Fast Algorithms for phase diversity and phase retrieval

Russell Luke[1] James V. Burke[2] and Richard Lyon[3]

July 26, 2000

Abstract

In this paper we develop the mathematical foundation of, and numerical solution techniques for the problem of deconvolution and wavefront reconstruction. This problem is fundamentally ill-posed. Following the work of previous investigators, this issue is addressed by solving an approximation problem using regularized least-squares and Fourier transform techniques. We show that steepest descent optimization methods applied to a least-squares objective is equivalent to a specific implementation of an iterative transform algorithm. The convergence properties of the method can be derived from standard results in the optimization literature. Convergence is accelerated by using limited memory techniques with trust-regions. CPU time for the method is reduced by using multi-resolution techniques. Numerical experiments on simulated data suggest that the method is both efficient and robust.

[1] *Correspondence should be sent to the first author at the Department of Applied Mathematics Box 352420 University of Washington, Seattle, WA 98195.* EMAIL: *russell@amath.washington.edu*
The first author's work was supported by the NASA GSRP grant NGT5-66
[2] *Department of Mathematics, University of Washington, Seattle, WA 98195*
[3] *NASA Goddard Space Flight Center, Code 930.5, Greenbelt MD 20771*

1 Introduction

In this paper we develop numerical methods for the problem of non-parametric deconvolution and wavefront reconstruction. This problem is fundamentally ill-posed. Following the work of previous investigators, this issue is addressed by solving an approximate regularized problem using least-squares and Fourier transform techniques. An algorithmic approach based on limited memory, trust-region, and multiresolution techniques is proposed. The convergence properties of the method can be derived from standard results in the optimization literature. The results of numerical experimentation on simulated data suggest that the method is both efficient and robust.

Phase reconstruction and deconvolution problems arise in diverse fields such as microscopy, optical design, crystallography, and astronomy. The physical setting is that of a wave generated from an incoherent, monochromatic, far-field source depicted in figure (1). The wave passes through a diffraction grating and is focused with a thin lens onto an array of receptors that measure intensity in a plane parallel to the diffraction grating. The plane in which the receptors lie is referred to as the *image plane*, and the plane in which the diffraction grating lies is referred to as the *pupil plane*. The intensity mapping resulting from a point source is the Green's function or the *point spread function* of the optical system. What is often referred to as the *phase retrieval* problem involves recovering the phase of an electromagnetic wave from intensity measurements alone when the source is a point source. If the image plane lies within a certain region relative to the focal point of the lens, to first order (*i.e.* in the Fraunhoffer approximation), the intensity mapping is the the modulus squared of the Fourier transform of the wavefront on the support of the diffraction grating [14]. In this case, the problem of phase retrieval is one of determining the phase of the wavefront from amplitude measurements in the spatial and Fourier domains. The problem is greatly complicated when the wave source is an unknown extended object rather than a point source. This situation often arises in earth-based optical astronomy where the atmosphere causes aberrations in the wavefront. In this situation one wishes to find both the unknown source and the wavefront aberration *simultaneously*.

Until the 1970's the problem of phase retrieval was thought to be hopeless for a number of reasons. In particular, in one dimension the discrete problem has a multitude of solutions. Indeed there are as many solutions to the problem as there are grid points. To address this problem a number of researchers have proposed the addition of constraints to narrow the number of potential solutions [21, 16, 24]. In 1972 Gerchberg and Saxton [12] proposed a particularly simple and successful projection algorithm for solving phase retrieval problems in two dimensions, though it's convergence properties are not well understood. In 1981 Hayes [15] showed that the solution to the two dimensional phase retrieval problem, if it exists, is almost surely unique up to rotations and linear shifts. One year later Fienup [11] generalized the Gerchberg-Saxton algorithm and analyzed many of it's convergence properties, showing, in particular, that the directions of the projections in the Gerchberg-Saxton algorithm are similar to directions of steepest descent for a least squares performance measure. We show below that the steepest descent direction is *equivalent* to an averaged, simultaneous projection algorithm. Also in 1982, Gonsalves [13] proposed a stable solution for the more general problem of simultaneous phase reconstruction and deconvolution which involved finding the least squares solution for several *diversity* measurements. The solution to the deconvolution problem over several measurements is the *phase diversity* solution. Since the introduction of the Gerchberg-Saxton algorithm and phase diversity numerous papers have been published on both of these problems, too many to provide a comprehensive list here. With a few notable exceptions, *e.g.* [10], all of the methods for solving the wavefront reconstruction and (simultaneous) deconvo-

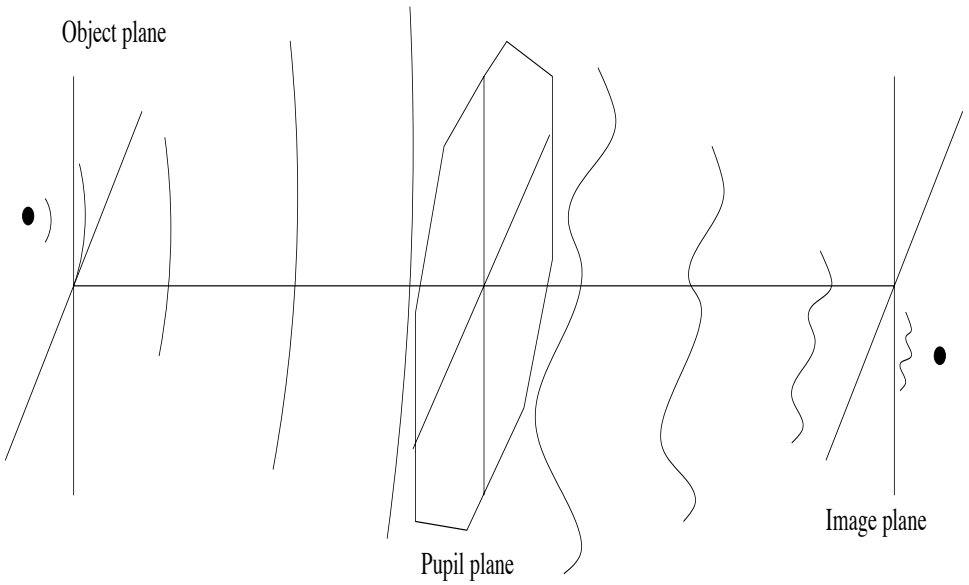

Figure 1: Model optical system

lution problems have at their core the Gerchberg Saxton algorithm and phase diversity. Showing convergence of these algorithms has proven to be notoriously difficult. In particular, with regard to projection algorithms, the nonconvexity of the underlying sets is a fundamental limitation. What are often referred to as convergence results for projection algorithms are statements that the error between iterations will not increase. To date, mathematical convergence for the nonconvex case with convergence rates has not been shown. In what follows we provide a heuristic foundation for stronger, *global* convergence results for a class of non-convex projection algorithms. A detailed convergence proof is beyond the scope of this paper, but can be found in [18]. Our main focus is on algorithms and implementation.

In section 2 we derive the mathematical model for diffraction imaging. In the same section we formulate the abstract optimization problem associated with image recovery. To place our algorithm in the context of previous work we detail the specific optimization problem for the least squares performance measure. Other performance measures, however, are possible. Paxman *et al* [23] have noted that the stochastic model implied by the least squares formulation is not appropriate for all situations. The algorithmic approach we study in this paper is not dependent on the performance measure and therefore can be applied to other statistical models and other performance measures. In Section 3 we detail the limited memory algorithm and the incorporation of trust regions. Section 4 details the performance of the algorithm on simulated data.

2 Mathematical formulation

2.1 Mathematical Model of Diffraction Imaging

What follows is a development of the formalism necessary for a precise discussion of the mathematical model. For ease of exposition we formulate the problem in the continuum leaving discretization until the very end. We represent the wavefront as a complex phasor, $fe^{i\theta}$, with

amplitude f and phase θ. We limit our discussion to ordered pairs $(f,\theta) \in S$ where $S \subset (L^1 \cap L^2)[\mathbb{R}^2,\mathbb{R}] \times (L^1 \cap L^2)[\mathbb{R}^2,\mathbb{R}]$. We denote the source by the function $\varphi : \mathbb{R}^2 \to \mathbb{R}$ and the image by the function $\psi : \mathbb{R}^2 \to \mathbb{R}$. For reasons that will become clear below, we restrict our attention to sources $\varphi \in U \subset (L^1 \cap L^2)[\mathbb{R}^2,\mathbb{R}]$. We show that the images therefore belong to $(L^1 \cap L^2)[\mathbb{R}^2,\mathbb{R}]$.

The optical system we model includes all of space from the pupil plane to the image plane. In the perfect imaging system of figure (1) we assume that the true wavefront is planar and that any deviations from this are *aberrations* in the optical system. Perturbations in the wavefront are reflected in the phase term, θ, and may occur at any point along the axis of propagation. To first order, the data recorded in the image plane is a geometric projection along the axis of propagation. Thus the locations of the aberrations along this axis are not important. Accordingly, we consider the phase of the wavefront to be the phase or aberration of the *generalized pupil function*, $P : \mathbb{R}^2 \to \mathbb{C}$, $P(\underline{x}) = f(\underline{x})e^{i\theta(\underline{x})}$, lying in the pupil plane (see figure 1). The support of f describes the geometry of the aperture in the pupil plane, and θ describes imperfections in the optics that cause changes in the optical path length. For a thorough treatment of the physics see [4, 14]. A simple example of an aberration is defocus. Defocus is often added to the optical system to stabilize numerical schemes.

The ideal optical system of figure (1) is modeled as a convolution operator. The image ψ resulting from a source φ emitting an incoherent, monochromatic, planar, optical wave is given by a Fredholm integral equation of the first kind:

$$\mathcal{K}[f,\theta]\varphi = \psi. \tag{1}$$

Here $\mathcal{K}[f,\theta]$ is a convolution operator whose kernel, $|\kappa[f,\theta]|^2$, is parameterized by f and θ:

$$\mathcal{K}[f,\theta]\varphi \equiv |\kappa[f,\theta]|^2 * \varphi, \tag{2}$$

where

$$\kappa[f,\theta] = (fe^{i\theta})^{\wedge}. \tag{3}$$

The kernel of the convolution operator is the *pointspread function* of the optical system. We denote convolution by $*$. The symbols $^{\wedge}$ and $^{\vee}$ denote the Fourier transform and it's inverse respectively. Since we have limited our discussion to functions on $(L^1 \cap L^2)[\mathbb{R}^2,\mathbb{R}]$ we can define the Fourier transform in the usual sense:

$$u^{\wedge}(\underline{\xi}) \equiv \int_{\mathbb{R}^n} u(\underline{x})e^{-2\pi i \underline{x}\cdot\underline{\xi}}\, d\underline{x}.$$

The Fourier transform is an isomorphism. Consequently,

$$\kappa \in (L^1 \cap L^2)[\mathbb{R}^2,\mathbb{C}]$$

and

$$\mathcal{K} : (L^1 \cap L^2)[\mathbb{R}^2,\mathbb{R}] \to (L^1 \cap L^2)[\mathbb{R}^2,\mathbb{R}].$$

Thus we require $\psi \in (L^1 \cap L^2)[\mathbb{R}^2,\mathbb{R}]$. The problem of recovering φ as well as θ from data ψ and the magnitude of the wavefront, $f = \mathcal{A}$, is an ill-posed inverse problem, linear in φ (deconvolution),

and nonlinear in θ (wavefront reconstruction). Phase retrieval is a specialization of (1) to the case where $\varphi = \delta$, the Dirac delta function.

In phase diversity several data sets, $\psi_j,\ j = 1, \ldots, m$, are collected with the goal of finding the unknown phase common to all. The data sets are called *diversity* images. The diversity images are generated by adding known phase aberrations, $\phi_j,\ j = 1, \ldots, m$, to the unknown phase aberration on the support of the pupil:

$$|\kappa[(\mathcal{X}f, \theta + \phi_j)]|^2 * \varphi = \psi_j, \tag{4}$$

where the indicator function $\mathcal{X}$ denotes the support of the pupil. We formulate the diversity problem as a system of operator equations

$$\underline{\mathcal{K}}[f, \theta]\varphi = \underline{\psi}, \tag{5}$$

where $\underline{\mathcal{K}}[f, \theta] : (L^1 \cap L^2)[\mathbb{R}^2, \mathbb{R}] \to (L^1 \cap L^2)^m[\mathbb{R}^2, \mathbb{R}]$ and $\underline{\psi} \in (L^1 \cap L^2)^m[\mathbb{R}^2, \mathbb{R}]$. Here $\underline{\mathcal{K}}[f, \theta]$ is a linear operator parameterized by the functions f and θ, and defined by

$$\begin{aligned} \underline{\mathcal{K}}[f, \theta] &= (\mathcal{K}_1[f, \theta],\ \ldots,\ \mathcal{K}_m[f, \theta]) \\ &\equiv (\mathcal{K}[(\mathcal{X}f, \theta + \phi_1)],\ \ldots,\ \mathcal{K}[(\mathcal{X}f, \theta + \phi_m)])\,. \end{aligned} \tag{6}$$

Thus $\underline{\mathcal{K}}\varphi = (\mathcal{K}_1\varphi,\ \ldots,\ \mathcal{K}_m\varphi)$. The space $(L^1 \cap L^2)^m[\mathbb{R}^2, \mathbb{R}]$ inherits the usual product space topology.

One must regularize the problem in order to obtain numerical stability since (5) is a Fredholm integral equation of the first kind. We achieve this by perturbing (5) to a "nearby" Fredholm integral equation of the *second* kind. For small $\alpha > 0$,

$$(\underline{\mathcal{K}}[f, \theta] + \alpha\underline{\mathcal{N}}[f, \theta])\varphi = \underline{\psi} \tag{7}$$

where the zero function is the only function contained in the null space of the components of both $\underline{\mathcal{K}}$ and $\underline{\mathcal{N}}$. This is a slightly different way of arriving at some of the more common regularization techniques, among which is the well known Tikhonov regularization. Regularization is a rich topic in and of itself, but we will not touch on it here. For our purposes we require the operator $\underline{\mathcal{N}}$ to be a convolution operator with a real valued kernel satisfying

$$\text{null}\left[(\underline{\mathcal{K}}[f, \theta] + \alpha\underline{\mathcal{N}}[f, \theta])^* (\underline{\mathcal{K}}[f, \theta] + \alpha\underline{\mathcal{N}}[f, \theta])\right] = \{0\},$$

where $(\cdot)^*$ denotes the adjoint. The kernel of the jth component of $\underline{\mathcal{N}}$, $\nu_j \in L^1 \cap L^2[\mathbb{R}^2, \mathbb{R}]$, is parameterized by the functions f and θ:

$$\mathcal{N}_j[f, \theta]\varphi = \nu_j[f, \theta] * \varphi.$$

Problem (5) is still ill-posed, even with regularization. We must therefore seek the best object estimate, φ_*, and the best wavefront estimate, $f_* e^{i\theta_*}$, for a given performance measure, ρ:

$$\begin{array}{cll} (\mathcal{P}) & \text{minimize} & \rho\left[\underline{\psi},\ (\underline{\mathcal{K}}[f, \theta] + \underline{\mathcal{N}}[f, \theta])\,\varphi\right] \\ & \text{over} & \varphi \in U,\ (f, \theta) \in S, \end{array}$$

where ϕ_j, and ψ_j $(j = 1, 2, \ldots, m)$ are given and $U \subset (L^1 \cap L^2)[\mathbb{R}^2, \mathbb{R}]$ and $S \subset (L^1 \cap L^2)^2[\mathbb{R}^2, \mathbb{R}]$. If in addition the amplitude f is assumed known, problem $(\mathcal{P})$ is the phase diversity problem.

We diagonalize the system of equations (7) by transforming the equation to it's Fourier dual. We define the *Fourier dual* to a functional equation as the Fourier transform of both sides of the equation. For example, by the convolution theorem any convolution operator, $\mathcal{G}$, with kernel $g \in L^1$, is associated with a dual multiplication operator G, with "kernel" $g^\wedge$, defined by the Fourier dual to the corresponding operator equation:

$$\mathcal{G}\varphi \equiv g * \varphi = \psi \qquad \stackrel{\mathcal{F}}{\longleftrightarrow} \qquad G\varphi^\wedge \equiv \left(g^\wedge\right)\left(\varphi^\wedge\right) = \psi^\wedge.$$

Given that $\underline{\mathcal{K}} + \alpha\underline{\mathcal{N}}$ is a convolution operator with real kernel, we can apply the convolution theorem to define a dual operator that is diagonal with a Hermitian kernel. By *Hermitian* we mean a function $u : \mathbb{R}^n \to \mathbb{C}$ satisfying $u(\underline{x}) = \overline{u}(-\underline{x})$. Equivalently, u is Hermitian if and only if $u^\wedge$ is a real valued function. As we will see below, this will greatly simplify calculations of the gradient of the objective. As a simple example, let $\nu = \delta$, the Dirac delta function, then $\mathcal{N}$ is the identity operator. Alternatively, if we define $\nu = (1 - \mathcal{X}_\Omega)^\vee$, where $\mathcal{X}_\Omega$ is the indicator function of the support, Ω, of $[|\kappa[f,\theta]|^2]^\wedge$, then $\mathcal{N}[f,\theta]$ is the projection onto the null space of $\mathcal{K}[f,\theta]$.

From equations (2) and (3) it is clear that kernel of the convolution operator $\mathcal{K}$ is the modulus squared of the Fourier transform of the generalized pupil functional, $|\kappa[f,\theta]|^2 = \left(fe^{i\theta}\right)^\wedge \overline{(fe^{i\theta})^\wedge}$. Using the identity $u^{\wedge\wedge} = u(-\underline{x}) = u^{\vee\vee}$, it is straight forward to verify that, for any complex valued scalar functions $u, v \in L^1 \cap L^2[\mathbb{R}^n, \mathbb{C}]$, one has

$$\left(\left(u^\wedge\right)\left(v^\vee\right)\right)^\wedge = u \star v, \tag{8}$$

where $\star$ is the correlation operator defined by

$$u \star v(\underline{x}) \equiv \int_{\mathbb{R}^n} u(\underline{x}')v(\underline{x} + \underline{x}')d\underline{x}'. \tag{9}$$

Thus

$$\left[|\kappa[f,\theta]|^2\right]^\wedge = (fe^{i\theta}) \star \left(\overline{fe^{i\theta}}\right). \tag{10}$$

We denote the Fourier dual of $\underline{\mathcal{K}}$ by $\underline{K}$ where

$$\begin{aligned}
\underline{K}[f,\theta]\varphi^\wedge &= \left(\left[\left(\mathcal{X}fe^{i\theta+\phi_1}\right) \star \left(\overline{\mathcal{X}fe^{i\theta+\phi_1}}\right)\right], \ \ldots, \ \left[\left(\mathcal{X}fe^{i\theta+\phi_m}\right) \star \left(\overline{\mathcal{X}fe^{i\theta+\phi_m}}\right)\right]\right)\varphi^\wedge \\
&= \left(\left[\left(\mathcal{X}fe^{i\theta+\phi_1}\right) \star \left(\overline{\mathcal{X}fe^{i\theta+\phi_1}}\right)\right]\varphi^\wedge, \ \ldots, \ \left[\left(\mathcal{X}fe^{i\theta+\phi_m}\right) \star \left(\overline{\mathcal{X}fe^{i\theta+\phi_m}}\right)\right]\varphi^\wedge\right).
\end{aligned} \tag{11}$$

Similarly, the Fourier dual to $\underline{\mathcal{N}}$ is a multiplication operator denoted by $\underline{N}$ with

$$\underline{N}[f,\theta]\varphi^\wedge = \left(\left(\nu_1[\mathcal{X}f, \theta + \phi_1]\right)^\wedge, \ \ldots, \ \left(\nu_1[\mathcal{X}f, \theta + \phi_m]\right)^\wedge\right)\varphi^\wedge.$$

We can thus write the Fourier dual of (7) as

$$\left(\underline{K} + \alpha\underline{N}\right)\varphi^\wedge = \underline{\psi}^\wedge. \tag{12}$$

The diagonalization of the convolution operator is a crucial property for numerical solutions and the reason for choosing $\mathcal{N}$ to be a convolution operator. Note also that the kernel of $\underline{K} + \alpha\underline{N}$ is Hermetian.

Using the Fourier dual representation we can also write the pointspread function as a quadratic in the dual function $\tilde{f} \in S^\wedge \subset (L^1 \cap L^2)[\mathbb{R}^2, \mathbb{R}]$ where $\tilde{f}e^{i\tilde{\theta}}$ is the Fourier dual of the wavefront $fe^{i\theta}$:

$$|\kappa|^2 = \tilde{f}^2. \tag{13}$$

It is useful to interpret the functions $\tilde{f}e^{i\tilde{\theta}}$ in terms of wave propagation. In geometric optics $\tilde{f}e^{i\tilde{\theta}}$ represents the distribution of *ray* components, *i.e.* the directions of propagation, of the wave $fe^{i\theta}$. In studies of wave propagation in which the Wigner distribution plays a role, the domain of interest is the product space including the physical domain and the frequency domain. The wavefront exists in the physical domain and the distribution of rays exists in the frequency domain. This product space is called *phase space*. Analogously we can formulate problem $(\mathcal{P})$ in phase-space,

$$\begin{array}{lll} (\mathcal{P}') & \text{minimize} & \rho\left[\underline{\psi}, \underline{\mathcal{K}}[f,\theta]\varphi, \underline{\widetilde{\mathcal{K}}}[\tilde{f},\tilde{\theta}]\varphi\right] \\ & \text{subject to} & (fe^{i\theta})^\wedge = \tilde{f}e^{i\tilde{\theta}} \\ & & (f,\theta) \in \mathcal{S}, \quad (\tilde{f},\tilde{\theta}) \in \mathcal{S}^\wedge \\ & & \varphi \in U \end{array}$$

Here $\underline{\widetilde{\mathcal{K}}}[\tilde{f},\tilde{\theta}] : (L^1 \cap L^2)[\mathbb{R}^2,\mathbb{R}] \to (L^1 \cap L^2)^m[\mathbb{R}^2,\mathbb{R}]$ is a convolution operator defined by

$$\underline{\widetilde{\mathcal{K}}}[\tilde{f},\tilde{\theta}]\varphi = \left(\left| \left[\mathcal{X} e^{i\phi_1} (\tilde{f}e^{i\tilde{\theta}})^\vee \right]^\wedge \right|^2 * \varphi, \ \ldots, \ \left| \left[\mathcal{X} e^{i\tilde{\phi_m}} (\tilde{f}e^{i\tilde{\theta}})^\vee \right]^\wedge \right|^2 * \varphi \right). \tag{14}$$

For a general review of this theory see [1, 2].

2.2 Least Squares Minimization

The system of equations (7) is linear in φ and nonlinear in (f,θ). This structure allows us to split the corresponding optimization problem using a Benders decomposition [3]. Benders decompositions are common techniques for splitting large-scale optimization problems such as $(\mathcal{P})$ into smaller problems which can be solved independently of one another in sequence. The least squares performance measure admits a particularly simple way to split the problem. This was first recognized by Gonsalves [13] and later generalized by Paxman *et al* [23].

A common assumption in adaptive optics is that the amplitude of the wavefront across the support of the pupil is constant with unit magnitude. We add this constraint as a penalty in the objective function. Using the least squares performance measure, the problem $(\mathcal{P})$ becomes

$$\begin{array}{lll} (\mathcal{P}_{LS}) & \text{minimize} & J_{LS} = \frac{1}{2}\left\| \left(\underline{\mathcal{K}}[f,\theta] + \alpha \underline{\mathcal{N}}[f,\theta]\right)\varphi - \underline{\psi} \right\|^2 + \frac{1}{2}\left\| f^2 - \mathcal{X}^2 \right\|^2 \\ & \text{over} & \varphi \in U, \ (f,\theta) \in S, \end{array}$$

Benders decomposition involves first obtaining φ_* by optimizing over φ for fixed f and θ. Next we optimize over f and θ holding φ_* fixed. The process is repeated until the iterates exceed some tolerance. Let $(\mathcal{P}_{LS})_{(f,\theta)}$ denote the least squares problem $(\mathcal{P})$ with (f,θ) fixed. For the least

squares performance measure, the optimal solution φ_* to $(\mathcal{P}_{LS})_{(f,\theta)}$ can be written in closed form. We obtain this by considering the Fourier dual to $(\mathcal{P}_{LS})_{(f,\theta)}$

$$\begin{array}{lll} \left(\mathcal{P}_{\widehat{LS}}\right)_{(f,\theta)} & \text{minimize} & J_{LS}^{\wedge} = \frac{1}{2}\left\|(\underline{K}[f,\theta] + \alpha\underline{N}[f,\theta])\,\varphi^{\wedge} - \underline{\psi}^{\wedge}\right\|^2 + \frac{1}{2}\left\|f^2 - \mathcal{X}^2\right\|^2. \\ & \text{over} & \varphi^{\wedge} \in U \end{array}$$

Assume that $\mathcal{N}$ has been chosen so that it's Fourier dual satisfies

$$\text{null}\left[(\underline{K}[f,\theta] + \underline{N}[f,\theta])^*\,(\underline{K}[f,\theta] + \underline{N}[f,\theta])\right] = \{0\}.$$

The optimal dual object estimate, denoted $\varphi_*^{\wedge}$, is the solution to the normal equations:

$$\varphi_*^{\wedge} = \left[(\underline{K} + \underline{\alpha N})^*\,(\underline{K} + \underline{\alpha N})\right]^{-1}(\underline{K} + \underline{\alpha N})^*\underline{\psi^{\wedge}}, \tag{15}$$

By Parseval's relation the optimal values of $(\mathcal{P}_{\widehat{LS}})_{(f,\theta)}$ and $(\mathcal{P}_{LS})_{(f,\theta)}$ are equivalent. The optimal solution to $(\mathcal{P}_{LS})$ is thus $\varphi_* \equiv [\varphi_*^{\wedge}]^{\vee}$. The first step of Benders decomposition can therefore be executed implicitly by substituting $[\varphi_*^{\wedge}]^{\vee}$ directly into $(\mathcal{P}_{LS})$ and optimizing over f and θ alone. Expression (15) is a generalized Wiener filter for diffraction limited systems. The regularization reflects the diffraction limits of the optical system.

If J_{LS} is Fréchet differentiable and the feasible set $U \cap S$ is nonempty and compact then a solution to $(\mathcal{P}_{LS})$ exists. The analytic properties of the objective are detailed in a forthcoming paper [19]. For our purposes we assume that a solution exists and that the objective is uniformly differentiable with uniformly continuous derivative.

2.3 Phase Retrieval

Projection algorithms, among which are iterative transform methods, are central to current numerical techniques for solving the phase retrieval problem [5, 8, 11, 12, 17, 20, 29, 28]. In this section we derive the minimization problem that these projection algorithms attempt to solve. Before we do this, however, we show the relationship between the phase retrieval problem as it is commonly posed and $(\mathcal{P}_{LS})$.

When $\varphi = \delta$, the Dirac delta function, the objective J_{LS} in $(\mathcal{P}_{LS})$, minus the magnitude constraint, simplifies to

$$\frac{1}{2}\left\|(\underline{\mathcal{K}}[f,\theta] + \alpha\underline{\mathcal{N}}[f,\theta])\,\delta - \underline{\psi}\right\|^2 = \frac{1}{2}\sum_{j=1}^{m}\left\|\,|\kappa_j|^2 + \alpha\nu_j - \psi_j\right\|^2.$$

Here $\kappa_j = \kappa[\mathcal{X}f, \theta + \phi_j]$. Regularization of the linear convolution problem is no longer necessary, so we take $\alpha = 0$. It has been noted, however, that the linearized least squares phase retrieval problem is in some sense still ill-posed. To address this a Tikhonov regularization strategy has been suggested. The regularization is critical to theoretical convergence results, though its effect on numerical performance is not always crucial. For details see [10].

The norm squared and modulus squared make the objective effectively a quartic in the unknowns which will tend to flatten out the objective and slow convergence. We therefore consider the following quadratic objective

$$\tilde{J} = \frac{1}{2}\sum_{j=1}^{m}\left\|\,|\kappa_j| - \psi_j^{1/2}\right\|^2. \tag{16}$$

We know ahead of time that the data ψ is non-negative, so the square root is not problematic. Adding to $\tilde{J}$ a quadratic constraint on the amplitude yields the sum of set distance errors:

$$J_{sde} = \frac{1}{2}\sum_{j=1}^{m}\left\| |\kappa_j| - \psi_j^{1/2} \right\|^2 + \frac{1}{2}\left\| f - \mathcal{X} \right\|^2. \tag{17}$$

The phase retrieval problem is posed as finding the minimum of the set distance error J_{sde}:

$$\begin{array}{cll} \left(\mathcal{P}_{sde}\right) & \text{minimize} & J_{sde}[f,\theta] \\ & \text{over} & (f,\theta) \in S. \end{array}$$

The objective J_{sde} is not differentiable in the usual sense. What is needed is a less restrictive notion of differentiation. Variational analysis [25] provides just such tools, however this is beyond the scope of this paper. For our purposes it suffices to perturb the problem $(\mathcal{P}_{sde})$ to a nearby optimization problem with an objective function which *is* differentiable. The gradient is well defined for the perturbed objective function J^{ϵ}_{sde}:

$$J^{\epsilon}_{sde} = \frac{1}{2}\sum_{j=1}^{m}\left\| \frac{|\kappa_j|^2}{|\kappa_j| + \epsilon} - \psi_j^{1/2} \right\|^2 + \frac{1}{2}\left\| \frac{f^2}{f+\epsilon} - \mathcal{X} \right\|^2. \tag{18}$$

We solve the optimization problem

$$\begin{array}{cll} \left(\mathcal{P}^{\epsilon}_{sde}\right) & \text{minimize} & J^{\epsilon}_{sde}[f,\theta] \\ & \text{over} & (f,\theta) \in S. \end{array}$$

As with the regularization suggested in [10], while the regularization above is necessary for theoretical purposes it does not have dramatic numerical effect.

3 Numerical Methods

3.1 Limited Memory BFGS with Trust regions

In this section we briefly review limited memory techniques with trust regions for solving the equation $\nabla J[f_*,\theta_*] = 0$. An optimal solution (f_*,θ_*) to $(\mathcal{P}'_{LS})$ or $(\mathcal{P}_{sde})$, if it exists, will satisfy $\nabla J[f_*,\theta_*] = 0$. The objective functions J_{LS} and J^{ϵ}_{sde} are not convex in f and θ. Therefore all that can be said of the point (f_*,θ_*) is that it is a critical point. This is fundamental to the theory of optimization and is covered at length in [18]. For the purposes of this paper we will be satisfied with solving the equation $\nabla J[f,\theta] = 0$, recognizing, however, that the solution to this equation is not necessarily a solution to the corresponding optimization problem.

The method we propose uses as much information about the objective function as possible while preserving the non-parametric nature of the formulation. Newton's method is an efficient iterative algorithm for solving equations and is based on the first two terms of the Taylor series expansion. For ∇J this can be written formally as

$$\nabla J[(f,\theta) + \underline{d}] \approx \nabla J[f,\theta](\underline{d}) + \nabla^2 J[f,\theta](\underline{d},\underline{d}). \tag{19}$$

Near a solution Newton iterates converge quadratically, *i.e.* if at step i the value of the function is 10^{-2}, the value of the function at the next iterate will be 10^{-4}; the following iterate will yield a function value of 10^{-8} and so forth. The Hessian, $\nabla^2 J[f,\theta;\underline{d}]$, of a function $J:\mathbb{R}^n\to\mathbb{R}$ is an $n\times n$ matrix; thus for large systems it is not feasible to represent the Hessian explicitly in a computer. The approach we follow involves a discretization in terms of pixels. A 512×512 image, for example, yields an optimization problem with 2^{10} variables (recall that we treat *both* the phase θ and amplitude f as unknowns so we have 2×2^9 unknowns). The corresponding Hessian, assuming it exists, is a $2^{10}\times 2^{10}$ matrix. Limited memory methods provide an efficient way to use approximate Hessian information without explicitly forming the matrix. These techniques are derived from quasi-Newton, or matrix secant methods. Quasi-Newton methods are two step quadratic methods that approximate the Hessian by divided differences of the gradients. Newton and Newton-like methods, however, cannot be expected to converge for starting guesses far from the solution or when the curvature is nearly zero. These methods are made robust with the introduction of trust regions. For a thorough treatment of matrix secant and trust region methods see [9].

We denote the discretized unknown functions (f,θ) by $u\in\mathbb{R}^n$. Matrix secant iterates are generated by

$$u_{k+1}=u_k-(\lambda_k M_k)^{-1}\nabla J(u_k) \tag{20}$$

where $\lambda_k\in\mathbb{R}$ is a scaling factor for the step length and $M_k\in\mathbb{R}^{n\times n}$ is an approximation to $\nabla^2 J[f,\theta;\underline{d}]$ satisfying the quasi-Newton equation:

$$M_k(u_{k-1}-u_k)=\nabla J(u_{k-1})-\nabla J(u_k). \tag{21}$$

Equation (21)) is a system of n equations in n^2 unknowns, thus infinitely many solutions are possible. Broyden's update, symmetric rank one (SR1) matrices, and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) update are common choices for the matrix secant M_k. The BFGS update has been shown empirically to be superior in many cases. We will review the limited memory techniques for BFGS matrices, however similar methods for alternative updates are possible.

The BFGS approximation to the true Hessian is given by

$$M_k=M_{k-1}+\frac{y_k y_k^T}{y_k^T s_k}-\frac{M_{k-1}s_k s_k^T M_{k-1}}{s_k^T M_{k-1}s_k},\ k=1,2,\dots \tag{22}$$

where

$$y_k\equiv\nabla J(u_{k+1})-\nabla J(u_k),\qquad\qquad s_k\equiv u_{k+1}-u_k.$$

The BFGS approximation is symmetric positive definite as long as $s_k^T y_k>0$.

In [7] Byrd *et al* derive compact representations of the BFGS approximation. Just as with conjugate gradients, these representations allow one to compute the product $M_k^{-1}\nabla J(u_k)$ without actually forming the matrix M_k^{-1}. Let $S_k\equiv[s_{k-m},\dots,s_{k-1}]\in\mathbb{R}^{n\times m}$ denote a matrix of steps from previous iterations. Similarly we store previous gradient differences, $Y_k\equiv[y_{k-m},\dots,y_{k-1}]\in\mathbb{R}^{n\times m}$. Limited memory techniques amount to generating *at each iteration* the BFGS matrix from the m most recent of the pairs $\{y_i,s_i\}_{i=1}^{k-1}$ and the generating matrix $M_k^{(0)}$. Typically $m\in[5,10]$. The choice of $M_k^{(0)}$ that is often used is $M_k^{(0)}=\lambda_k I$ where I is the identity matrix and λ_k is some scaling (see [26]). With this generating matrix limited memory BFGS is equivalent to doing m steps of

conjugate gradients at each iteration. Using limited memory BFGS, denoted L-BFGS, it can be shown that the complexity of calculating (20) is on the order of $mn + m^3$. See [7] for details.

Before accepting the step to the next iterate one usually checks the accuracy of the quadratic approximation

$$\widetilde{J}(u_{k+1}) = J(u_k) + \nabla J(u_k)^T \cdot s_k + \frac{1}{2} s_k^T M_k s_k \tag{23}$$

against the true function value $J(u_{k+1})$. Usually what is computed is the ratio of the actual change in the function value between iterates u_k and u_{k+1} and the predicted change, *i.e.*

$$r(s_k) = \frac{\text{predicted change}_k}{\text{actual change}_k}. \tag{24}$$

If the ratio is below some tolerance, the quadratic model is constrained by a *trust region, i.e.* a ball around the current iterate u_k that contains the largest "trustworthy" step length to the next iterate given the unreliable quadratic model. The trust region subproblem with trust region radius Δ is given by

$$TR(\Delta) \quad \begin{array}{ll} minimize & \nabla J(u_k)^T s + \frac{1}{2} s^T M_k s. \\ & \|s\| \leq \Delta \end{array}$$

The Lagrangian of $TR(\Delta)$ yields the following unconstrained, *implicit* trust region subproblem

$$TR'(\mu) \quad \begin{array}{ll} minimize & \nabla J(u_k)^T s + \frac{1}{2} s^T \left(M_k + \mu I\right) s. \\ & s \in \mathbb{R}^n \end{array}$$

A solution, $s_*(\mu)$, to $TR'(\mu)$ corresponds to a solution to $TR(\Delta)$ with $\Delta = \|s_*(\mu)\|$. The larger μ the smaller the trust region radius Δ.

In [6] Burke and Wiegmann derive a compact representation of the inverse of the matrix $\mu I + M_k$ for solving the trust region subproblem. This inverse can be computed with the same computational complexity as the computation of M_k^{-1}. The basic assumption of the trust region step is that the step s_k is at least as effective as the step obtained by moving in the direction of steepest descent. Thus the strategy is "global" in the sense that far from the solution the method will perform at least as well as gradient descent while near a solution the second order information will accelerate convergence.

3.2 Algorithms and Implementation

As noted in [6], for the proper scaling, λ_k, the trust region is required only a small fraction of the time. This scaling is key to the success of the algorithm. There are many definitions for the optimal scaling λ_k [22]. We follow the scaling suggested by Shanno and Phua [26]

$$\lambda_k = \frac{y_{k-1}^T y_{k-1}}{s_{k-1}^T y_{k-1}}. \tag{25}$$

In our implementations we default to the unconstrained L-BFGS method, *i.e.* by default $\mu = 0$ in $TR'(\mu)$ at the beginning of each iteration. The trust region is invoked only if the ratio $r(s_k)$ falls below a given tolerance, indicating that the quadratic model (23) is not reliable, *i.e.* $r(s_k) < \text{tol}_{TR}$.

Numerical experiments indicate, however, that when a step does not give sufficient decrease in the objective value, or even causes an *increase* in the objective, it is worthwhile to keep that step direction and use it to update the L-BFGS matrix, even though the step is not taken. The rational behind this is that bad steps still contain curvature information that is useful in approximations to the Hessian of the objective.

Algorithm 3.1 (Limited Memory Trust Region BFGS) *For $k \geq 1$, given the vectors $u_{k-1} \in \mathbb{R}^n$, $\nabla J(u_{k-1})$, the scalars $J(u_{k-1})$, $\|\nabla J(u_{k-1})\|_2^2$, $\overline{m} \in \{1, 2, \ldots, n\}$, and the tolerance tol_{MS}:*

0. *Set $S_{k-2} = Y_{k-2} = [\,]$, the empty matrix, and $m = 0$. Use some line search method (*e.g. *line search with backtracking) to find a direction of descent to the point u_k. Calculate the vectors s_{k-1}, y_{k-1}, $\nabla J(u_k)$, and the scalar $\|\nabla J(u_k)\|_2^2$.*

1. *If $s_{k-1}^T y_{k-1} \leq 0$, goto step 0.; otherwise set $m = min\{m+1, \overline{m}\}$, $\lambda_k = y_{k-1}^T y_{k-1}/(s_{k-1}^T y_{k-1})$ and update S_{k-1}, Y_{k-1}, $Y_{k-1}^T Y_{k-1}$, and $S_{k-1}^T Y_{k-1}$ (see algorithm 3.1 of [7]).*

2. *If $r(s_k) < tol_{TR}$ restrict trust region Δ and solve the trust region subproblem (see [6]); otherwise calculate $\nabla J(u_{k+1})$, $\|\nabla J(u_{k+1})\|_2^2$, and update S_k, Y_k and the limited matrix BFGS matrix (see [7]).*

3. *If $\|\nabla J(u_{k+1})\|_2^2 < tol_{MS}$, end; otherwise calculate the actual change, the predicted change, $r(s_k)$ via (24), set $k = k+1$ and return to 1.*

3.3 Gradients

In the appendix we outline the derivation for the formal derivative of J_{LS} in the direction $\underline{d} = (\delta f, \delta\theta, \delta\varphi)$. For a regularization operator $\mathcal{N}$ that is independent of (f, θ) we have:

$$\nabla J_{LS}[f,\theta,\varphi](\underline{d}) = \left\langle \underline{d},\ 2Re\left[\sum_{j=1}^{m} \underline{g}_j\right]\right\rangle + \left\langle \underline{d},\ 2f \cdot r_{\mathcal{X}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\rangle \tag{26}$$

where

$$\underline{g}_j = \left(\mathcal{X} e^{-i(\theta+\phi_j)}\left[(\varphi^{\vee} r_j^{\wedge}) \star \left(\mathcal{X} f e^{i(\theta+\phi_j)}\right)\right],\ -i\mathcal{X} f e^{-i(\theta+\phi_j)}\left[(\varphi^{\vee} r_j^{\wedge}) \star \left(\mathcal{X} f e^{i(\theta+\phi_j)}\right)\right],\ \frac{1}{2}\tilde{\kappa}_j \star r_j\right) \tag{27}$$

and

$$r_j \equiv \left(|\kappa_j|^2 + \alpha\mathcal{N}\right) * \varphi - \psi_j, \qquad r_{\mathcal{X}} \equiv f^2 - \mathcal{X}, \qquad \tilde{\kappa}_j = |\kappa_j|^2 + \alpha\mathcal{N}. \tag{28}$$

Solving for φ in terms of f and θ via (15), the directional derivative $\nabla J_{LS}[f,\theta](\underline{d})$ for $\underline{d} = (\delta f, \delta\theta)$ takes the form

$$\nabla J_{LS}[f,\theta](\underline{d}) = \left\langle \underline{d},\ 2Re\left[\sum_{j=1}^{m} \tilde{\underline{g}}_j\right]\right\rangle + \left\langle \underline{d},\ 2f \cdot r_{\mathcal{X}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\rangle \tag{29}$$

where

$$\tilde{\underline{g}}_j \equiv \mathcal{X} e^{-i(\phi_j+\theta)} \left[\left(\varphi_*^\vee r_j^\wedge\right) \star \left(\mathcal{X} f e^{i(\theta+\phi_j)} \right) \right] \begin{pmatrix} 1 \\ -if \end{pmatrix}. \tag{30}$$

A Cartesian representation of the wavefront yields a simpler expression than (30). Let $w+iz = fe^{i\theta}$. The directional derivative of J_{LS} with respect to w and z is

$$\begin{aligned} \nabla J_{LS}[w,z](\underline{d}) &= \left\langle \underline{d},\ 2Re \left[\sum_{j=1}^{m} \mathcal{X} e^{-i\phi_j} \left[\left(\varphi_*^\vee r_j^\wedge\right) \star \left(\mathcal{X} e^{i\phi_j}(w+iz) \right) \right] \begin{pmatrix} 1 \\ -i \end{pmatrix} \right] \right\rangle \\ &\quad + \left\langle \underline{d},\ 2Re \left[(w+iz) \cdot r_{\mathcal{X}} \begin{pmatrix} 1 \\ -i \end{pmatrix} \right] \right\rangle. \end{aligned} \tag{31}$$

Note that the components of the gradient with respect to w and z are the real and imaginary parts of

$$2 \sum_{j=1}^{m} \mathcal{X} e^{-i\phi_j} \left[\left(\varphi_*^\vee r_j^\wedge\right) \star \left(\mathcal{X} e^{i\phi_j}(w+iz) \right) \right] + 2(w+iz) \cdot r_{\mathcal{X}}$$

respectively. Expression (31) is more efficient to calculate and does not suffer from 2π ambiguities in the phase.

The gradient of the set distance error objective J^{ϵ}_{sde}, also derived in the appendix, is formally

$$\begin{aligned} \nabla J^{\epsilon}_{sde}[f,\theta](\underline{d}) &= \left\langle \underline{d},\ Re \left[\sum_{j=1}^{m} \mathcal{X} e^{-i(\theta+\phi_j)} \left[\left(\frac{|\kappa_j|+2\epsilon}{(|\kappa_j|+\epsilon)^2} \right) \kappa_j \cdot r_j \right]^\vee \begin{pmatrix} 1 \\ -if \end{pmatrix} \right] \right\rangle \\ &\quad + \left\langle \underline{d},\ \left(\frac{f+2\epsilon}{(f+\epsilon)^2} \right) f \cdot r_{\mathcal{X}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle. \end{aligned} \tag{32}$$

where $r_j = \frac{|\kappa_j|^2}{|\kappa_j|+\epsilon} - \psi^{1/2}$ and $r_{\mathcal{X}} = \frac{f^2}{f+\epsilon} - \mathcal{X}$. In the limit as $\epsilon \to 0$ this expression formally becomes

$$\nabla J_{sde}[f,\theta](\underline{d}) = \left\langle \underline{d},\ Re \left[\sum_{j=1}^{m} \mathcal{X} e^{-i(\theta+\phi_j)} \left[\frac{\kappa_j}{|\kappa_j|} r_j \right]^\vee \begin{pmatrix} 1 \\ -if \end{pmatrix} \right] \ +\ r_{\mathcal{X}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle. \tag{33}$$

Note that if $\kappa_j(\underline{x}) > 0$ for all $\underline{x}$, then (33) is simply the sum of the *simultaneous* projections onto all the magnitude constraints. Thus, at least heuristically, a gradient descent method for this objective can be viewed as a particular case of a projection method. Convergence results for projection methods applied to this problem do not exist since the problem is non-convex. Convergence results for gradient descent algorithms, however, can be derived from standard results in the optimization literature.

4 Results

In this section we report the performance of the algorithm described in section (3.2) on simulated data. We discretize the functions above into a pixel basis. Thus for a 512×512 image the

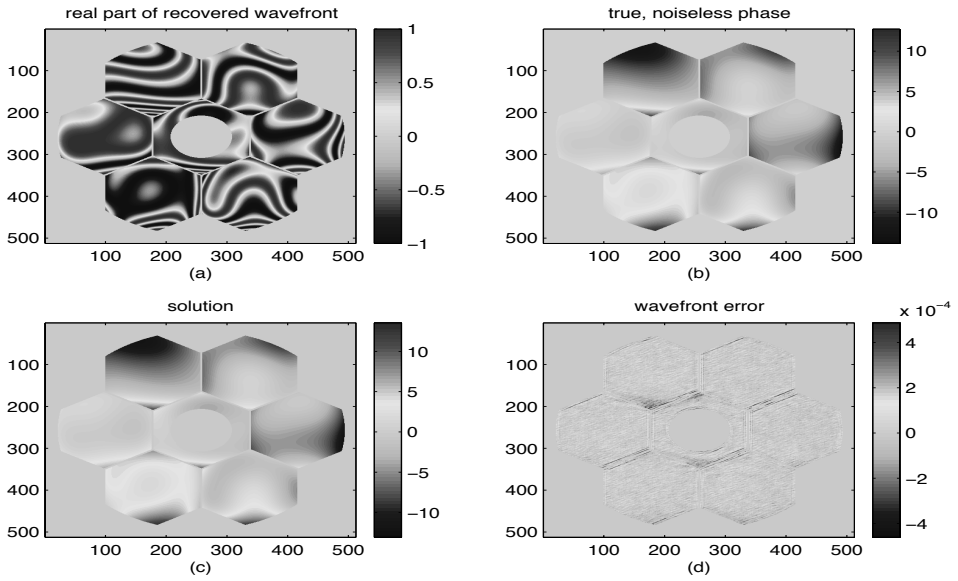

Figure 2: Seven paneled, segmented aperture with phase aberrations.

optimization problem ($\mathcal{P}_{LS}$) involves 786432 variables. In all simulations we used little or no regularization. We begin with the simplest problem, phase retrieval.

The first example shows the performance of the limited memory algorithm with the objective J^{ϵ}_{sde} against a common iterative transform algorithm called the Misell algorithm [5, 20]. The Misell algorithm projects between several diversity images in a sequential manner. The same information, two out of focus images and one in focus image, is available to the Misell algorithm that is available to the limited memory algorithm. We take the parameter ϵ in J^{ϵ}_{sde} to be machine precision, 10^{-16}. With this small regularization, a steepest descent algorithm applied to J^{ϵ}_{sde} is practically the same as a simultaneous, weighted projection algorithm. Given the connection between the gradient (33) and evenly weighted, simultaneous projections, the difference between the two algorithms is the weighting of the projections, and, in the case of limited memory, the use of previous iterations for acceleration. Figure (3a) shows the performance of the two algorithms on a single panel of the aperture shown in figure (2b). Figure (3b) shows the performance of the algorithms on the full segmented pupil with a resolution of 512×512 pixels. The limited memory algorithm has trouble finding a descent direction in the middle iterations, but converges rapidly toward the end. The Misell algorithm has trouble resolving the boundaries of the pupil. This test was terminated after 1000 iterations, at which point the Misell algorithm had made no further progress than it had in the first 200 iterations.

Figure (4) shows the performance of various implementations of the limited memory algorithm. In all examples the given test problem was the full segmented aperture shown in figure(2b) at 512×512 resolution. In figure(4a) limited memory with the objective function J_{LS} is compared to limited memory with the objective J^{ϵ}_{sde}. Both examples started from the same zero phase initial guess. The polar wavefront representation was used with f fixed. For random initial guesses the Cartesian representation outperformed all other formulations on average, though performance is comparable.

Figure (4b) illustrates the "flat" spot in the objective J_{sde}. In the middle iterations the gradient

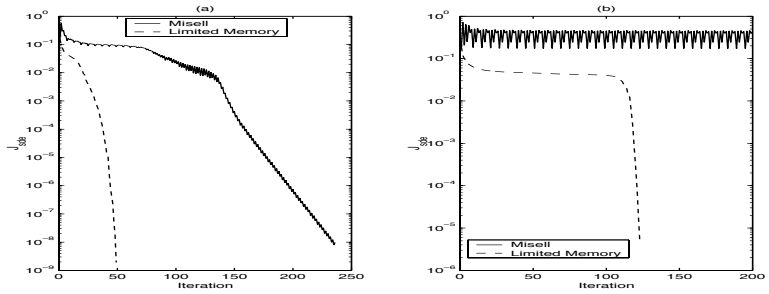

Figure 3: Comparison of the Misell algorithm with limited memory.

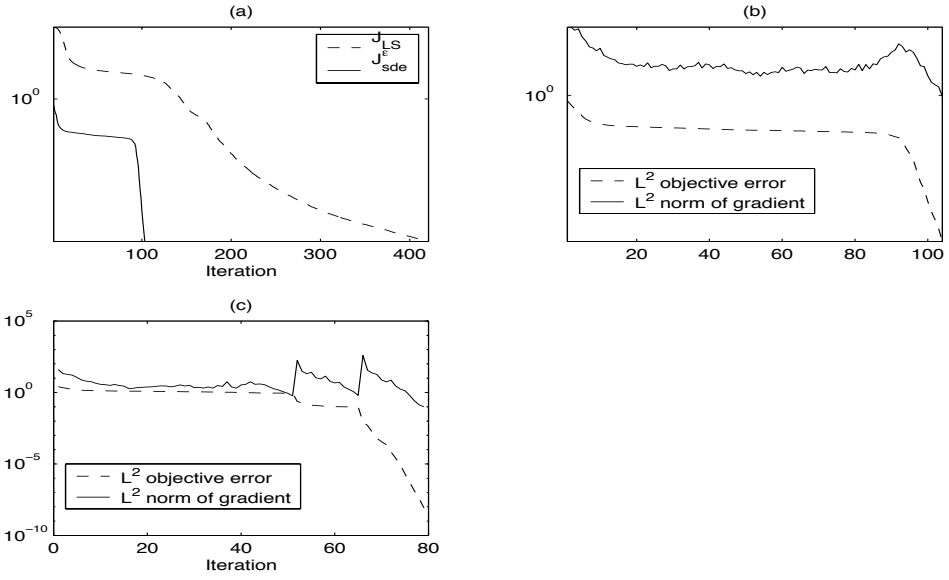

Figure 4: Comparison of limited memory implementations.

becomes very small, indicating that a strong direction of descent is not being found. Lacking better curvature information, we make use of a multi-resolution strategy that ensures that the "hard" work of the middle iterations is performed at low resolutions where the computations are inexpensive. The more expensive, higher resolution computations are only executed after the low resolution images have converged. This is shown in figure (4c).

The multi-resolution strategy is implemented by iterating only on the center pixels of the data image ψ_j and expanding the portion included pixels by factors of 2 to allow for fullest use of the fast Fourier transform. The solution at resolution 2^n is used as the first guess for the problem at resolution 2^{n+l}. The solution to the multiresolution example is shown in figure(2). In figure (4c) we started with the center 32×32 pixels of ψ_j. We used the solution at this resolution as the initial guess for the problem at resolution 128×128. Once a solution at 128×128 was found, we used this solution as the initial guess for the full resolution problem. The jumps in the gradient indicate where the resolution has changed. At resolutions of 128×128 and 512×512 less than 15 iterations were necessary to achieve the desired tolerance. At the tolerance shown in figure (4c) the maximum per

pixel wavefront error is less than 10^{-3} wavelengths, well below the tolerance of most operational systems (see figure (2d)). The computation time for the multi-resolution implementation from random initial guesses is on average 3.5 minutes using Matlab on a dec alpha workstation.

Figure (5) shows the numerical solution to the phase diversity problem. The method successfully converges to a diffraction limited solution, however, the rate of convergence is extremely slow, as is indicated by the performance with the J_{LS} objective in figure (4a). It has been noted that the BFGS approximation to second order information is not a competitive substitute for true second order information [27]. A method for incorporating second order information has been proposed by Dobson [10]. This is an active area of research.

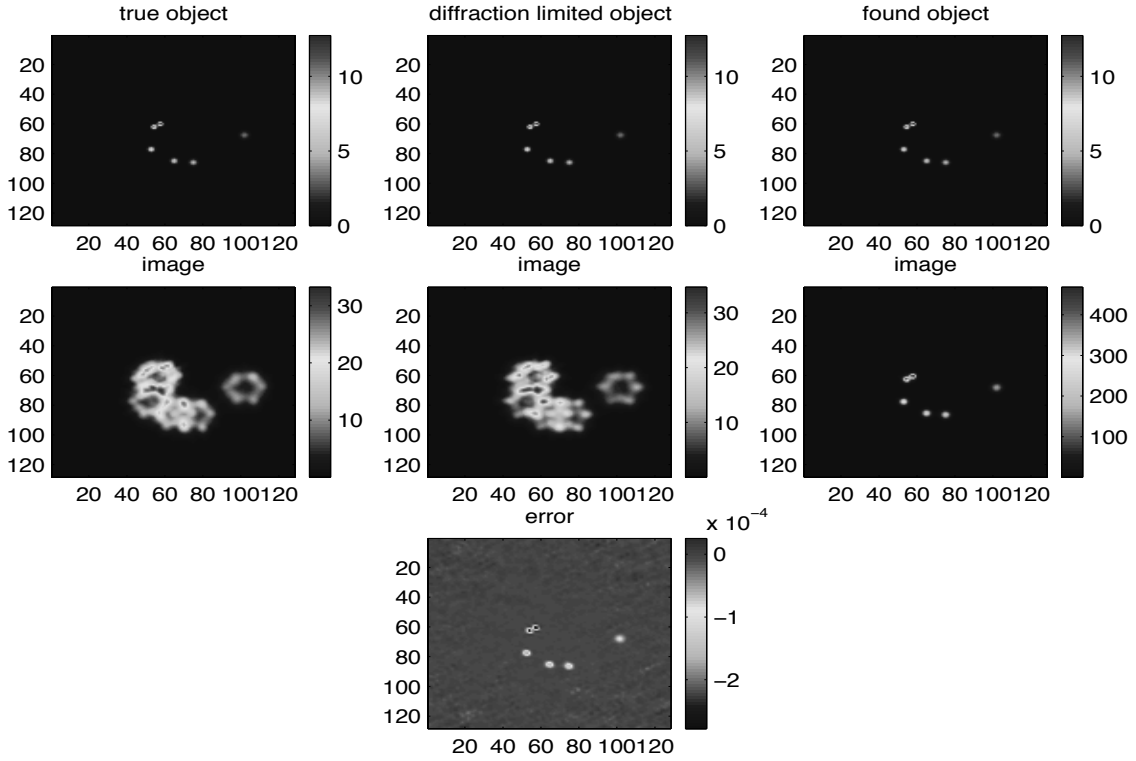

Figure 5: Phase diversity.

5 Conclusion

The phase retrieval and phase diversity problems are fundamentally ill-posed. Not only is the mathematical model ill-posed, but the underlying functions are not differentiable. A precise mathematical framework is necessary for a thorough treatment of these issues. The proposed regularization strategy for the phase diversity algorithm involves perturbing the original ill-posed problem to a nearby, numerically stable problem. Similarly, the non-differentiability of the set distance error used in phase retrieval is addressed by perturbing the objective to a nearby differentiable function which is asymptotically consistent with the original problem. Dobson [10] shows that the gradient of the J_{LS} objective is a compact operator. This introduces instabilities into the linearized problem which

necessitates further regularization. The non-parametric approach we take here allows researchers the freedom to choose a regularization that is physically justified. More work on regularization is necessary not only for theoretical results, but for faster algorithms.

Prototype test examples indicate that limited memory methods with trust regions are more robust than iterative transform methods and have faster rates of convergence. The performance measure is crucial to fast convergence. One critical advantage of the trust region limited memory techniques over iterative transform methods is in the careful scaling of the step size (25). This in conjunction with multi-resolution implementations decrease cpu time dramatically. Further speed up in run time is possible by processing the expressions for the gradients (31) and (32) in parallel. The limited memory BFGS algorithm with trust regions converges to diffraction limited solutions for the phase diversity problem, though convergence is slow. More work is needed to find efficient algorithms that make use of second order information.

A Gradient Calculations

The formulas for the gradients follow from tedious, though elementary vector calculus. Wherever possible we detail the operations in general, leaving the particulars of the specific operators for the reader. The reader will not, therefore, find complete explicit formulations in what follows. Where there is no chance for confusion, we omit the arguments from the operators.

- **Convolution and Correlation identities:** Define the convolution operator, $*$, by

$$u * v(\underline{x}) \equiv \int_{\mathbb{R}^n} u(\underline{x}')v(\underline{x} - \underline{x}')d\underline{x}'. \tag{34}$$

Define the correlation operator, $\star$, by

$$u \star v(\underline{x}) \equiv \int_{\mathbb{R}^n} u(\underline{x}')v(\underline{x} + \underline{x}')d\underline{x}'. \tag{35}$$

We define the inner product to be $\langle u, v\rangle \equiv \int_{\mathbb{R}^n} u(\underline{x})\overline{v}(\underline{x})d\underline{x}$. It is elementary to verify the following identities.

$$\begin{aligned}
u \star v &= u^{\wedge\wedge} * v; && (36)\\
u * v &= [u^{\wedge} \cdot v^{\wedge}]^{\vee} = [u^{\vee} \cdot v^{\vee}]^{\wedge}; && (37)\\
u \star v &= [u^{\vee} \cdot v^{\wedge}]^{\vee} = [u^{\wedge} \cdot v^{\vee}]^{\wedge}; && (38)\\
\langle u * v, h\rangle &= \langle v, \overline{u} \star h\rangle; && (39)\\
\langle u \star v, h\rangle &= \langle \overline{h} \star v, \overline{u}\rangle; && (40)\\
\text{if } h^{\vee} \text{ is real then } \langle v \star u, h\rangle &= \langle h \star v, \overline{u}\rangle. && (41)
\end{aligned}$$

- **Least squares gradient:** For $\underline{p} \in \left[L^2[\mathbb{R}^n, \mathbb{C}]\right]^m$, an m-vector of complex valued, Fréchet differentiable functionals $p_j : \mathcal{S} \to L^2[\mathbb{R}^n, \mathbb{C}]$ $(j = 1, \ldots, m)$, where $\mathcal{S} \subset \left[L^2[\mathbb{R}^n, \mathbb{R}]\right]^m$, the variational derivative of $\frac{1}{2}\left\|\underline{p}[\underline{u}]\right\|^2$ at $\underline{u} \in \mathcal{S}$ in the direction $\underline{d} \in \mathcal{S}$ is given by

$$\begin{aligned}
\left[\nabla \frac{1}{2}\|\underline{p}[\underline{u}]\|^2\right](\underline{d}) &= \sum_{j=1}^{m} Re\,\langle \nabla p_j[\underline{u}](\underline{d}),\ p_j[\underline{u}]\rangle\\
&= \sum_{j=1}^{m} \langle \underline{d},\ Re\,[\nabla p_j[\underline{u}]]^*\,(p_j[\underline{u}])\rangle\\
&= \left\langle \underline{d},\ Re \sum_{j=1}^{m} [\nabla p_j[\underline{u}]]^*\,(p_j[\underline{u}]) \right\rangle && (42)
\end{aligned}$$

- **Convolution adjoints**: Let $p,\ q \in L^2[\mathbb{R}^n, \mathbb{C}]$, and given $\mathcal{S} \subset \left[L^2[\mathbb{R}^n, \mathbb{R}]\right]^m$ let $g : \mathcal{S} \to L^2[\mathbb{R}^n, \mathbb{C}]$ be the Fréchet differentiable kernel of the convolution operator $\mathcal{G}$. The directional derivative of $\mathcal{G}$ at $[\underline{u}]$ in the direction $\underline{d}$, where $[\underline{u}],\ \underline{d} \in \mathcal{S}$ is given by $(\nabla\mathcal{G}(\underline{d}))\,p = [\nabla g[\underline{u}](\underline{d})] * p$. Applying Tonelli's theorem we define the adjoint with respect to $\underline{d}$ by

$$\begin{aligned}
\langle [\nabla\mathcal{G}[\underline{u}](\underline{d})]\,p,\ q\rangle &\equiv \langle \nabla g[\underline{u}](\underline{d}) * p,\ q\rangle\\
&= \langle \nabla g[\underline{u}](\underline{d}),\ \overline{p} \star q\rangle\\
&= \langle \underline{d},\ [\nabla g[\underline{u}]]^*\,(\overline{p} \star q)\rangle\,. && (43)
\end{aligned}$$

Thus we define

$$[\nabla \mathcal{G} p]^* (q) \equiv [\nabla g[\underline{u}]]^* (\overline{p} \star q) . \tag{44}$$

Note that while $(\nabla \mathcal{G}(\underline{d}))\, p \in L^2[\mathbb{R}^n, \mathbb{C}]$, the adjoint $[\nabla g[\underline{u}]]^* (\overline{p} \star q) \in \left[L^2[\mathbb{R}^n, \mathbb{C}]\right]^m$.

- **Multiplicative adjoints**: Let $p,\ q \in L^2[\mathbb{R}^n, \mathbb{C}]$, and let $g : \mathcal{S} \to L^2[\mathbb{R}^n, \mathbb{C}]$, where $\mathcal{S} \subset \left[L^2[\mathbb{R}^n, \mathbb{R}]\right]^m$, be a Fréchet differentiable "kernel" of the multiplication operator G. The directional derivative of the multiplication operator at $[\underline{u}]$ in the direction $\underline{d}$, where $[\underline{u}],\ \underline{d} \in \mathcal{S}$ is given by $(\nabla G(\underline{d}))\, p = [\nabla g[\underline{u}](\underline{d})]\, p$. The adjoint with respect to $\underline{d}$ is defined by

$$\begin{aligned} \langle [\nabla G[\underline{u}](\underline{d})]\, p,\ q \rangle &\equiv \langle [\nabla g[\underline{u}](\underline{d})]\, p,\ q \rangle \\ &= \langle \nabla g[\underline{u}](\underline{d}),\ \overline{p} q \rangle \\ &= \langle \underline{d},\ [\nabla g[\underline{u}]]^* (\overline{p} q) \rangle . \end{aligned} \tag{45}$$

Thus we define

$$[\nabla G p]^* (q) \equiv [\nabla g[\underline{u}]]^* (\overline{p} q) . \tag{46}$$

Again, note that while $(\nabla G(\underline{d}))\, p \in L^2[\mathbb{R}^n, \mathbb{C}]$, the adjoint $[\nabla g[\underline{u}]]^* (\overline{p} q) \in \left[L^n[\mathbb{R}^2, \mathbb{C}]\right]^m$.

- **The Lebesuge dominated convergence theorem and it's consequences:** The formulas for the gradient given by (26, (29), (32), and (33) assume that the generalized pupil function $f e^{i\theta}$ satisfies $|f(\underline{x})| \to 0$ as $|\underline{x}| \to \infty$. In the case of the Cartesian representation this condition is $|w(\underline{x}) + i z(\underline{x})| \to 0$ as $|\underline{x}| \to \infty$. The Lebesgue dominated convergence theorem then implies that the variational directional derrivative commutes with the Fourier transform, *i.e* that $\nabla_{w,z}[(w + iz)^\wedge](\underline{d}) = [\nabla_{w,z}[w + iz](\underline{d})]^\wedge$ where $\nabla_{w,z}[\cdot](\underline{d})$ is defined as the variation with respect to the functions w and z in the direction $\underline{d}$.

References

[1] M. J. Bastiaans. The wigner distribution function and hamilton's characteristics of a geometric-optical system. *Optics Communications*, 30(3):321–326, 1979.

[2] M. J. Bastiaans. Application of the wigner distribution function to partially coherent light. *J.Opt.Soc.Am.A*, 3(8):1227–1238, 1986.

[3] J. F. Benders. Partitioning proceedures fro solving mixed variables programming problems. *Numerische Mathematik*, 4:238–252, 1962.

[4] M. Born and E. Wolf. *Principles of Optics*. 6 edition, 1980.

[5] R. H. Boucher. Convergence of algorithms for phase retrieval from two intensity measurements. *SPIE*, 231:130–141, 1980.

[6] J. V. Burke and A. Wiegmann. Low-dimensional quasi-newton updating strategies for large-scale unconstrained optimization. Submitted to the SIAM Journal on Optimization, July 1996.

[7] R. H. Byrd, J. Nocedal, and R. B. Schnabel. Representations of quasi-newton matrices and their use in limited memory methods. *Math. Prog.*, 63:129–156, 1994.

[8] J. Dainty and J. Fienup. Phase retrieval and image reconstruction for astronomy. In H. Stark, editor, *Image Recovery: Theory and Application*. Academic Press, 1987.

[9] J. E. Dennis and R. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice Hall, 1983.

[10] D. C. Dobson. Phase reconstruction via nonlinear least squares. *Inverse Problems*, 8:541–557, 1992.

[11] J. Fienup. Phase retrieval algorithms: a comparison. *Appl.Opt.*, 21(5):2758–2768, 1982.

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

[13] R. Gonsalves. Phase retrieval and diversity in adaptive optics. *Opt.Eng*, 21(5):829–832, 1982.

[14] J. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, 1968.

[15] M. H. Hayes. *Signal Reconstruction from Phase or Magnitude*. PhD thesis, Massachusetts Institute of Technology, 1981.

[16] M. H. Hayes and A. V. Oppenheim. Signal reconstruction from phase or magnitude. *IEEE Trans. Acc. Sp. and Sig. Proc.*, 1980.

[17] A. Levi and H. Stark. Image restoration by the method of generalized projections with application to restoration from magnitude. *J.Opt.Soc.Am.A*, 1(9):932–943, 1984.

[18] D. R. Luke. PhD thesis, University of Washington, in preparation.

[19] D. R. Luke, J. V. Burke, R. Lyon, and T. Murphy. Non-parametric phase retrieval part ii: Numerical implementation and results. in preparation.

[20] D. L. Misell. An examination of an iterative method for the solution of the phase problem in optics and electron optics i. test calculations. *J.Phys.D.*, 1973.

[21] A.V. Oppenheim and R.W. Schafer. *Digital Signal Processing*. Prentice-Hall, 1975.

[22] S. S. Oren and E. Spedicato. Optimal conditioning of self-scaling variable metric algorithms. *Math.Prog.*, 10:70–90, 1976.

[23] R. Paxman, T. Schultz, and J. Fienup. Joint estimation of object and aberrations by using phase diversity. *J.Opt.Soc.Am.A*, 9(7):1072–1085, 1992.

[24] T. Quatieri and A.V. Oppenheim. Iterative techniques for minimum phase signal reconstruction from phase or magnitude. *IEEE Trans. on Acc.,Sp.and Sig. Proc.*, ASSP-29(6):1187–1193, Dec. 1981.

[25] R. T. Rockafellar and R. J. Wetts. *Variational Analysis*. Springer, 1998.

[26] D. F. Shanno and K. Phua. Matrix conditioning and nonlinear optimization. *Math. Prog.*, 14:149–160, 1978.

[27] C. R. Vogel, T. Chan, and R Plemmons. Fast algorithms for phase diversity-based blind deconvolution. Technical report, Department of Mathematical Sciences, Montana State University, 1999.

[28] D. C. Youla. Mathematical theory of image restoration by the method of convex projections. In H. Stark, editor, *Image Recovery: Theory and Applications*, pages 29–77. Academic Press, 1987.

[29] D. C. Youla and H. Web. Image restoration by the method of convex projections: Part i - theory. *IEEE Trans Med. Im.*, MI-1(2):81–94, Oct. 1982.