$ dB (where SNR $= 10 \log(\|\,|\mathbf{Fx}|^2\|_2 \,/\, \|\mathbf{y^2} - |\mathbf{Fx}|^2\|_2))$. These measurements were used to obtain the initial HIO reconstructions at the output

34

(a) Penguin  (b) Wall  (c) Men

(d) Portland Head
Light  (e) House  (f) Turtle
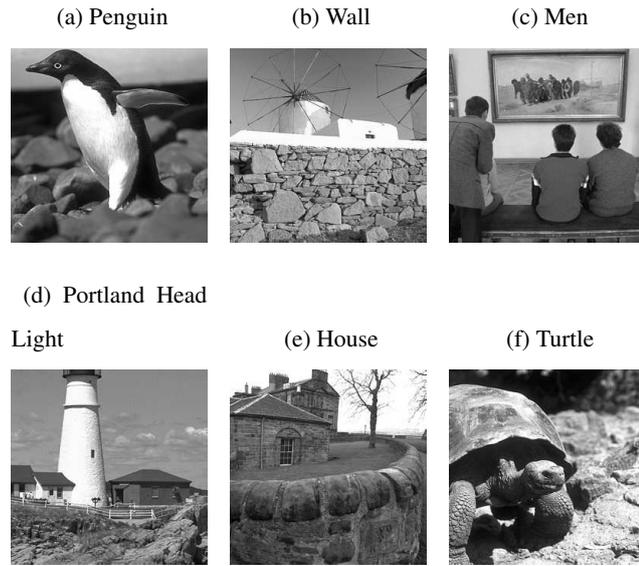
Figure 3.3: Sample natural images of size 256 x 256 from the test dataset.

(a) Butterfly Nebula  (b) E. Coli  (c) Pillars of Creation

(d) Pollen  (e) Tadpole Galaxy  (f) Yeast
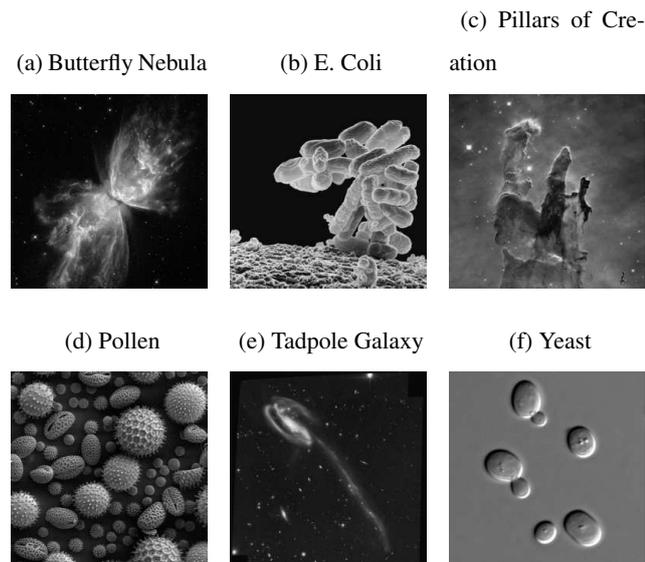
Figure 3.4: Sample unnatural images of size 256 x 256 from the test dataset. [2].

of the initialization stage. DNN-1 was trained using these reconstructions and the true images. Likewise, DNN-2 was trained using the true images and the HIO reconstructions of the iterative DNN-HIO stage. Although only natural images were used in training, the developed approach with the trained DNNs was tested using both natural and unnatural images.

Training was performed by minimizing the MSE-based loss between the true images and the reconstructions at the output of each DNN. Stochastic gradient descent algorithm with momentum was used for the optimization [63]. All computations were done using MATLAB with MatConvNet toolbox [64] and NVIDIA Geforce GTX TI-TAN X GPU. The total training times for DNN-1 and DNN-2 were about $38$ hours (for $251$ iterations) and $51$ hours (for $201$ iterations), respectively.

In the initialization stage, the HIO method was first run with $m = 50$ different random initialization for $k = 50$ iterations. Then, the reconstruction with the lowest residual was used for another HIO run for $n = 1000$ iterations. The resulting reconstruction was input to the iterative DNN-HIO stage as shown in Fig. 3.1. In this stage, each time the HIO method was run for $t = 5$ iterations.

After the testing phase, the reconstructions of the developed approach were compared with the true images using peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [65]. For comparison, the reconstructions of the HIO method and prDeep [2], one of the state-of-the-art deep learning-based phase retrieval algorithms, were also obtained. Both the developed algorithm and prDeep were initialized with the output of the initialization stage. The HIO reconstruction used for comparison was the output of this initialization stage.

In Table 3.1, the average reconstruction performance of the algorithms for $236$ test images and $5$ Monte Carlo runs are given for different amount of Poisson noise ($\alpha = 2, 3, 4$). As seen in the table, for all cases, the developed method outperforms the HIO and prDeep methods in terms of both PSNR and SSIM, while requiring little additional runtime compared to HIO. As another benchmark, the results at the output of DNN-1 and iterative DNN-HIO stages are also provided in the table to show performance gains obtained by the iterative approach. The results illustrate that, by utilizing a DNN in an iterative manner with the HIO method, many of the HIO artifacts can be successfully removed while preserving the image characteristics. This iterative approach with the additional DNN (DNN-2) is the overall method, which provides the best reconstruction performance.

Sample reconstructions for a natural image in the test dataset are shown in Fig. 3.5. As seen from the figures, the developed approach provides the best reconstruction

Table 3.1: The average reconstruction and runtime performances for 236 test images (5 Monte Carlo runs)

| $\alpha = 2$ (Avg. SNR: 33.39 dB) | Avg. PSNR (dB) | | | Avg. SSIM | | | Avg. runtime (sec.) |
|---|---|---|---|---|---|---|---|
| | Overall | Natural | Unnatural | Overall | Natural | Unnatural | |
| The HIO method | 18.97 | 18.92 | 20.78 | 0.28 | 0.29 | 0.26 | **55.40** |
| DNN-1 | 20.76 | 20.77 | 20.33 | 0.33 | 0.33 | 0.20 | 55.47 |
| Iterative DNN-HIO | 21.63 | 21.60 | 22.75 | 0.47 | 0.47 | 0.26 | 59.07 |
| PrDeep | 23.45 | 23.49 | 21.72 | 0.51 | 0.51 | 0.24 | 169.81 |
| Developed method | **23.61** | **23.60** | **24.02** | **0.53** | **0.53** | **0.31** | 59.14 |
| $\alpha = 3$ (Avg. SNR: 31.66 dB) | Avg. PSNR (dB) | | | Avg. SSIM | | | Avg. runtime (sec.) |
| | Overall | Natural | Unnatural | Overall | Natural | Unnatural | |
| The HIO method | 18.07 | 18.02 | 19.97 | 0.21 | 0.21 | 0.14 | **55.61** |
| DNN-1 | 19.69 | 19.68 | 20.06 | 0.26 | 0.26 | 0.18 | 55.69 |
| Iterative DNN-HIO | 21.07 | 21.03 | 22.82 | 0.41 | 0.42 | 0.25 | 60.29 |
| PrDeep | 22.06 | 22.09 | 20.91 | 0.44 | 0.44 | 0.22 | 171.02 |
| Developed method | **22.87** | **22.85** | **23.50** | **0.47** | **0.48** | **0.29** | 60.35 |
| $\alpha = 4$ (Avg. SNR: 30.40 dB) | Avg. PSNR (dB) | | | Avg. SSIM | | | Avg. runtime (sec.) |
| | Overall | Natural | Unnatural | Overall | Natural | Unnatural | |
| The HIO method | 17.34 | 17.30 | 18.72 | 0.16 | 0.17 | 0.10 | **55.78** |
| DNN-1 | 18.75 | 18.76 | 18.65 | 0.21 | 0.21 | 0.14 | 55.86 |
| Iterative DNN-HIO | 20.08 | 20.03 | 22.22 | 0.35 | 0.36 | 0.20 | 60.99 |
| PrDeep | 20.69 | 20.70 | 20.38 | 0.37 | 0.38 | 0.18 | 172.47 |
| Developed method | **21.80** | **21.77** | **22.79** | **0.41** | **0.41** | **0.25** | 61.05 |

visually as well as in terms of used quantitative image quality measures (PSNR and SSIM). In fact, our approach generally does not introduce artifacts and errors like the HIO and prDeep methods. As mentioned before, removing artifacts sometimes causes the side effect of smoothing.

For the same test image, Fig. 3.6 shows several intermediate reconstructions obtained with the developed approach. The reconstructions at the output of each stage, including the initialization stage, iterative DNN-HIO stage, and the final DNN stage, are shown here, together with their respective PSNR and SSIM values. These clearly illustrate the contribution of each stage. For example, the improvement obtained with the final DNN-2 stage can be understood by comparing the final reconstruction in Fig. 3.6f with the reconstructions at the output of the iterative stage as given in Figures 3.6d and 3.6e. In fact, this final reconstruction is much better than all the other reconstructions both visually and quantitatively. Moreover, to demonstrate the use-

(a) HIO reconstruction (PSNR:17.68, SSIM:0.32)

(b) PrDeep reconstruction (PSNR:25.35, SSIM:0.71)

(e) Developed method (PSNR:26.49, SSIM:0.73)

(f) Ground truth
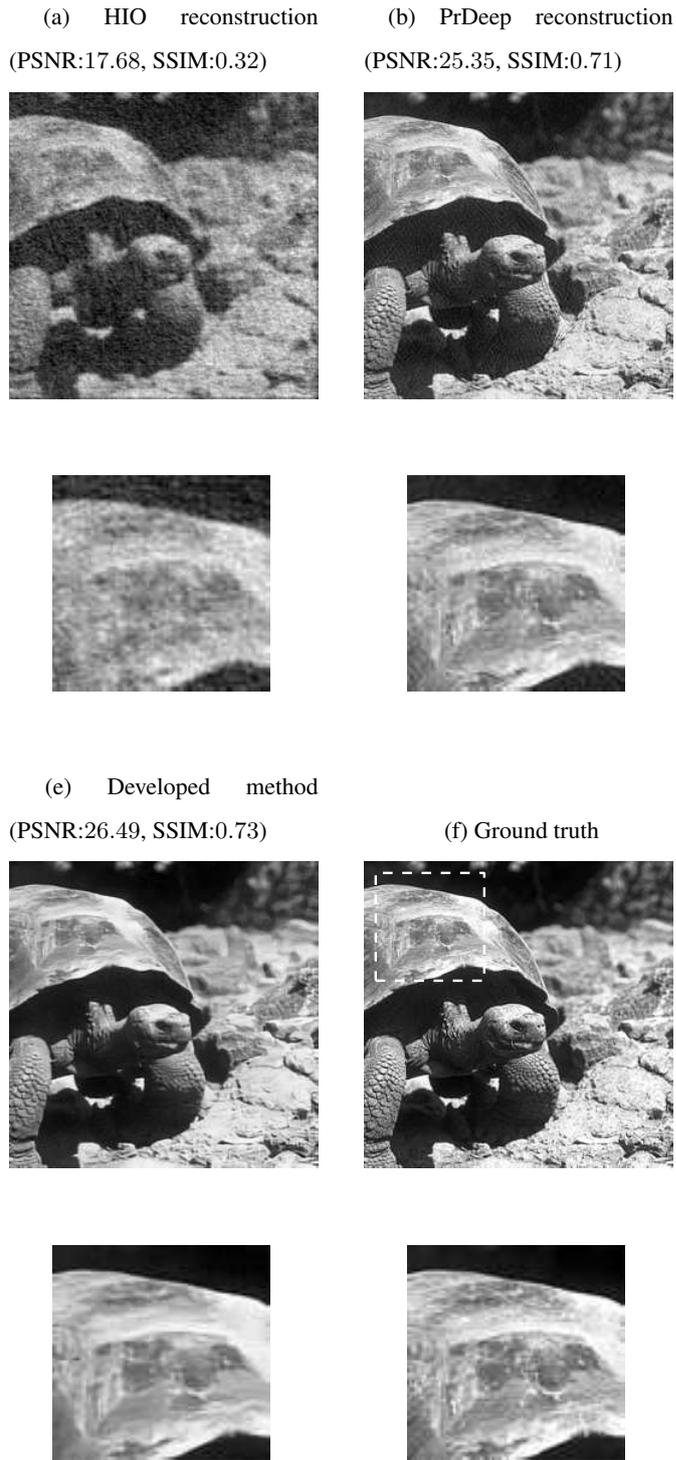
Figure 3.5: The reconstructions of the different algorithms for the "Turtle" test image for $\alpha$=3 case.

fulness of the iterative use of HIO with DNN-1, the reconstructions obtained after the first iteration are also provided in Figures 3.6b and 3.6c. Comparing these with Figures 3.6d and 3.6e illustrates that, although even a single iteration helps to improve the initial HIO reconstruction, iterations until convergence can provide much significant improvement. Note that after DNN-1 (see Fig. 3.6b), the reconstruction suffers from over-smoothing, and when this is input to HIO (see Fig. 3.6c) some high frequency information is recovered but with artifacts. As the iterations proceed, both over-smoothing and artifacts are reduced.

To assess the performance of different algorithms in terms of image generality, the results for both natural and unnatural test images are separately provided in Table 3.1. As seen in the table, although the DNNs were trained by using only natural images, the developed method shows the best reconstruction performance not only for natural images but also for unnatural images, which have distinct statistics from natural images. In particular, the performance of the prDeep method substantially degrades for unnatural images, as expected, since its reconstruction relies on a regularization prior learned from natural images. To illustrate these points, sample reconstructions for an unnatural image in the test dataset are shown in Fig. 3.7.

The developed approach also appears to be robust to different noise levels. As seen from the table, the reconstruction performance of the approach surpasses the other methods for different noise levels ($\alpha = 2, 4$) as well, even though the DNNs were trained only for a specific noise level ($\alpha = 3$).

As mentioned before, phase retrieval algorithms are generally sensitive to initialization because of the nonlinearity involved in the problem. To illustrate the robustness of the developed approach to different initialization and image characteristics, the PSNR and SSIM histograms are provided in Fig. 3.8 for each method (when $\alpha = 3$). These include reconstructions obtained with 236 distinct test images and 5 Monte Carlo runs, which means that 5 different initialization is used for each test image. As seen from the histograms, although the histogram for the prDeep reconstructions has more counts in higher PSNR and SSIM values, our method attains a higher average PSNR and SSIM, as well as a smaller spread around these averages. These results suggest that the performance of the developed approach is more robust to different

(a) Initial HIO recons. (PSNR:17.68, SSIM:0.32)

(b) DNN-1 recons. in the $1^{st}$ iteration (PSNR:21.30, SSIM:0.43)



(c) HIO recons. in the $1^{st}$ iteration (PSNR:19.47, SSIM:0.45)

(d) Final DNN-1 recons. of the iterative DNN-HIO stage (PSNR:24.34, SSIM:0.61)



(e) Final HIO recons. of the iterative DNN-HIO stage (PSNR:21.27, SSIM:0.62)

(f) Final recons. of the developed method (PSNR:26.49, SSIM:0.73)
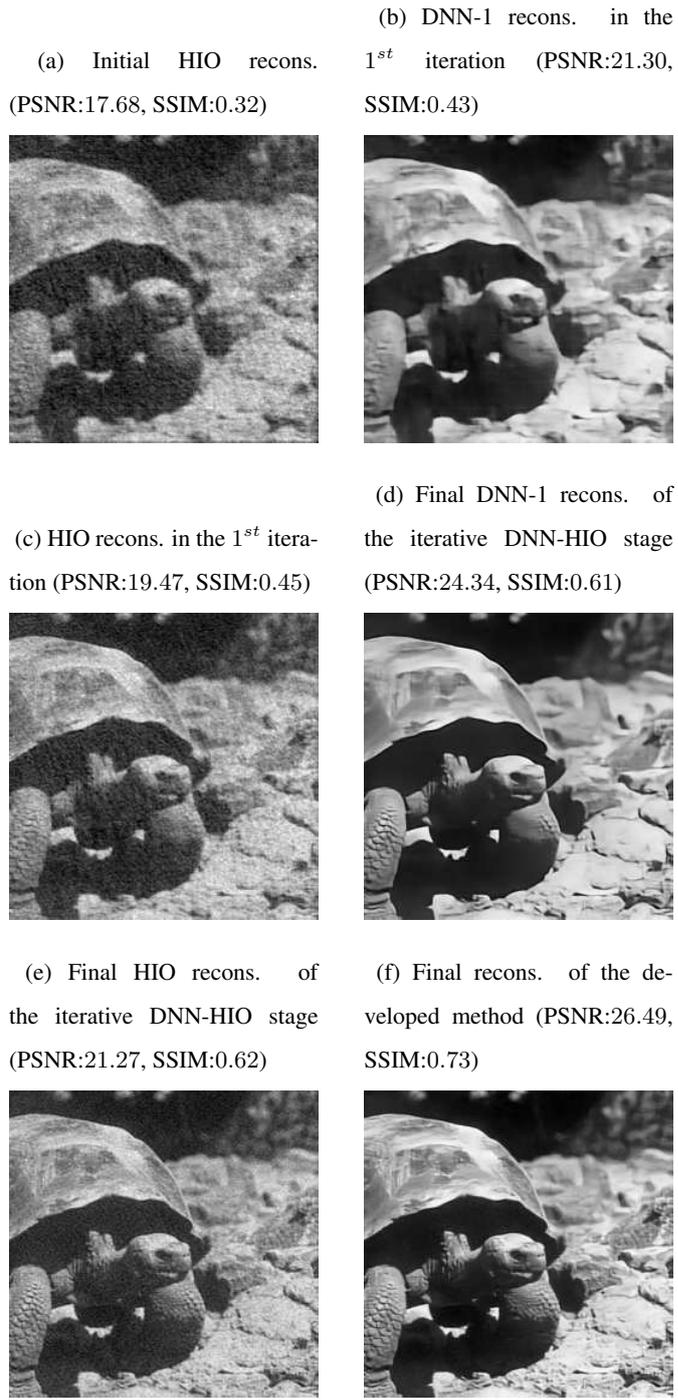


Figure 3.6: The intermediate reconstructions of the developed approach for the "Turtle" test image for $\alpha = 3$ case. These images are also used in the illustration of the method in Fig 3.1.

(a) HIO reconstruction (PSNR:21.63, SSIM:0.39)

(b) PrDeep reconstruction (PSNR:19.37, SSIM:0.41)

(e) Developed method (PSNR:25.33, SSIM:0.67)

(f) Ground truth
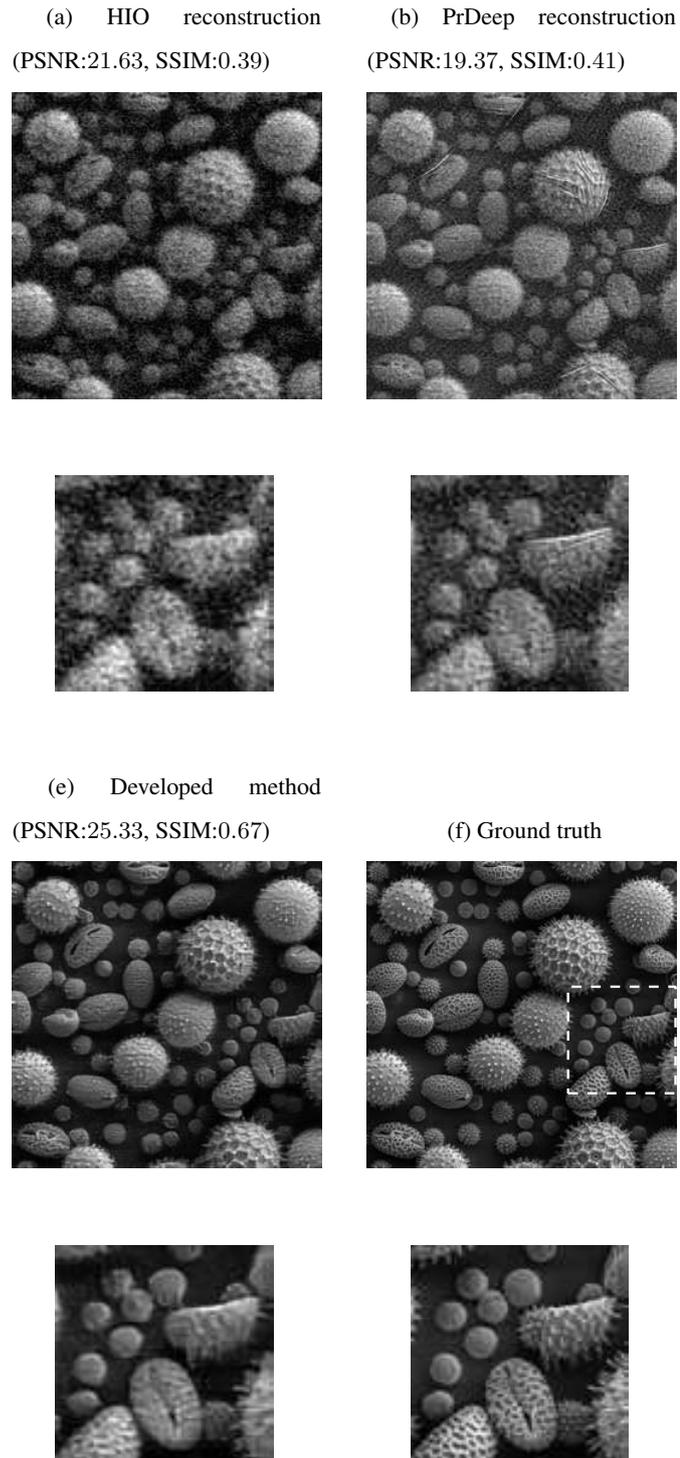
Figure 3.7: The reconstructions of the different algorithms for the "Pollen" test image for $\alpha$=3 case.

initialization and image statistics compared to HIO and prDeep.
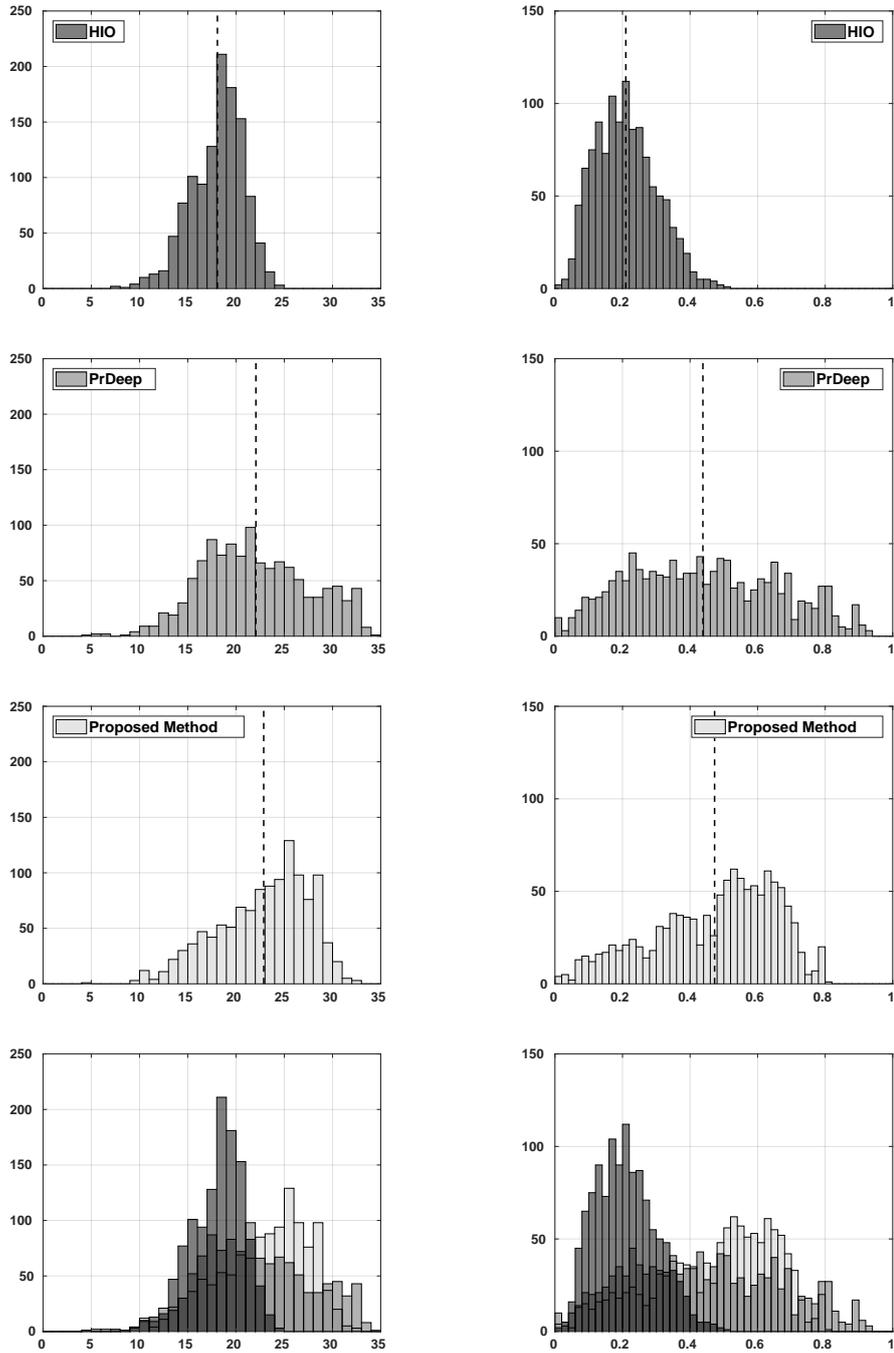


Figure 3.8: The PSNR (left column) and SSIM (right column) histograms for the reconstructions of the different methods for 236 test images and 5 Monte Carlo runs for $\alpha = 3$ case. Vertical dashed lines present the average PSNR and SSIM values. At the bottom, overlapping histograms are given for each column.

Sample reconstructions illustrating the performance of the developed approach for different initialization are shown in Fig. 3.9. Here different HIO reconstructions of the same image are used as an initialization for prDeep and the developed method. As seen, for the HIO initialization with the lower PSNR and SSIM values, prDeep reconstruction has more artifacts than the developed method. Hence, Fig. 3.8 and 3.9 together demonstrate that the developed method is more robust to initialization than prDeep.

The average runtime of each method is also given in Table 3.1. As seen, the HIO and the developed method are roughly three-fold faster than prDeep. In fact, the runtime of the HIO initialization stage approximately corresponds to $92\%$ of the runtime of the developed method. Hence our approach not only outperforms the prDeep and HIO methods in terms of reconstruction quality but also is computationally more efficient than prDeep and achieves a computational efficiency almost comparable with the HIO method.

## 3.6    Conclusions

In this chapter, we developed a phase retrieval approach that utilizes two DNNs with the model-based HIO method. The key idea in the approach is the iterative use of a DNN with the HIO method, which simultaneously incorporates the physical model and the constraints into the solution, while avoiding the reconstruction artifacts. The performance of the developed approach is also compared with the classical and state-of-the-art methods through various numerical simulations. The results demonstrate the effectiveness of our approach both in terms of reconstruction quality and computational efficiency. Our approach not only achieves state-of-the-art reconstruction performance but also is more robust to initialization, different noise levels, and image statistics. Moreover, the developed approach achieves a computational efficiency almost comparable with the HIO method.

Note that the developed method contains two DNNs, DNN-1 and DNN-2, each of which is trained to remove HIO artifacts. That is, DNN-1 is trained to remove the artifacts of HIO reconstructions at the output of the initialization stage, and DNN-2

is trained to remove the artifacts of HIO reconstructions at the output of the iterative DNN-HIO stage. These reconstructions have different amounts of artifacts, as one can observe from the PSNR and SSIM values in Table 1. One would expect the trained weights of these two DNNs to vastly differ since each DNN is trained to remove different amount of HIO artifacts. To explore this, we analyzed the frequency response of the first 64 convolution filters in each trained DNN. Almost half of these filters in DNN-1 have similar characteristics, which effectively correspond to low pass filters. On the other hand, DNN-2 filters have varying frequency responses, and a very small fraction of these filters are low-pass. This indicates that the detailed differences between the input and the desired output images (like edges) are lost more in the first convolution filters of DNN-1 and do not propagate much through the network. This is expected since DNN-1 is trained using inputs with larger amount of HIO artifacts. The low-pass behavior of many of the input filters of DNN-1 can be the reason why DNN-1 is less successful in learning the details and leads to over-smoothed images at its output. A more detailed analysis of the filters in DNNs would provide a better understanding of the developed approach, which will be a topic for future study.

The developed approach in this chapter has promising reconstruction. The performance can be further improved through joint training of DNNs and the HIO method. Recurrent neural network models [66] can also be used to improve the reconstruction. Moreover, different loss functions for training can be utilized to improve the method.

(a) HIO reconstruction (PSNR:18.30, SSIM:0.21)

(b) Another HIO reconstruction (PSNR:14.65, SSIM:0.14)

(c) PrDeep reconstruction (PSNR:25.47, SSIM:0.50)

(d) Another prDeep reconstruction (PSNR:17.75, SSIM:0.28)

(e) Developed method recons. (PSNR:25.40, SSIM:0.53)

(f) Another developed method recons. (PSNR:19.71, SSIM:0.43)

(g) Ground truth

(h) Ground truth
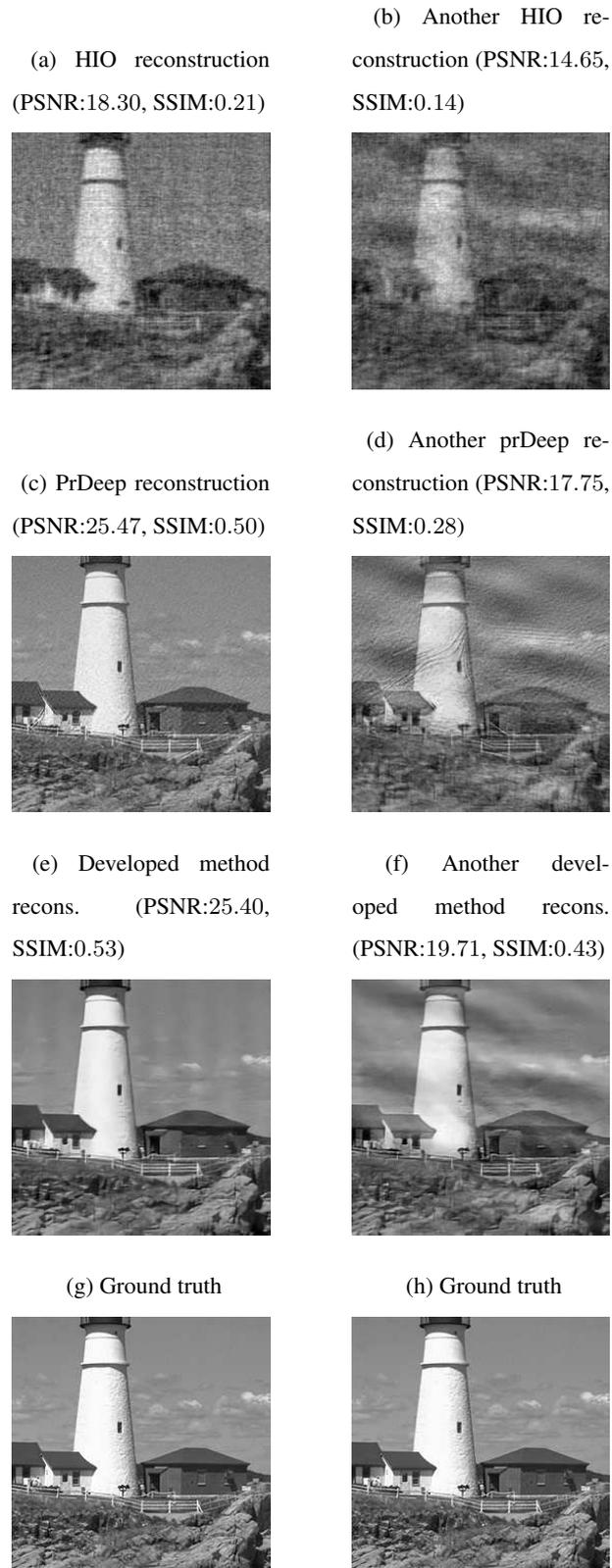
Figure 3.9: Each column represents the reconstructions of different algorithms with two different initialization for "Portland Head Light" test image for $\alpha$=3 case.

# CHAPTER 4

# MODEL-BASED PHASE RETRIEVAL WITH PLUG-AND-PLAY PRIORS

## 4.1 Introduction

As explained in Chapter 3, the classical phase retrieval problem is the recovery of a constrained signal from the magnitude of its Fourier transform, or equivalently from its autocorrelation. In this chapter, a model-based approach using plug-and-play regularization is developed for the solution of the classical phase retrieval problem.

Recently, plug-and-play regularization [16, 17] is successfully applied to many inverse problems which enables to exploit image priors resulting from the state-of-the-art denoisers. The classical regularization approach is to formulate the inverse problem as a Maximum A Posteriori (MAP) estimation problem by using the prior information about the unknown image. This can be considered as an optimization problem containing data fidelity term for the physical model and a regularization term for the prior knowledge about the unknowns. By variable-splitting techniques such as alternating direction method of multipliers (ADMM) [67] and half quadratic splitting (HQS) [18, 19], this problem can be divided into several sub-problems to handle the data fidelity and regularization terms separately. The sub-problem containing regularization term can be solved with any denoiser algorithm. Therefore, different type of denoisers can be easily plugged into the method to exploit different image priors. This plug-and-play approach has been applied to many inverse problems [17, 68].

In the last few years, DNNs, in particular DNN-based denoisers, have been utilized for the regularization of model-based inversion methods by using plug-and-play regularization and its variants [16, 17]. One of the benefits of this approach is that DNN-based image priors can be easily exploited in different inverse problems

while conventional use of DNNs in an end-to-end fashion requires a time-consuming training process for different forward models in inverse problems. The regularization with DNN-based image priors have been also applied for several inverse problems [61, 69, 70, 71] including phase retrieval [2].

In this chapter, an ad hoc model-based reconstruction method using plug-and-play regularization is developed for the solution of the classical phase retrieval problem. HQS is used to decouple the inverse problem into two separate sub-problems containing the data fidelity and prior terms. The sub-problem containing the data fidelity term is solved in an ad hoc manner with the HIO method for a small number of iterations. The other sub-problem containing the prior term is solved with a deep learning-based denoiser. The iterative solution of these sub-problems with a varying split parameter provides the final reconstruction. The developed method has the flexibility to be utilized for different magnitude measurements including CDP measurements and over-sampled Fourier magnitude measurements. Moreover, any denoiser algorithm, which exploits different image priors can be easily incorporated into the developed method. Here, we present our method for Fourier magnitude measurements. The performance of the developed approach is compared with the classical and state-of-the-art methods through numerical simulations. The results demonstrate the effectiveness of our approach both in terms of reconstruction accuracy and computational efficiency.

The rest of this chapter is organized as follows. Related work on phase retrieval and plug-and-play priors for image reconstruction are discussed in Section 4.2. Section 4.3 presents the developed approach. The performance of the approach is compared with the classical and state-of-the-art methods in Section 4.4 through simulations. Finally, we summarize the results and conclude in Section 4.5.

## 4.2 Related Work

### 4.2.1 Hybrid Input-Output (HIO) method for Phase Retrieval

As discussed in more detail in Chapter 3, the most commonly used alternating projection-based method is the HIO algorithm [15], which is developed based on the ER algo-

rithm. The iterations of the HIO method can be summarized as follows:

$$\mathbf{x}_{k+1}[n] = \begin{cases} \mathbf{x}'_k[n] & \text{for} \quad n \notin \gamma \\ \mathbf{x}_k[n] - \beta \mathbf{x}'_k[n] & \text{for} \quad n \in \gamma \end{cases} \tag{41}$$

where

$$\mathbf{x}'_k = \mathbf{F}^{-1} \left\{ \mathbf{y} \odot \frac{\mathbf{F}\mathbf{x}_k}{|\mathbf{F}\mathbf{x}_k|} \right\} \tag{42}$$

Here, $\mathbf{x}_k \in \mathbb{R}^{N^2}$ is the reconstruction at the $k^{th}$ iteration, $\mathbf{F}^{-1}$ denotes the inverse DFT matrix, $\odot$ represents the element-wise (Hadamard) multiplication operation, $\beta$ is a constant parameter (with a typical value of $0.9$) and $\gamma$ is the set of indices $n$ for which $\mathbf{x}'_k[n]$ violates the space domain constraints [15].

### 4.2.2 Plug-and-play Priors for Inverse Problems

The classical regularization approach is to formulate the inverse problem as a maximum posterior (MAP) estimation problem by using the prior statistical knowledge about the unknown image and measurements. This corresponds to an optimization problem containing a data fidelity term for the model and a regularization term for the prior information about the unknowns. By variable splitting techniques, this problem can be divided into several sub-problems to handle the data fidelity and regularization terms separately. In particular, the sub-problem containing the prior term corresponds to a denoising problem, which can be solved with any denoising algorithm. This is the main idea in plug-and-play regularization [16, 17]. This regularization is a flexible model-based approach since the same denoiser can be used for the solution of different inverse problems. This approach has been applied to many inverse problems [17, 68].

There are several variable splitting techniques including alternating direction method of multipliers (ADMM) [67] and half quadratic splitting (HQS) [18, 19], which are utilized for plug-and-play regularization. ADMM is used for plug-and-play regularization to solve several inverse problems [16, 72]. Moreover, plug-and-play regularization with HQS is also applied to several inverse problems [61, 73].

Recently, DNNs are utilized for the regularization of model-based reconstruction methods because of its success in various inverse problems [46]. In particular, DNN-

based denoisers are used for plug-and-play regularization because they provide state-of-the-art performance in denoising. The plug-and-play regularization with DNN-based denoiser has been applied to many inverse problems including deconvolution, denoising, superresolution and demosaicking [61, 69, 70, 71], as well as phase retrieval [2, 74].

For the classical phase retrieval problem, a variant of plug-and-play regularization has been utilized in the PrDeep method [2]. This variant is called regularization by denoising (RED) [74]. By proposing a different regularization term, the RED approach penalizes the residual difference between the unknown image and its denoised version from an independent denoiser. Also, it penalizes the correlation between the unknown image and this residual. PrDeep [2] method uses RED approach with a deep learning-based denoiser to solve the classical phase retrieval problem. In this method, the forward-backward splitting algorithm is used for the resulting optimization problem [75]. This method is also applied to CDP measurements.

## 4.3 Hybrid Input-output Approach with Plug-and-play Regularization

We developed a model-based method using plug-and-play regularization for the solution of the classical phase retrieval problem. The main idea is to incorporate a learning-based prior to the HIO method through plug-and-play regularization [16, 17]. The developed method is inspired from the HQS method, which is used for plug-and-play regularization. HQS is used to decouple the inverse problem into two separate sub-problems containing the data fidelity and prior terms [18, 19]. The sub-problem containing the data fidelity term is solved in an ad hoc manner with the HIO method for a small number of iterations. The other sub-problem containing the prior term is solved with a deep learning-based denoiser. The iterative solution of these sub-problems with a varying split parameter provides the final reconstruction. The developed method not only achieves state-of-the-art reconstruction performance but also has flexibility to be utilized for different magnitude measurements including CDP measurements and oversampled Fourier magnitude measurements. Also, any denoiser algorithm can be easily incorporated into our method. Here we formulated our method for Fourier magnitude measurements, however, it can also be utilized for

CDP measurements.

In the developed method, the splitting approach in the HQS method is adapted to use the HIO method and a denoiser algorithm together. Now, the HQS method is explained to clarify the idea behind the developed method. A linear inverse problem is formulated as an optimization problem by using MAP estimation. This formulation can be given as

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \lambda\Phi(\mathbf{x}) \tag{43}$$

Here, $\mathbf{y} \in \mathbb{R}^{M^2}$ denotes the noisy measurements, $\mathbf{A}$ is the forward model matrix for the inverse problem, $\mathbf{x} \in \mathbb{R}^{N^2}$ represents the unknown image of interest, $\lambda$ denotes the regularization parameter and $\Phi(\mathbf{x})$ represents the regularization function related to the prior statistical knowledge about the unknown image $\mathbf{x}$. The problem in Eqn. 43 is split into sub-problems by using HQS to utilize plug-and-play regularization. By defining an auxiliary variable $\mathbf{z}$, Eqn. 43 is reformulated as follows:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \lambda\Phi(\mathbf{z}) \quad \text{s.t.} \quad \mathbf{z} = \mathbf{x} \tag{44}$$

Then HQS can be used to solve the problem in Eqn. 44 by expressing it as follows:

$$\hat{\mathbf{x}}, \hat{\mathbf{z}} = \underset{\mathbf{x}, \mathbf{z}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \lambda\Phi(\mathbf{z}) + \frac{\mu}{2}\|\mathbf{x} - \mathbf{z}\|^2 \tag{45}$$

where $\mu$ is a non-decreasing penalty parameter which varies iteratively. The problem in 45 can be solved by iteratively handling the following sub-problems:

$$\mathbf{x}_{k+1} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \mu\|\mathbf{x} - \mathbf{z}_k\|^2 \tag{46a}$$

$$\mathbf{z}_{k+1} = \underset{\mathbf{z}}{\operatorname{argmin}} \frac{\mu}{2}\|\mathbf{z} - \mathbf{x}_{k+1}\|^2 + \lambda\Phi(\mathbf{z}) \tag{46b}$$

As seen from Eqn. 46, the data fidelity and regularization terms are split into two sub-problems by using the HQS approach. This approach is used to solve several linear inverse problems with deep learning-based denoisers [61]. For linear inverse problems, the sub-problem in Eqn. 46a has a closed form solution. The second sub-problem in Eqn. 46b corresponds to a denoising problem, which is solved with a deep learning-based denoiser approach in [61]. Hence, instead of an analytical regularization function, a deep learning-based denoiser approach is used. The input of this denoiser approach is $\mathbf{x}_{k+1}$ and the output is $\mathbf{z}_{k+1}$. The denoiser networks in

this approach are trained to solve Gaussian denoising problems with different noise levels by using noisy images and their corresponding ground truth images.

If the HQS method is applied to the classical phase retrieval problem, the optimization problem becomes non-convex and the sub-problem in Eqn. 46a does not have any closed form solution for the classical phase retrieval problem. For this purpose, the HIO method is used in the developed method although it does not exactly solve this sub-problem. The HIO method minimizes an unknown distance between the Fourier magnitude of the reconstruction and the noisy measurement $\Delta(\mathbf{y} - |\mathbf{Fx}|)$ and enforcing space domain constraints including non-negativity, real valuedness, and finite support. One of the motivations for using the HIO method here is that error reduction (ER) algorithm, which is a descendant of HIO, corresponds to the steepest descent method for optimization of $\|\mathbf{y} - |\mathbf{Fx}|\|^2$ [15]. The HIO method is the improved version of the ER algorithm, and also the most commonly used method for phase retrieval. Therefore, in the developed method, the HIO method is used to incorporate the physical model and the constraints into the reconstruction. For each iteration, $\mathbf{z}_k$ is chosen as the initialization of the HIO method. As seen in HQS iterations (Eqn. 46), the $\mu$ parameter is increased with iterations since $\mathbf{x}$ and $\mathbf{z}$ should be equal to each other at the end of the iterations. To incorporate this effect into the HIO method, the noisy measurement $\mathbf{y}$ is updated at each iteration as the linear (convex) combination of the Fourier magnitude of the current estimate $\mathbf{z}_k$ and the measurement $\mathbf{y}$:

$$\mathbf{y}'_k \leftarrow \eta\mathbf{y} + (1 - \eta)|\mathbf{Fz}_k| \tag{47}$$

where $\mathbf{y}'_k$ denotes the measurements used in the HIO method at the $k^{th}$ iteration, the parameter $\eta$ is a non-increasing parameter which varies between $[0, 1]$ iteratively. It is directly related to the $\mu$ parameter. The relation is explained with details below.

The sub-problem in Eqn. 46b can be reformulated as follows:

$$\mathbf{z}_{k+1} = \underset{\mathbf{z}}{\mathrm{argmin}} \; \frac{1}{2(\sqrt{\lambda/\mu})^2}\|\mathbf{x}_{k+1} - \mathbf{z}\|^2 + \Phi(\mathbf{z}) \tag{48}$$

This corresponds to a Gaussian denoising sub-problem with a standard deviation of $\sqrt{\lambda/\mu}$. Therefore, any Gaussian denoiser $D(\cdot, \cdot)$ can be used to solve this sub-problem. As a result, the developed method can be summarized as follows:

$$\mathbf{y}'_k \leftarrow \eta\mathbf{y} + (1 - \eta)|\mathbf{Fz}_k| \tag{49a}$$

52

$$\mathbf{x}_{k+1} = HIO(\mathbf{y}'_k, \mathbf{z}_k) \tag{49b}$$

$$\mathbf{z}_{k+1} = D(\mathbf{x}_{k+1}, \sqrt{\lambda/\mu}) \tag{49c}$$

Since $\mu$ is a non-decreasing parameter, $\sigma = \sqrt{\lambda/\mu}$ is a non-increasing parameter for the denoiser. Therefore, we used the normalized $\sigma/\sigma_{max}$ value as $\eta$ parameter for the measurement update. The implementation of the overall method is given in Algorithm 1. In the implementation, $\sigma$ is chosen as a logarithmically decreasing parameter with iterations.

A DNN-based denoiser, which was used for plug-and-play regularization in [61], is chosen as a denoiser algorithm for the developed method. The main reason here is that the DNN-based denoisers have become state-of-the-art for different denoising problems in recent years [46]. Nevertheless, different denoising algorithms can be also easily incorporated into the developed method.

For initialization of our method, the HIO method was first run with $50$ different random initialization for $50$ iterations. Then, the reconstruction with the lowest residual $\|\mathbf{y} - |\mathbf{F}\hat{\mathbf{x}}|\|_2$ was used for another HIO run for $1000$ iterations. The resulting reconstruction was used as an initialization of the developed method. The overall initialization procedure is illustrated in Fig. 4.1.
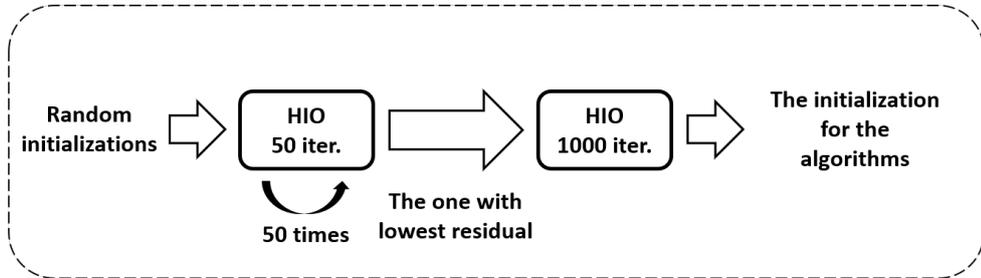


Figure 4.1: The HIO initialization procedure for the developed method and the competing algorithms (Residual:$\|\mathbf{y} - |\mathbf{F}\hat{\mathbf{x}}|\|_2$).

## 4.4 Numerical Results

In this section, the effectiveness of the developed method is illustrated through numerical simulations. For this, we consider the same image dataset in Chapter 3 and

---
**Algorithm 1** Plug and Play HIO
---
**Input:** $\mathbf{y}, K, T, \beta,$

    $\sigma \in \mathbb{R}^T$, which is a logarithmically decreasing vector between $[\sigma_{max}, \sigma_{min}]$.

**Output:** $\hat{\mathbf{x}}_T$

    *Initialization* :

1: $\hat{\mathbf{x}}_0 \leftarrow$ HIO initialization procedure

    *Main iterations*

2: **for** $k = 0$ to $T - 1$ **do**

3:     $\eta \leftarrow \sigma[k]/\sigma_{max}$

4:     $\mathbf{y'}_k \leftarrow \eta \mathbf{y} + (1 - \eta)|\mathbf{F}\hat{\mathbf{x}}_k|$

5:     **for** $i = 0$ to $K - 1$ **do**

6:         $\hat{\mathbf{x}}'_k \leftarrow \mathbf{F}^{-1} \left\{ \mathbf{y'}_k \odot \mathbf{F}\hat{\mathbf{x}}_k / |\mathbf{F}\hat{\mathbf{x}}_k| \right\}$

7:         $\gamma \leftarrow$ the set of indices for which $\hat{\mathbf{x}}'_k$ violates the space domain constraints (such as support and non-negativity)

8:         $\hat{\mathbf{x}}''_k[n] = \begin{cases} \hat{\mathbf{x}}'_k[n] & \text{for} \quad n \notin \gamma \\ \hat{\mathbf{x}}_k[n] - \beta\hat{\mathbf{x}}'_k[n] & \text{for} \quad n \in \gamma \end{cases}$

9:     **end for**

10:     $\hat{\mathbf{x}}_{k+1} \leftarrow D(\hat{\mathbf{x}}''_k, \sigma[k])$

11: **end for**

12: **return** $\hat{\mathbf{x}}_T$
---

compare the reconstruction performance of the developed approach with the classical and state-of-the-art phase retrieval methods. The effect of the method parameters on the reconstruction performance is also investigated.

To compare the algorithms in terms of noise tolerance, image generality, and computational efficiency, the reconstruction performance is investigated using two different kind of images, which are called natural and unnatural images.

For testing, both natural and unnatural images are used as in Chapter 3. This test dataset consists of 236 images containing 230 natural and 6 unnatural images. These include 200 test images of BSD, 24 Kodak dataset images [58], 6 natural and 6 unnatural images taken from [2]. The unnatural image dataset consists of images acquired by scanning electron microscopes and telescopes, as shown in Fig. 3.4. The pixel values of all images are between 0 and 255, and all are of size $256 \times 256$.

The noisy Fourier measurements were simulated using Eqn. 31 with $\alpha = 2, 3, 4$, resulting in an average SNR of 33.39, 31.66, 30.40 dB respectively (where SNR $= 10 \log(\| |\mathbf{Fx}|^2 \|_2 / \| \mathbf{y}^2 - |\mathbf{Fx}|^2 \|_2))$. These measurements were used to obtain the HIO reconstructions by using the initialization procedure given in Fig. 4.1. These are

regarded as the final HIO reconstruction for comparison, and also used as an initialization for all other competing algorithms. All computations were done using MATLAB with MatConvNet toolbox [64] and NVIDIA Geforce GTX TITAN X GPU.

The reconstructions of the developed approach were compared with the true images using peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [65]. For comparison, the reconstructions of the HIO method and prDeep [2], one of the state-of-the-art deep learning-based phase retrieval algorithms, were also obtained.

A DNN-based denoiser, which is developed in [61], is used as a denoiser in the developed method. The DNN architecture that is used in this denoiser contains dilated convolution layers, batch normalization, and residual learning to increase the training and testing performance [61, 46]. This DNN architecture provides state-of-the-art denoising performance [47]. Moreover, the plug-and-play regularization with this denoiser achieves state-of-the-art performances in image deconvolution and superresolution [61]. For this denoiser model, 25 DNNs are trained for Gaussian noise level $[0, 50]$ with a step size of 2 for each model [61]. More details about the architecture of these DNNs are given in [61].

To train DNNs for this denoiser, the training dataset consists of $400$ images of Berkeley segmentation dataset (BSD) [59], $400$ selected images from validation set of ImageNet database [60, 61], and $4,744$ images of Waterloo Exploration Database [62]. These images were cropped into small patches of size 35×35 and N=256×4,000 patches for training. Training was performed by minimizing the mean squared error (MSE) based loss between the reconstructions at the output of each DNN and the residue between true images and noisy images. In other words, the aim of the DNN was to learn the residue between noisy and clean images. ADAM solver was used for optimization [76].

In the developed method, there are several parameters including the number of HIO iterations $K$, the number of total iterations $T$, the maximum and minimum noise levels $\sigma_{max}$, $\sigma_{min}$, and $\beta$. As discussed earlier in Section 3.3.1, the parameter $\beta$ in the HIO method is set to its common value of 0.9. In fact, $\sigma_{max}$ and $\sigma_{min}$ have to be between $[0, 50]$ because of the chosen denoiser model. Before optimizing $\sigma_{max}$ and $\sigma_{min}$ jointly, which are directly related to the regularization parameter $\lambda$, the

parameters $K$ and $T$ are selected as $5$ and $200$ respectively because it is observed that these iteration numbers are generally sufficient for the algorithm to converge. We consider the parameter space for the search as $\sigma_{max} = [30, 40, 50]$ and $\sigma_{min} = [1, 3, 5, 7]$. The average reconstruction performance of the method for $236$ test images and $5$ Monte Carlo runs are analyzed for the Poisson noise case with $\alpha = 3$.

The average PSNR and SSIM values of the developed method are given for $\sigma_{max} = [30, 40, 50]$ and $\sigma_{min} = [1, 3, 5, 7]$ in Fig. 4.2 and 4.3. These figures suggest that the optimal $\sigma_{max}$ and $\sigma_{min}$ values are $40$ and $5$ respectively. Hence, we set $\sigma_{max}$ and $\sigma_{min}$ to these values for the subsequent simulations.
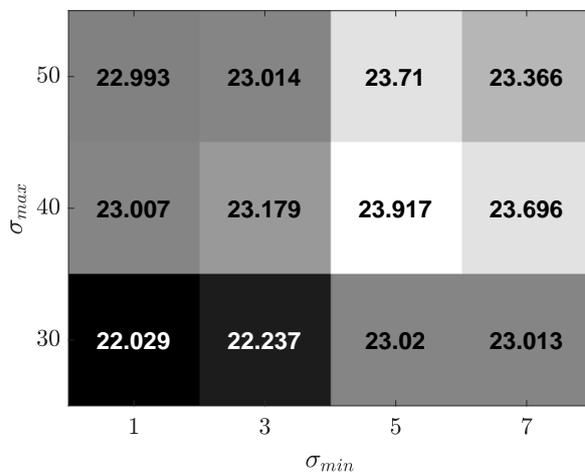


Figure 4.2: The PSNR values for different $\sigma_{max} = [30, 40, 50]$ and $\sigma_{min} = [1, 3, 5, 7]$

In Table 4.1, the average reconstruction performance of the algorithms for $236$ test images and $5$ Monte Carlo runs are given for different amount of Poisson noise ($\alpha = 2, 3, 4$). As seen in the table, for all cases, the developed method outperforms the HIO and prDeep methods in terms of both PSNR and SSIM, while requiring little additional runtime compared to HIO. By utilizing the developed method, many of the HIO artifacts can be successfully removed while preserving the image characteristics. The table illustrates that the developed method provides the best reconstruction performance.

Sample reconstructions for a natural image in the test dataset are shown in Fig. 4.4. As seen from the figures, the developed approach provides the best reconstruction
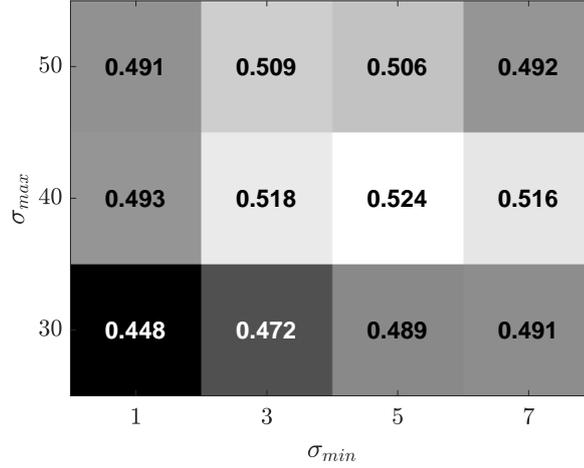
Figure 4.3: The SSIM values for different $\sigma_{max} = [30, 40, 50]$ and $\sigma_{min} = [1, 3, 5, 7]$

Table 4.1: The average reconstruction and runtime performances for $236$ test images ($5$ Monte Carlo runs)

| $\alpha = 2$ (Avg. SNR: 33.39 dB) | Avg. PSNR (dB) | | | Avg. SSIM | | | Avg. runtime (sec.) |
|---|---|---|---|---|---|---|---|
| | Overall | Natural | Unnatural | Overall | Natural | Unnatural | |
| The HIO method | 18.97 | 18.92 | 20.78 | 0.28 | 0.29 | 0.26 | **55.40** |
| PrDeep | 23.45 | 23.49 | 21.72 | 0.51 | 0.51 | 0.24 | 169.81 |
| Developed method | **24.87** | **24.86** | **25.56** | **0.57** | **0.57** | **0.33** | 69.83 |
| $\alpha = 3$ (Avg. SNR: 31.66 dB) | Avg. PSNR (dB) | | | Avg. SSIM | | | Avg. runtime (sec.) |
| | Overall | Natural | Unnatural | Overall | Natural | Unnatural | |
| The HIO method | 18.07 | 18.02 | 19.97 | 0.21 | 0.21 | 0.14 | **55.61** |
| PrDeep | 22.06 | 22.09 | 20.91 | 0.44 | 0.44 | 0.22 | 171.02 |
| Developed method | **23.92** | **23.92** | **23.98** | **0.52** | **0.53** | **0.32** | 70.27 |
| $\alpha = 4$ (Avg. SNR: 30.40 dB) | Avg. PSNR (dB) | | | Avg. SSIM | | | Avg. runtime (sec.) |
| | Overall | Natural | Unnatural | Overall | Natural | Unnatural | |
| The HIO method | 17.34 | 17.30 | 18.72 | 0.16 | 0.17 | 0.10 | **55.78** |
| PrDeep | 20.69 | 20.70 | 20.38 | 0.37 | 0.38 | 0.18 | 172.47 |
| Developed method | **22.41** | **22.39** | **23.09** | **0.45** | **0.46** | **0.26** | 70.49 |

visually as well as in terms of used quantitative image quality measures (PSNR and SSIM). In fact, our approach generally does not introduce artifacts and errors like the HIO and prDeep methods.

Another sample reconstructions for a different natural image are illustrated in Fig. 4.5. The developed approach shows the best reconstruction performance for this sample image as well.

(a) HIO reconstruction
(PSNR:17.68, SSIM:0.32)

(b) PrDeep reconstruction
(PSNR:25.35, SSIM:0.71)

(c) Developed method
(PSNR:27.87, SSIM:0.77)

(d) Ground truth
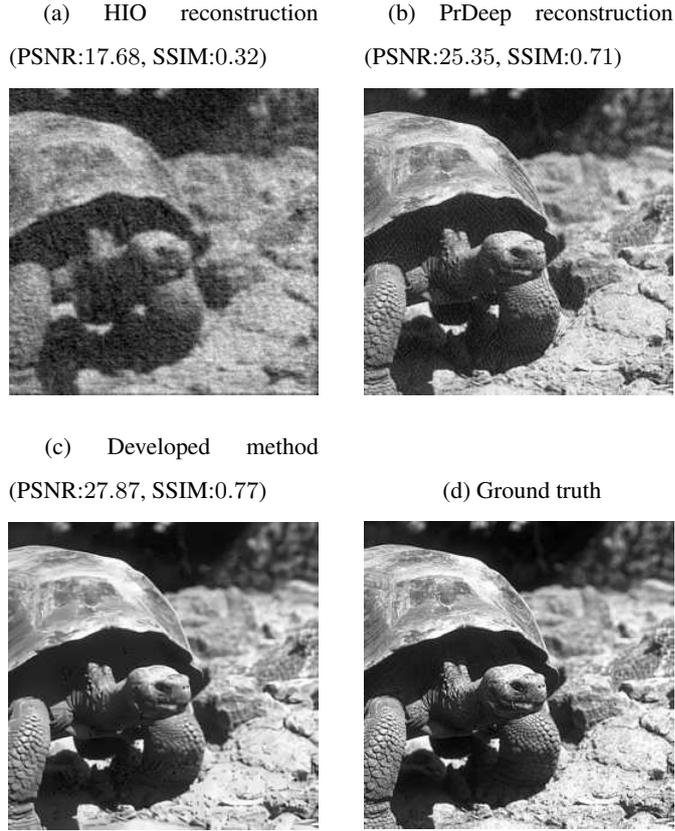
Figure 4.4: The reconstructions of the different algorithms for the "Turtle" test image for $\alpha$=3 case.

To assess the performance of different algorithms in terms of image generality, the results for both natural and unnatural test images are separately provided in Table 4.1. Although the denoiser model in the developed method is trained with only natural images, the developed method shows the best reconstruction performance not only for natural images but also for unnatural images, which have distinct statistics from natural images. In particular, the performance of the prDeep method substantially degrades for unnatural images, as expected, since its reconstruction relies on a regularization prior learned from natural images as seen in the table. To illustrate these points, sample reconstructions for an unnatural image in the test dataset are shown in Fig. 4.6.

The developed approach also appears to be robust to different noise levels. As seen from the table, the reconstruction performance of the approach surpasses the other

(a) HIO reconstruction (PSNR:18.30, SSIM:0.21)

(b) PrDeep reconstruction (PSNR:25.47, SSIM:0.50)

(c) Developed method (PSNR:29.30, SSIM:0.58)

(d) Ground truth
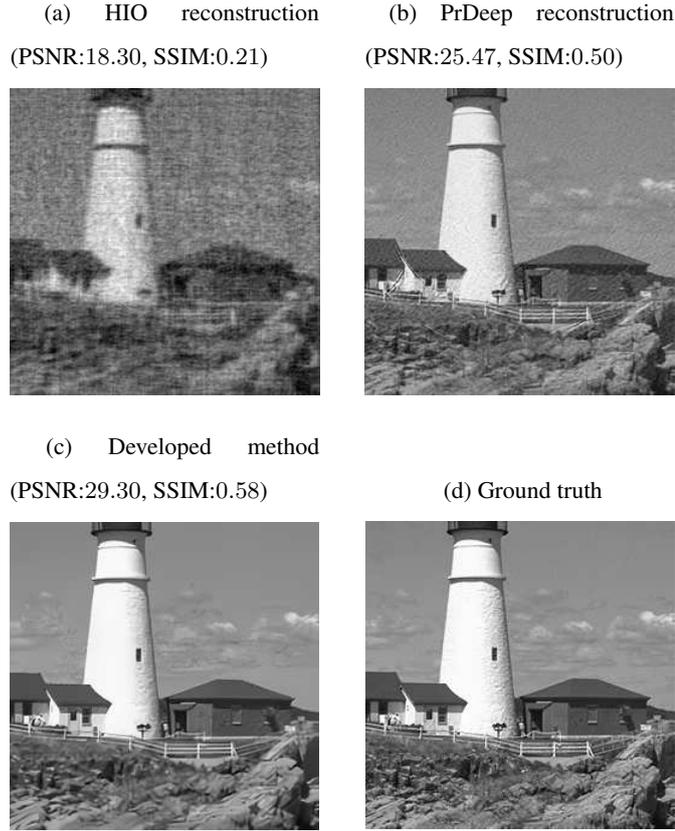
Figure 4.5: The reconstructions of the different algorithms for a sample "Portland Head Light" test image for $\alpha$=3 case.

methods for different noise levels ($\alpha = 2, 4$) as well, even though the parameters of the developed method are optimized for $\alpha = 3$.

As mentioned before, phase retrieval algorithms are generally sensitive to initialization because of the nonlinearity involved in the problem. To illustrate the robustness of the developed approach to different initialization and image characteristics, the PSNR and SSIM histograms are provided in Fig. 4.7 for each method (when $\alpha = 3$). These include reconstructions obtained with $236$ distinct test images and $5$ Monte Carlo runs, which means that $5$ different initialization is used for each test image. As seen from the histograms, although the histogram for the prDeep reconstructions has more counts in higher PSNR and SSIM values, our method (Plug and Play HIO) attains a higher average PSNR and SSIM, as well as a smaller spread around these averages. These results suggest that the performance of the developed approach is more

robust to different initialization and image statistics compared to HIO and prDeep.

The average runtime of each method is also given in Table 4.1. As seen, the HIO and the developed method are roughly 2.5-fold faster than prDeep. In fact, the runtime of the HIO initialization stage approximately corresponds to $80\%$ of the runtime of the developed method. Hence our approach not only outperforms the prDeep and HIO methods in terms of reconstruction quality, but also is computationally more efficient than prDeep and achieves a computational efficiency almost comparable with the HIO method.



(a) HIO reconstruction (PSNR:21.63, SSIM:0.39)

(b) PrDeep reconstruction (PSNR:19.37, SSIM:0.41)

(c) Developed method (PSNR:26.28, SSIM:0.71)

(d) Ground truth
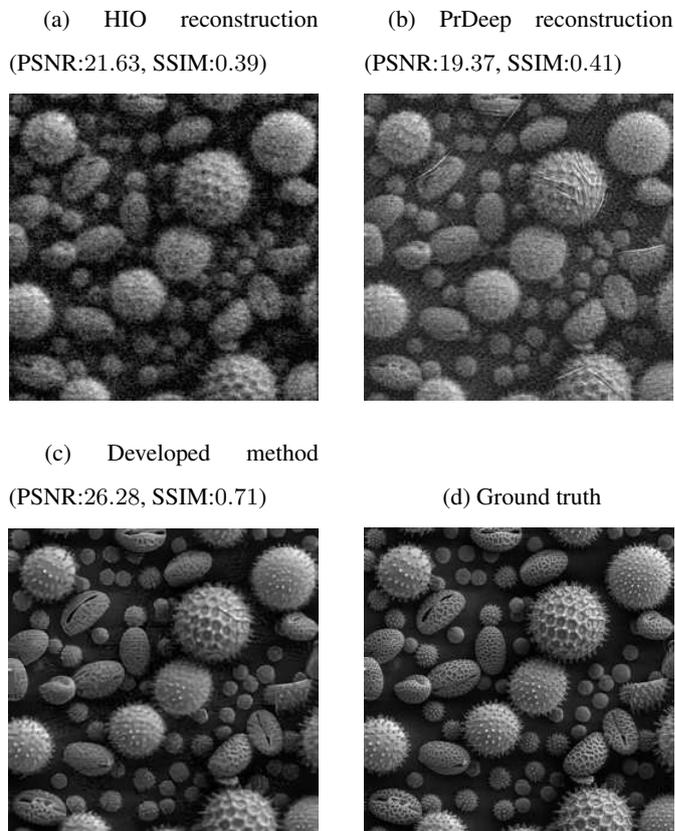
Figure 4.6: The reconstructions of the different algorithms for the "Pollen" test image for $\alpha$=3 case.
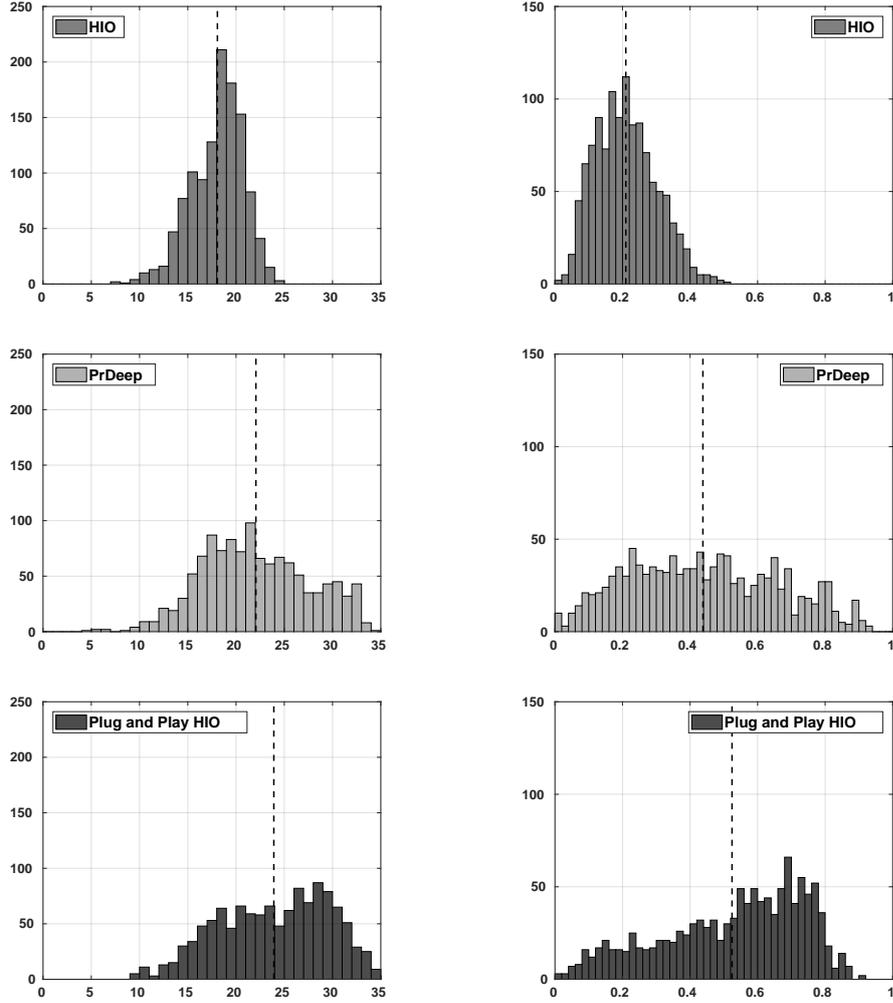
Figure 4.7: The PSNR (left column) and SSIM (right column) histograms for the reconstructions of the different methods for $236$ test images and $5$ Monte Carlo runs for $\alpha = 3$ case. Vertical dashed lines present the average PSNR and SSIM values.

## 4.5 Conclusions

In this chapter, an ad hoc model-based reconstruction method using plug-and-play regularization is developed for the solution of the classical phase retrieval problem. HQS is used to decouple the inverse problem into two separate sub-problems containing the data fidelity and prior terms. The sub-problem containing the data fidelity term is solved in an ad hoc manner with the HIO method for a small number of iterations. The other sub-problem containing the prior term is solved with a deep learning-based

denoiser. The iterative solution of these sub-problems with a varying split parameter provides the final reconstruction.

Any denoiser algorithm can be incorporated into the developed method. Unlike the use of DNNs in an end-to-end fashion, as in Chapter 3, the developed method has a flexibility to be utilized for different magnitude measurements including CDP measurements and oversampled Fourier magnitude measurements. The reconstruction performance of the developed method is compared with the classical and state-of-the-art approaches. The results demonstrate the effectiveness of our approach both in terms of reconstruction accuracy and computational efficiency. As future work, the developed method can be compared with the PrDeep algorithm when the same denoiser models are used. A more detailed numerical analysis and mathematical reasoning are also left as future work.

The developed method in this chapter also provides better reconstruction performance than the method in Chapter 3 in terms of both visual and quantitative comparison. However, each of these developments contributes to our understanding in a different way. This chapter illustrates that the HIO method and possibly other model-based phase retrieval algorithms can be significantly improved by joint use of learning-based priors through plug-and-play regularization. On the other hand, Chapter 3 illustrates that hybrid use of DNNs with model-based methods in an end-to-end fashion can also lead to the development of robust phase retrieval algorithms. Based on these observations, we believe that the joint use of DNNs with model-based methods (such as HIO) deserves further study.

# CHAPTER 5

# CONCLUSIONS

In this thesis, novel analysis and reconstruction methods are developed for coherent imaging systems. First of all, we developed a novel analysis approach for diffraction-limited coherent optical imaging systems. Then, the classical phase retrieval problem, which is one of the fundamental problems in coherent imaging, is considered and two novel reconstruction techniques are developed.

In Chapter 2, a phase-space approach, which utilizes phase-space (space-frequency) window, is developed to analyze diffraction-limited coherent imaging systems. We show how the phase-space window is linked to important imaging parameters of the system including cut-off frequency, diffraction-limited resolution, effective focal length, and magnification. Different than the widely used geometrical-optics methods for analysis, the proposed approach takes into account diffraction effects from all apertures, and also requires a simple computation. To illustrate the utility of the approach, a single-lens system and an objective lens of a microscope are analyzed, and the consistency of the results with the known properties and specifications of the inspected systems is shown.

In Chapter 3 and 4, we focus on the classical phase retrieval problem, which is a fundamental problem in coherent imaging. Two different novel phase retrieval algorithms are developed by jointly exploiting DNNs and the HIO method, which is a traditional model-based inversion method for phase retrieval

In particular, in Chapter 3, a hybrid phase retrieval algorithm utilizing DNNs with the HIO method is developed. The main idea in the developed method is to use a DNN in an iterative manner with the HIO method to improve the HIO reconstructions. Then,

a second DNN is trained to remove the remaining artifacts of the reconstructions. The results demonstrate the effectiveness of our approach, which has relatively little additional computational cost compared to HIO. This approach has promising reconstruction performance. As future work, the performance can be further improved through joint training of DNNs and the HIO method. Moreover, different loss functions can be utilized for training to improve the method.

In Chapter 4, an ad hoc model-based reconstruction method using plug-and-play regularization is developed for the solution of the classical phase retrieval problem. The main idea is to incorporate a learning-based prior to the HIO method through plug-and-play regularization [16, 17] to obtain a flexible model-based approach for phase retrieval. The developed method is flexible such that it can be used with different image priors coming from different denoisers. This method also has the flexibility to be used for different phase retrieval data including coded diffraction patterns and Fourier magnitude measurements. The results demonstrate the effectiveness of our approach both in terms of reconstruction accuracy and computational efficiency. As future work, the developed method can be compared with the PrDeep algorithm when the same denoiser models are used. A more detailed numerical analysis and mathematical reasoning are also left as future work.

Although both of the developed phase retrieval approaches achieve state-of-the-art reconstruction performance, and also more robust to different initialization and noise levels, the developed model-based method in Chapter 4 shows better reconstruction performance than the developed method in Chapter 3 in terms of both visual and quantitative comparison. In fact, the main idea of the method in Chapter 3 is different than the one in Chapter 4. Chapter 4 demonstrates that any image prior coming from different denoiser algorithms can be used with the HIO method through plug-and-play regularization. In particular, a deep learning-based denoiser is used to show the reconstruction performance of the developed method in Chapter 4. Chapter 3 also illustrates that the iterative use of the DNNs with the HIO method may play a key role in developing more reliable algorithms for phase retrieval. Moreover, the main ideas in Chapter 3 and 4 can even be combined to obtain a better method. As another possible future research direction, both methods can also be tested on experimental phase retrieval data.

To conclude, in this thesis, novel analysis and reconstruction methods are developed for coherent imaging systems. For the analysis of the coherent systems, phase-space approach can be a worthy alternative because of its simple but powerful features. For the reconstruction methods, we believe that the hybrid use of DNNs with model-based approaches, as illustrated in Chapter 3 and 4, may play a key role in developing more reliable algorithms for phase retrieval and nonlinear inverse problems in general.

# REFERENCES

[1] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Transactions on Image Processing*, vol. 26, pp. 4509–4522, Sept 2017.

[2] C. Metzler, P. Schniter, A. Veeraraghavan, and R. Baraniuk, "prDeep: Robust phase retrieval with a flexible deep network," in *Proceedings of the 35th International Conference on Machine Learning*, pp. 3498–3507, JMLR.org, 2018.

[3] J. Goodman, *Introduction to Fourier Optics*. Roberts and Company Publishers, 2005.

[4] E. Hecht, *Optics*. Pearson Education, Incorporated, 2017.

[5] R. E. Blahut, *Theory of Remote Image Formation*. Cambridge University Press, 2004.

[6] H. M. Ozaktas and F. S. Oktem, "Phase-space window and degrees of freedom of optical systems with multiple apertures," *J. Opt. Soc. Am. A*, vol. 30, pp. 682–690, Apr 2013.

[7] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: a contemporary overview," *IEEE Signal Processing Magazine*, vol. 32, no. 3, pp. 87–109, 2015.

[8] R. P. Millane, "Phase retrieval in crystallography and optics," *J. Opt. Soc. Am. A*, vol. 7, pp. 394–411, Mar 1990.

[9] G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution fourier ptychographic microscopy," *Nature Photonics*, vol. 7, no. 9, p. 739, 2013.

[10] Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light: Science & Applications*, vol. 7, no. 2, p. 17141, 2018.

[11] J. C. Dainty and J. Fienup, "Phase retrieval and image reconstruction for astronomy," *Image Recovery: Theory Appl*, vol. 13, pp. 231–275, 01 1987.

[12] A. Walther, "The question of phase retrieval in optics," *Optica Acta: International Journal of Optics*, vol. 10, no. 1, pp. 41–49, 1963.

[13] T. J. Schulz and D. L. Snyder, "Image recovery from correlations," *J. Opt. Soc. Am. A*, vol. 9, pp. 1266–1272, Aug 1992.

[14] L. Rabiner and B.-H. Juang, *Fundamentals of speech recognition*. Prentice-Hall, Inc., 1993.

[15] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.*, vol. 21, pp. 2758–2769, Aug 1982.

[16] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *IEEE Global Conference on Signal and Information Processing*, pp. 945–948, IEEE, 2013.

[17] A. Danielyan, V. Katkovnik, and K. Egiazarian, "Deblurring by augmented langrangian with BM3D frame prior," in *Workshop on Information Theoretic Methods in Science and Engineering*, pp. 16–18, WITMSE, 2010.

[18] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Transactions on Image Processing*, vol. 4, no. 7, pp. 932–946, 1995.

[19] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 479–486, IEEE, 2011.

[20] Ç. Işıl and F. S. Oktem, "A phase-space approach to diffraction-limited resolution," in *Imaging and Applied Optics 2018*, p. JM4A.32, Optical Society of America, 2018.

[21] Ç. Işıl and F. S. Oktem, "Simple phase-space approach for the analysis of coherent imaging systems with apertures," *submitted for publication*.

[22] S. A. Collins, "Lens-system diffraction integral written in terms of matrix optics," *JOSA*, vol. 60, no. 9, pp. 1168–1177, 1970.

[23] W. T. Rhodes, "Simple procedure for the analysis of coherent imaging systems," *Optics letters*, vol. 19, no. 19, pp. 1559–1561, 1994.

[24] M. E. Testorf, "A matter of style: Adolf lohmann's approach to optical phase space," *Appl. Opt.*, vol. 56, pp. A69–A78, Jan 2017.

[25] H. Ozaktas, M. Kutay, and Z. Zalevsky, *The Fractional Fourier Transform: With Applications in Optics and Signal Processing*. Wiley Series in Pure and Applied Optics, Wiley, 2001.

[26] F. S. Oktem and H. M. Ozaktas, "Equivalence of linear canonical transform domains to fractional fourier domains and the bicanonical width product: a generalization of the space–bandwidth product," *J. Opt. Soc. Am. A*, vol. 27, pp. 1885–1895, Aug 2010.

[27] A. W. Lohmann, R. G. Dorsch, D. Mendlovic, Z. Zalevsky, and C. Ferreira, "Space–bandwidth product of optical signals and systems," *JOSA A*, vol. 13, no. 3, pp. 470–473, 1996.

[28] A. J. den Dekker and A. van den Bos, "Resolution: a survey," *J. Opt. Soc. Am. A*, vol. 14, pp. 547–557, Mar 1997.

[29] J. Greivenkamp, *Field Guide to Geometrical Optics*. Field Guide Series, Society of Photo Optical Instrumentation Engineers, 2004.

[30] F. L. Pedrotti, L. M. Pedrotti, and L. S. Pedrotti, *Introduction to optics*. Cambridge University Press, 2017.

[31] B. Walker, *Optical Design for Visual Systems*, vol. 45 of *Spie Press Monograph*. SPIE Press, 2000.

[32] M. Laikin, *Lens Design*. Optical Science and Engineering, CRC Press, 2006.

[33] Ç. Işıl, F. S. Oktem, and A. Koç, "A deep learning-based hybrid approach for phase retrieval," in *Imaging and Applied Optics 2019*, Optical Society of America, 2019.

[34] Ç. Işıl, F. S. Oktem, and A. Koç, "Deep iterative reconstruction for phase retrieval," *Applied Optics*, vol. 58, pp. 5422–5431, Jul 2019.

[35] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval from coded diffraction patterns," *Applied and Computational Harmonic Analysis*, vol. 39, no. 2, pp. 277–299, 2015.

[36] M. Hayes, "The reconstruction of a multidimensional sequence from the phase or magnitude of its fourier transform," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 30, pp. 140–154, Apr 1982.

[37] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–250, 11 1972.

[38] J. R. Fienup, "Reconstruction of an object from the modulus of its fourier transform," *Optics letters*, vol. 3, no. 1, pp. 27–29, 1978.

[39] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM Review*, vol. 57, no. 2, pp. 225–251, 2015.

[40] I. Waldspurger, A. d'Aspremont, and S. Mallat, "Phase recovery, maxcut and complex semidefinite programming," *Mathematical Programming*, vol. 149, no. 1-2, pp. 47–81, 2015.

[41] T. Goldstein and C. Studer, "Phasemax: Convex phase retrieval via basis pursuit," *IEEE Transactions on Information Theory*, vol. 64, no. 4, pp. 2675–2689, 2018.

[42] A. M. Tillmann, Y. C. Eldar, and J. Mairal, "Dolphin—dictionary learning for phase retrieval," *IEEE Transactions on Signal Processing*, vol. 64, no. 24, pp. 6485–6500.

[43] V. Katkovnik and J. Astola, "Phase retrieval via spatial light modulator phase modulation in 4f optical setup: numerical inverse imaging with sparse regularization for phase and amplitude," *J. Opt. Soc. Am. A*, vol. 29, pp. 105–116, Jan 2012.

[44] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval via wirtinger flow: Theory and algorithms," *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1985–2007, 2015.

[45] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[46] A. Lucas, M. Iliadis, R. Molina, and A. K. Katsaggelos, "Using deep neural networks for inverse problems in imaging: Beyond analytical methods," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 20–36, 2018.

[47] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian de-noiser: Residual learning of deep cnn for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.

[48] C. Schuler, M. Hirsch, S. Harmeling, and B. Scholkopf, "Learning to deblur.," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 7, p. 1439, 2016.

[49] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4681–4690, IEEE, 2017.

[50] A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica*, vol. 4, no. 9, pp. 1117–1125, 2017.

[51] T. Nguyen, Y. Xue, Y. Li, L. Tian, and G. Nehmetallah, "Deep learning approach for fourier ptychography microscopy," *Optics Express*, vol. 26, no. 20, pp. 26470–26484, 2018.

[52] Ç. Işıl, M. Yorulmaz, B. Solmaz, A. B. Turhan, C. Yurdakul, S. Ünlü, E. Ozbay, and A. Koç, "Resolution enhancement of wide-field interferometric microscopy by coupled deep autoencoders," *Applied Optics*, vol. 57, no. 10, pp. 2545–2552, 2018.

[53] E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-storm: super-resolution single-molecule microscopy by deep learning," *Optica*, vol. 5, no. 4, pp. 458–464, 2018.

[54] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-assisted Intervention*, pp. 234–241, Springer, 2015.

[55] X. Mao, C. Shen, and Y.-B. Yang, "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections," in *Advances in Neural Information Processing Systems*, pp. 2802–2810, Curran Associates, Inc., 2016.

[56] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2536–2544, IEEE, 2016.

[57] A. Dosovitskiy, P. Fischer, E. Ilg, P. Hausser, C. Hazirbas, V. Golkov, P. Van Der Smagt, D. Cremers, and T. Brox, "Flownet: Learning optical flow with convolutional networks," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 2758–2766, IEEE, 2015.

[58] R. W. Franzen, "True color kodak images," *http://r0k.us/graphics/kodak*.

[59] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, vol. 2, pp. 416–423, IEEE, 2001.

[60] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR)*, pp. 248–255, IEEE, 2009.

[61] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3929–3938, IEEE, 2017.

[62] K. Ma, Z. Duanmu, Q. Wu, Z. Wang, H. Yong, H. Li, and L. Zhang, "Waterloo Exploration Database: New challenges for image quality assessment models," *IEEE Transactions on Image Processing*, vol. 26, pp. 1004–1016, Feb. 2017.

[63] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Cognitive Modeling*, vol. 5, no. 3, p. 1, 1988.

[64] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab," in *Proceedings of the 23rd ACM International Conference on Multimedia*, pp. 689–692, ACM, 2015.

[65] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[66] H. K. Aggarwal, M. P. Mani, and M. Jacob, "Modl: Model-based deep learning architecture for inverse problems," *IEEE transactions on medical imaging*, vol. 38, no. 2, pp. 394–405, 2018.

[67] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[68] F. Heide, M. Steinberger, Y.-T. Tsai, M. Rouf, D. Pająk, D. Reddy, O. Gallo, J. Liu, W. Heidrich, K. Egiazarian, *et al.*, "Flexisp: A flexible camera image processing framework," *ACM Transactions on Graphics*, vol. 33, no. 6, p. 231, 2014.

[69] T. Meinhardt, M. Moller, C. Hazirbas, and D. Cremers, "Learning proximal operators: Using denoising networks for regularizing inverse imaging problems," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 1781–1790, IEEE, 2017.

[70] J. R. Chang, C. L. Li, B. Poczos, and B. V. Kumar, "One network to solve them all—solving linear inverse problems using deep projection models," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 5889–5898, IEEE, 2017.

[71] H. K. Aggarwal, M. P. Mani, and M. Jacob, "Modl: Model-based deep learning architecture for inverse problems," *IEEE Transactions on Medical Imaging*, vol. 38, pp. 394–405, Feb 2019.

[72] F. Heide, S. Diamond, M. Nießner, J. Ragan-Kelley, W. Heidrich, and G. Wetzstein, "Proximal: Efficient image optimization using proximal algorithms," *ACM Transactions on Graphics*, vol. 35, no. 4, p. 84, 2016.

[73] N. Yair and T. Michaeli, "Multi-scale weighted nuclear norm image restoration," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3165–3174, 2018.

[74] Y. Romano, M. Elad, and P. Milanfar, "The little engine that could: Regularization by denoising (RED)," *SIAM Journal on Imaging Sciences*, vol. 10, no. 4, pp. 1804–1844, 2017.

[75] T. Goldstein, C. Studer, and R. Baraniuk, "A field guide to forward-backward splitting with a fasta implementation," *arXiv preprint arXiv:1411.3406*, 2014.

[76] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.