Absolute uniqueness of phase retrieval with random illumination

Absolute uniqueness of phase retrieval with random illumination

Albert Fannjiang

Department of Mathematics, UC Davis, CA 95616-8633, USA

E-mail: fannjiang@math.ucdavis.edu

Abstract

Random illumination is proposed to enforce absolute uniqueness and resolve
all types of ambiguity, trivial or nontrivial, in phase retrieval. Almost sure
irreducibility is proved for *any* complex-valued object whose support set has
rank $\geqslant 2$. While the new irreducibility result can be viewed as a probabilistic
version of the classical result by Bruck, Sodin and Hayes, it provides a novel
perspective and an effective method for phase retrieval. In particular, almost
sure uniqueness, up to a global phase, is proved for complex-valued objects
under general two-point conditions. Under a tight sector constraint absolute
uniqueness is proved to hold with probability exponentially close to unity
as the object sparsity increases. Under a magnitude constraint with random
amplitude illumination, uniqueness modulo global phase is proved to hold
with probability exponentially close to unity as object sparsity increases. For
general complex-valued objects without any constraint, almost sure uniqueness
up to global phase is established with two sets of Fourier magnitude data
under two independent illuminations. Numerical experiments suggest that
random illumination essentially alleviates most, if not all, numerical problems
commonly associated with the standard phasing algorithms.

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

1. Introduction

Phase retrieval is a fundamental problem in many areas of physical sciences such as X-ray
crystallography, astronomy, electron microscopy, coherent light microscopy, quantum state
tomography and remote sensing. Because of loss of the phase information a central question
of phase retrieval is the uniqueness of solution which is the focus of the present work.

Researchers in phase retrieval, however, have long settled with the notion of *relative
uniqueness* (i.e. irreducibility) for generic (i.e. random) objects, without a practical means for
deciding the reducibility of a given (i.e. deterministic) object, and searched for various ad hoc
strategies to circumvent problems with stagnation and error in reconstruction. The common
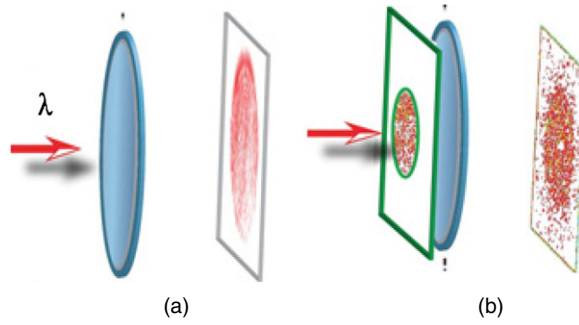
Figure 1. Illumination of a partially transparent object (the blue oval) with a deterministic (a) or random field λ created by a diffuser (b) followed by an intensity measurement of the diffraction pattern. In the case of wavefront reconstruction, the random modulator is placed at the exit pupil instead of the entrance pupil as in (b).

problem of stagnation may be due to the possibility of the iterative process to approach the object and its twin or shifted image, the support not tight enough or the boundary not sharp enough [13, 14, 18]. Besides the uniqueness issue, phase retrieval is also inherently nonconvex and many researchers have believed the lack of convexity in the Fourier magnitude constraint to be a main, if not the dominant, source of numerical problems with the standard phasing algorithms [4, 21, 30]. While there have been dazzling advances in applications of phase retrieval in the past decades [22], we still do not know just how much of the error and stagnation problems is attributable to to the lack of uniqueness or convexity.

We propose here to refocus on the issue of uniqueness as uniqueness is undoubtedly the first foundational issue of any inverse problem, including phase retrieval. Specifically we will first establish uniqueness in the absolute sense with *random* illumination under general, physically reasonable object constraints (figure 1) and secondly demonstrate that random illumination practically alleviates most numerical problems and drastically improves the quality of reconstruction.

To fix the idea, consider the discrete version of the phase retrieval problem: Let $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{Z}^d$ and $\mathbf{z} = (z_1, \ldots, z_d) \in \mathbb{C}^d$. Define the multi-index notation $\mathbf{z^n} = z_1^{n_1} z_2^{n_2} \cdots z_d^{n_d}$. Let $f(\mathbf{n})$ be a finite complex-valued function defined on $\mathbb{Z}^d$ vanishing outside the finite lattice

$$\mathcal{N} = \{\mathbf{0} \leqslant \mathbf{n} \leqslant \mathbf{N}\}$$

for $\mathbf{N} = (N_1, \ldots, N_d) \in \mathbb{N}^d$. We use the notation $\mathbf{m} \leqslant \mathbf{n}$ for $m_j \leqslant n_j, \forall j$. The *z*-transform of a finite sequence $f(\mathbf{n})$ is given by

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

The Fourier transform can be obtained from the *z*-transform as

$$F(\mathbf{w}) = F(\mathrm{e}^{\mathrm{i}2\pi w_1}, \ldots, \mathrm{e}^{\mathrm{i}2\pi w_d}) = \sum_{\mathbf{n}} f(\mathbf{n}) \mathrm{e}^{-\mathrm{i}2\pi \mathbf{n} \cdot \mathbf{w}}, \quad \mathbf{w} = (w_1, \ldots, w_d) \in [0, 1]^d$$

by some abuse of notation. The discrete phase retrieval problem is to determine $f(\mathbf{n})$ from the knowledge of the Fourier magnitude $|F(\mathbf{w})|, \forall \mathbf{w} \in [0, 1]^d$.

The question of uniqueness was partially answered in [3, 6, 16, 17] which says that in dimension two or higher and with the exception of a measure zero set of finite sequences phase retrieval has a unique solution up to the equivalence class of 'trivial associates'

(i.e. relative uniqueness). These trivial, but omnipresent, ambiguities include constant global phase,

$$f(\cdot) \longrightarrow e^{i\theta} f(\cdot), \quad \text{for some } \theta \in [0, 2\pi],$$

spatial shift

$$f(\cdot) \longrightarrow f(\cdot + \mathbf{m}), \quad \text{for some } \mathbf{m} \in \mathbb{Z}^d,$$

and conjugate inversion

$$f(\cdot) \longrightarrow f^*(\mathbf{N} - \cdot).$$

Conjugate inversion produces the so-called twin image.

This landmark uniqueness result, however, does not address the following issues. First, a *given* object array, there is no way of deciding *a priori* the irreducibility of the corresponding *z*-transform and the relative uniqueness of the phasing problem. Secondly, although visually no different from the true image the trivial associates (particularly spatial shift and conjugate inversion) nevertheless 'confuse' the standard numerical iterative processes and cause serious stagnation [13, 14, 24, 30].

In this paper, we study the notation of *absolute uniqueness*: if two finite objects f and g give rise to the same Fourier magnitude data, then $f = g$ unequivocally. More importantly, we present the approach of random (phase or amplitude) illumination to the absolute uniqueness of phase retrieval. The idea of random illumination is related to coded-aperture imaging whose utility in other imaging contexts than phase retrieval has been established experimentally [1, 2, 5, 7, 28, 31] as well as mathematically [8, 27].

Our basic tool is an improved version (theorem 2) of the irreducibility result of [16, 17] with, however, a completely different perspective and important practical implications. The main difference is that while the classical result [16, 17] works with generic (thus random) objects from a certain ensemble theorem 2 can deal with a *given, deterministic* object whose support has rank $\geqslant 2$. This improvement is achieved by endowing the probability measure on the ensemble of illuminations, which we can manipulate, instead of the space of objects, which we cannot control, as in the classical setting.

On the basis of almost sure irreducibility, the mere assumption that the phases or magnitudes of the object at two arbitrary points lie in a countable set enforces uniqueness, up to a global phase, in phase retrieval with a single random illumination (theorem 3). The absolute uniqueness can be enforced then by imposing the positivity constraint (corollary 1). For objects satisfying a tight sector condition, absolute uniqueness is valid with high probability depending on the object sparsity for either phase or amplitude illumination (theorem 4). For complex-valued objects under a magnitude constraint, uniqueness up to a global phase is valid with high probability (theorem 5). For general complex-valued objects, almost sure uniqueness, up to global phase, is proved for phasing with two independent illuminations (theorem 6).

The paper is organized as follows. In section 2 we discuss various sources of ambiguity. In section 3 we prove the almost sure irreducibility (theorem 2 and appendix). In section 4 we derive the uniqueness results (theorem 3–6 and corollary 1). We demonstrate phasing with random illumination in section 5. We conclude in section 6.

2. Sources of ambiguity

As commented before the phase retrieval problem does not have a unique solution. Nevertheless, the possible solutions are constrained as stated in the following theorem [16, 26].

Theorem 1. *Let the z-transform $F(\mathbf{z})$ of a finite complex-valued sequence $\{f(\mathbf{n})\}$ be given by*

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

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

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

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

To start, it is convenient to write

$$|F(\mathbf{w})|^2 = \sum_{\mathbf{n}=-\mathbf{N}}^{\mathbf{N}} \sum_{\mathbf{m} \in \mathcal{N}} f(\mathbf{m} + \mathbf{n}) f^*(\mathbf{m}) \, e^{-i2\pi \mathbf{n} \cdot \mathbf{w}}$$

$$= \sum_{\mathbf{n}=-\mathbf{N}}^{\mathbf{N}} \mathcal{C}_f(\mathbf{n}) \, e^{-i2\pi \mathbf{n} \cdot \mathbf{w}}, \tag{2}$$

where

$$\mathcal{C}_f(\mathbf{n}) = \sum_{\mathbf{m} \in \mathcal{N}} f(\mathbf{m} + \mathbf{n}) f^*(\mathbf{m}) \tag{3}$$

is the autocorrelation function of f. Note the symmetry $\mathcal{C}_f^*(\mathbf{n}) = \mathcal{C}_f(-\mathbf{n})$.

The theorem then follows straightforwardly from the equality between the autocorrelation functions of f and g, because $F(\mathbf{w})F^*(\mathbf{w}) = G(\mathbf{w})G^*(\mathbf{w})$, and the unique factorization of polynomials (see [26] for more details).

Remark 1. If the finite array $f(\mathbf{n})$ is known *a priori* to vanish outside the lattice $\mathcal{N}$, then by Shannon's sampling theorem for band-limited functions the sampling domain for $\mathbf{w}$ can be limited to the finite regular grid

$$\mathcal{M} = \left\{ (k_1, \ldots, k_d) : \forall j = 1, \ldots, d \& k_j = 0, \frac{1}{2N_j + 1}, \frac{2}{2N_j + 1}, \ldots, \frac{2N_j}{2N_j + 1} \right\} \tag{4}$$

since $|F(\mathbf{w})|^2$ is band-limited to the set $-\mathbf{N} \leqslant \mathbf{n} \leqslant \mathbf{N}$.

There are three sources of ambiguity. First, the linear phase term $\mathbf{z}^{-\mathbf{m}}$ in (1) remain undetermined because the autocorrelation operation destroys information about spatial shift. The unspecified constant phase θ is another source of ambiguity.

To understand the physical meaning of the operation

$$F(\mathbf{z}) \longrightarrow \mathbf{z}^{-\mathbf{N}} F^*(1/\mathbf{z}^*)$$

consider the case $d = 1$

$$z^{-N} F^*(1/z^*) = f^*(0) z^{-N} + f^*(1) z^{1-N} + \cdots + f^*(N),$$

which is the z-transform of the conjugate space-inversed array $\{f^*(N), f^*(N-1), \ldots, f^*(0)\}$. The same is true in multi-dimensions.

The subtlest form of ambiguity is caused by *partial* conjugate inversion on some, but not all, factors of a factorable object, with a reducible z-transform, without which the conjugate inversion, like spatial shift and global phase, is global in nature and considered 'trivial' in the literature (even though the twin image may have an opposite orientation).

In this paper, we consider both types, trivial and nontrivial, of ambiguity, as they both can degrade the performance of phasing schemes. Our main purpose is to show by rigorous analysis that with random illumination it is possible to eliminate all ambiguities at once.

3. Irreducibility

Random illumination amounts to replacing the original object $f(\mathbf{n})$ by

$$\tilde{f}(\mathbf{n}) = f(\mathbf{n})\lambda(\mathbf{n}), \tag{5}$$

where $\lambda(\mathbf{n})$, representing the incident field, is a *known* array of samples of random variables (r.v.s). The idea is to first modify the object by the encoding array $\lambda(\mathbf{n})$ so that phase retrieval has unique solution and then use the prior knowledge of λ to recover f.

Nearly independent random illumination can be produced by a diffuser placed near the object, cf figure 1. The illumination field can be randomly modulated in *phase* only with the use computer generated holograms [5], random phase plates [1, 28] and liquid crystal phase-only panels [7]. One of the best known amplitude masks is uniformly redundant array [11] and its variants [15]. The advantage of phase mask, compared to amplitude mask, is the lossless energy transmission of an incident wavefront through the mask. By placing either phase or amplitude mask at a distance from the object, one can create an illumination field modulated in both amplitude and phase in a way dependent on the distance [31].

Let $\lambda(\mathbf{n})$ be *continuous* r.v.s with respect to the Lebesgue measure on $\mathbb{S}^1$ (the unit circle), $\mathbb{R}$ or $\mathbb{C}$. The case of $\mathbb{S}^1$ can be facilitated by a random phase modulator with

$$\lambda(\mathbf{n}) = e^{i\phi(\mathbf{n})}, \tag{6}$$

where $\phi(\mathbf{n})$ are continuous r.v.s on $[0, 2\pi]$ while the case of $\mathbb{R}$ can be facilitated by a random amplitude modulator. The case of $\mathbb{C}$ involves simultaneously both phase and amplitude modulations. More generally, $\lambda(\mathbf{n})$ can be any continuous r.v. on a real algebraic variety $\mathcal{V}(\mathbf{n}) \subset \mathbb{C} \simeq \mathbb{R}^2$. For example $\mathbb{R}$ and $\mathbb{S}^1$ can be viewed as real projective varieties defined by the polynomial equations $y = 0$ and $x^2 + y^2 - 1 = 0$, respectively, on the complex plane identified as $\mathbb{R}^2$.

The support Σ of a polynomial $F(\mathbf{z})$ is the set of exponent vectors in $\mathbb{N}^d$ with nonzero coefficients. The rank of the support set is the dimension of its convex hull.

Theorem 2. *Let $\{f(\mathbf{n})\}$ be a finite complex-valued array whose support has rank $\geqslant 2$ and touches all the coordinate hyperplanes $\{n_j = 0 : j = 1, 2, \ldots, d\}$. Let $\{\lambda(\mathbf{n})\}$ be continuous r.v.s on nonzero real algebraic varieties $\{\mathcal{V}(\mathbf{n})\}$ in $\mathbb{C}(\simeq \mathbb{R}^2)$ with an absolutely continuous joint distribution with respect to the standard product measure on $\prod_{\mathbf{n} \in \Sigma} \mathcal{V}(\mathbf{n})$ where $\Sigma \subset \mathbb{N}^d$ is the support set of $\{f(\mathbf{n})\}$. Then the z-transform of $\tilde{f}(\mathbf{n}) = f(\mathbf{n})\lambda(\mathbf{n})$ is irreducible with probability 1.*

Remark 2. If the object support does not touch all the coordinate hyperplanes, then the the irreducibility holds true, *up to some monomial of* $\mathbf{z}$. In view of theorem 1 this is sufficient for our purpose.

Remark 3. The theorem does not hold if the rank-2 condition fails. For example, let $p(\mathbf{z})$ be any monomial and consider

$$F(\mathbf{z}) = \sum_j c_j p^j(\mathbf{z}), \tag{7}$$

which is reducible for any $c_j \in \mathbb{C}$, except when F is a monomial, by the fundamental theorem of algebra (of one variable). Another example is the *homogeneous* polynomials of a sum degree N

$$F(\mathbf{z}) = \sum_{i+j=N} c_{ij} z_1^i z_2^j, \tag{8}$$

which is factorable by, again, the fundamental theorem of algebra.

The proof of theorem 2 is given in the appendix.

Theorem 2 improves in several aspects on the classical result that the set of the reducible polynomials has zero measure in the space of multivariate polynomials with real-valued coefficients [16, 17]. The main improvement is that while the classical result works with generic (thus random) objects theorem 2 deals with *any deterministic* object with minimum (and necessary) conditions on its support set. By definition, deterministic objects belong to the measure zero set excluded in the classical setting of [16, 17]. It is both theoretically and practically important that theorem 2 places the probability measure on the ensemble of illuminations, which we can manipulate, instead of the space of objects, which we cannot control.

In the next section, we go further to show that with additional, but for all practical purposes sufficiently general, constraints on the values of the object, we can essentially remove all ambiguities with the only possible exception of global phase factor. This decisive step distinguishes our method from the standard approach.

4. Uniqueness

Without additional *a priori* knowledge on the object theorem 2, however, does not preclude the trivial ambiguities such as global phase, spatial shift and conjugate inversion. For example, we can produce another finite array $\{g(\mathbf{n})\}$ that yields the same measurement data by setting

$$g(\mathbf{n}) = e^{i\theta} f(\mathbf{n} + \mathbf{m})\lambda(\mathbf{n} + \mathbf{m})/\lambda(\mathbf{n}) \tag{9}$$

or

$$g(\mathbf{n}) = e^{i\theta} f^*(\mathbf{N} - \mathbf{n} + \mathbf{m})\lambda^*(\mathbf{N} - \mathbf{n} + \mathbf{m})/\lambda(\mathbf{n}) \tag{10}$$

for $\theta \in [0, 2\pi]$ and $\mathbf{m} \in \mathbb{Z}^d$. Expressions (9) and (10) are the remaining ambiguities to be addressed.

4.1. Two-point constraint

One important exception is the case of *real-valued* objects when the illumination is complex-valued (the case of $\mathbb{S}^1$ or $\mathbb{C}$). In this case, on the one hand (9) produces a complex-valued array with probability 1 unless $\mathbf{m} = 0, \theta = 0, \pi$ and, on the other hand, (10) is complex-valued with probability 1 regardless of $\mathbf{m}$. In this case, none of the trivial ambiguities can arise. Indeed, a stronger result is true depending on the nature of random illumination.

Theorem 3. *Suppose the object support has rank* $\geqslant 2$. *Suppose either of the following cases holds.*

(i) *The phases of the object* $\{f(\mathbf{n})\}$ *at two points, where f does not vanish, belong to a known countable subset of* $[0, 2\pi]$ *and* $\{\lambda(\mathbf{n})\}$ *are independent continuous r.v.s on real algebraic varieties in* $\mathbb{C}$ *such that their angles are continuously distributed on* $[0, 2\pi]$ *(e.g.,* $\mathbb{S}^1$ *or* $\mathbb{C}$*).*

(ii) *The amplitudes of the object* $\{f(\mathbf{n})\}$ *at two points, where f does not vanish, belong to a known measure zero subset of* $\mathbb{R}$ *and* $\{\lambda(\mathbf{n})\}$ *are independent continuous r.v.s on real algebraic varieties in* $\mathbb{C}$ *such that their magnitudes are continuously distributed on* $(0, \infty)$ *(e.g.* $\mathbb{R}$ *or* $\mathbb{C}$*).*

Then f is determined uniquely, up to a global phase, by the Fourier magnitude measurement on the lattice $\mathcal{M}$ *with probability 1.*

Remark 4. For the two-point constraint in case (i) to be convex, it is necessary for the constraint set to be a singleton, namely the phases of the object at two nonzero points must take on a single known value. On the other hand, the amplitude constraint in case (ii) can never be convex unless the set is a singleton and the object phases are the same at the two points.

Proof. By theorem 2 the z-transform of $\{\lambda(\mathbf{n})f(\mathbf{n})\}$ is irreducible with probability 1. We prove the theorem case by case.

Case (i). Suppose the phases of $f(\mathbf{n}_1)$ and $f(\mathbf{n}_2)$ belong to the countable set $\Theta \subset [0, 2\pi]$. Let us show the probability that the phase of $g(\mathbf{n})$ as given by (9) with $\mathbf{m} \neq 0$ takes on a value in Θ at two distinct points is zero.

Since $\lambda(\mathbf{n} + \mathbf{m})$, $\mathbf{m} \neq 0$, and the phases of $\lambda(\mathbf{n})$ are independent, continuous r.v.s on $[0, 2\pi]$, the phase of $g(\mathbf{n})$, $\forall \mathbf{n}$, is continuously distributed on $[0, 2\pi]$ for all θ.

Now suppose the phase of $g(\mathbf{n}_0)$ for some $\mathbf{n}_0$ lies in the set Θ. This implies that θ must belong to the countable set Θ' which is Θ shifted by the negative phase of $f(\mathbf{n}_0 + \mathbf{m})\lambda(\mathbf{n}_0 + \mathbf{m})/\lambda(\mathbf{n}_0)$. The phase of $g(\mathbf{n})$ at a different location $\mathbf{n} \neq \mathbf{n}_0$, however, almost surely does not take on any value in the set Θ for any fixed $\theta \in \Theta'$ unless $\mathbf{m} = 0$. Since a countable union of measure-zero sets has zero measure, the probability that the phases of g at two points lie in Θ is zero if $\mathbf{m} \neq 0$.

Likewise, $\lambda^*(\mathbf{N} - \mathbf{n} + \mathbf{m})/\lambda(\mathbf{n})$, $\forall \mathbf{m}$, has a random phase that is continuously distributed on $[0, 2\pi]$ and by the same argument the probability that the phases of g as given by (10) at two points lie in Θ is zero.

Case (ii). Suppose the amplitudes of $f(\mathbf{n}_1)$ and $f(\mathbf{n}_2)$ belong to the measure zero set $\mathcal{A}$. Since $\lambda(\mathbf{n} + \mathbf{m})$, $\mathbf{m} \neq 0$, and $\lambda(\mathbf{n})$ are independent and continuously distributed on $\mathbb{R}$ or $\mathbb{C}$, the amplitude of $g(\mathbf{n})$ as given by (9) is continuously distributed on $\mathbb{R}$ and hence the probability that the amplitude of $g(\mathbf{n})$ as given by (9) belongs to $\mathcal{A}$ at any $\mathbf{n}$ is zero.

Now consider $g(\mathbf{n})$ given by (10). Suppose that the amplitude of $g(\mathbf{n}_0)$ belongs to $\mathcal{A}$ at some $\mathbf{n}_0$. This is possible only for $\mathbf{n}_0 = (\mathbf{N} + \mathbf{m})/2$ in which case $g(\mathbf{n}_0) = e^{i\theta}f^*(\mathbf{n}_0)$. The amplitude of $g(\mathbf{n})$, $\mathbf{n} \neq \mathbf{n}_0$, has a continuous distribution on $\mathbb{R}$ and zero probability to lie in $\mathcal{A}$.

The global phase θ, however, cannot be determined uniquely in either case. $\square$

The global phase factor can be determined uniquely by additional constraint on the values of the object. For example, the following result follows immediately from theorem 3(i).

Corollary 1. *Suppose that $\{f(\mathbf{n})\}$ is real and nonnegative and its support has rank $\geqslant 2$. Suppose that $\{\lambda(\mathbf{n})\}$ are independent continuous r.v.s on real algebraic varieties in $\mathbb{C}$ such that their phases are continuously distributed on $[0, 2\pi]$ (e.g., $\mathbb{S}^1$ or $\mathbb{C}$). Then $\{f(\mathbf{n})\}$ can be determined absolutely uniquely with probability 1.*

Proof. With a real, positive object, the countable set for phase is the singleton $\{0\}$ and the global phase is uniquely fixed. $\square$

4.2. Sector constraint

More generally, we consider the *sector* constraint that the phases of $\{f(\mathbf{n})\}$ belong to $[a, b] \subset [0, 2\pi]$. For example, the class of complex-valued objects relevant to X-ray diffraction typically have nonnegative real and imaginary parts where the real part is the effective number of electrons coherently diffracting photons, and the imaginary part represents the attenuation [24]. For such objects, $[a, b] = [0, \pi/2]$.

Generalizing the argument for theorem 3 we can prove the following.

Theorem 4. *Suppose the object support has rank $\geqslant 2$. Let the finite object $\{f(\mathbf{n})\}$ satisfy the sector constraint that the phases of $\{f(\mathbf{n})\}$ belong to $[a, b] \subset [0, 2\pi]$. Let S be the sparsity (the number of nonzero elements) of the object.*

(i) Suppose $\{\lambda(\mathbf{n})\}$ are independent, identically distributed (i.i.d.) continuous r.v.s on real algebraic varieties in $\mathbb{C}$ such that their phases $\{\phi(\mathbf{n})\}$ are uniformly distributed on $[0, 2\pi]$ (e.g. the random phase illumination (6)). Then with probability at least $1 - |\mathcal{N}||b-a|^{[S/2]}(2\pi)^{-[S/2]}$ the object f is uniquely determined, up to a global phase, by the Fourier magnitude measurement. Here $[S/2]$ is the greatest integer at most $S/2$.

(ii) Consider the random amplitude illumination with i.i.d. continuous r.v.s $\{\lambda(\mathbf{n})\} \subset \mathbb{R}$ that are equally likely negative or positive, i.e. $\mathbb{P}\{\lambda(\mathbf{n}) > 0\} = \mathbb{P}\{\lambda(\mathbf{n}) < 0\} = 1/2, \forall \mathbf{n}$. Suppose $|b-a| \leqslant \pi$. Then with probability at least $1 - |\mathcal{N}|2^{-[(S-1)/2]}$ the object f is uniquely determined, up to a global phase, by the Fourier magnitude measurement.

In both cases, the global phase is uniquely determined if the sector $[a, b]$ is tight in the sense that no proper interval of $[a, b]$ contains all the phases of the object.

Proof.

Case (i). Consider first the expression (9) with any $\mathbf{m} \neq 0$ and the $[S/2]$ independently distributed r.v.s of $g(\mathbf{n})$ corresponding to $[S/2]$ nonoverlapping pairs of points $\{\mathbf{n}, \mathbf{n} + \mathbf{m}\}$. The probability for every such the phase of $g(\mathbf{n})$ to lie in the sector $[a, b]$ is $|b - a|/(2\pi)$ for any θ and hence the probability for all $g(\mathbf{n})$ with $\mathbf{m} \neq 0, \theta \neq 0$, to lie in the sector is at most $|b - a|^{[S/2]}(2\pi)^{-[S/2]}$. The union over $\mathbf{m} \neq 0$ of these events has probability at most $|\mathcal{N}||b - a|^{[S/2]}(2\pi)^{-[S/2]}$.

Likewise the probability for all $g(\mathbf{n})$ given by (10) to lie in the first quadrant for any $\mathbf{m}$ is at most $|\mathcal{N}||b - a|^{[S/2]}(2\pi)^{-[S/2]}$.

Case (ii). For (9) with any $\mathbf{m} \neq 0$ the $[S/2]$ independently distributed random variables $g(\mathbf{n})$ corresponding to $[S/2]$ nonoverlapping pairs of points $\{\mathbf{n}, \mathbf{n}+\mathbf{m}\}$, satisfy the sector constraint with probability at most $2^{-[S/2]}$ if $|b - a| \leqslant \pi$ for any θ. Hence the probability that all $g(\mathbf{n})$ with $\mathbf{m} \neq 0$ satisfy the sector constraint is at most $|\mathcal{N}|2^{-[S/2]}$.

For (10) with $\theta = 0$ and any $\mathbf{m}$, $g(\mathbf{n}_0) = f(\mathbf{n}_0)$ at $\mathbf{n}_0 = (\mathbf{N} + \mathbf{m})/2$ and hence $g(\mathbf{n}_0)$ lies in the first quadrant with probability 1. For $\mathbf{n} \neq \mathbf{n}_0$, $g(\mathbf{n})$ satisfies the sector constraint with probability $1/2$ if $|b - a| \leqslant \pi$. Now the $[(S - 1)/2]$ independently distributed r.v.s $g(\mathbf{n})$ corresponding to nonoverlapping pairs of points $\{\mathbf{n}, \mathbf{n} + \mathbf{m}\}, \mathbf{n} \neq \mathbf{n}_0$, satisfy the sector constraint with probability at most $2^{-[(S-1)/2]}$ if $|b - a| \leqslant \pi$. Hence the probability that all $g(\mathbf{n})$ given by (10) with arbitrary $\mathbf{m}$ satisfy the sector constraint is at most $|\mathcal{N}|2^{-[(S-1)/2]}$. □

4.3. Magnitude constraint

Likewise if the object satisfies a magnitude constraint then we can use random *amplitude* illumination to enforce uniqueness (up to a global phase).

Theorem 5. *Suppose that the object support has rank $\geqslant 2$. Suppose that K pixels of the complex-valued object f satisfy the magnitude constraint $0 < a \leqslant |f(\mathbf{n})| \leqslant b$ and that $\{\lambda(\mathbf{n})\}$ are i.i.d. continuous r.v.s on real algebraic varieties in $\mathbb{C}$ with $\mathbb{P}\{|\lambda(\mathbf{n})/\lambda(\mathbf{n}')| > b/a \text{ or } |\lambda(\mathbf{n})/\lambda(\mathbf{n}')| < a/b\} = 1 - p > 0$ for $\mathbf{n} \neq \mathbf{n}'$. Then the object f is determined uniquely, up to a global phase, by the Fourier magnitude data on $\mathcal{M}$, with probability at least $1 - |\mathcal{N}|p^{[(K-1)/2]}$.*
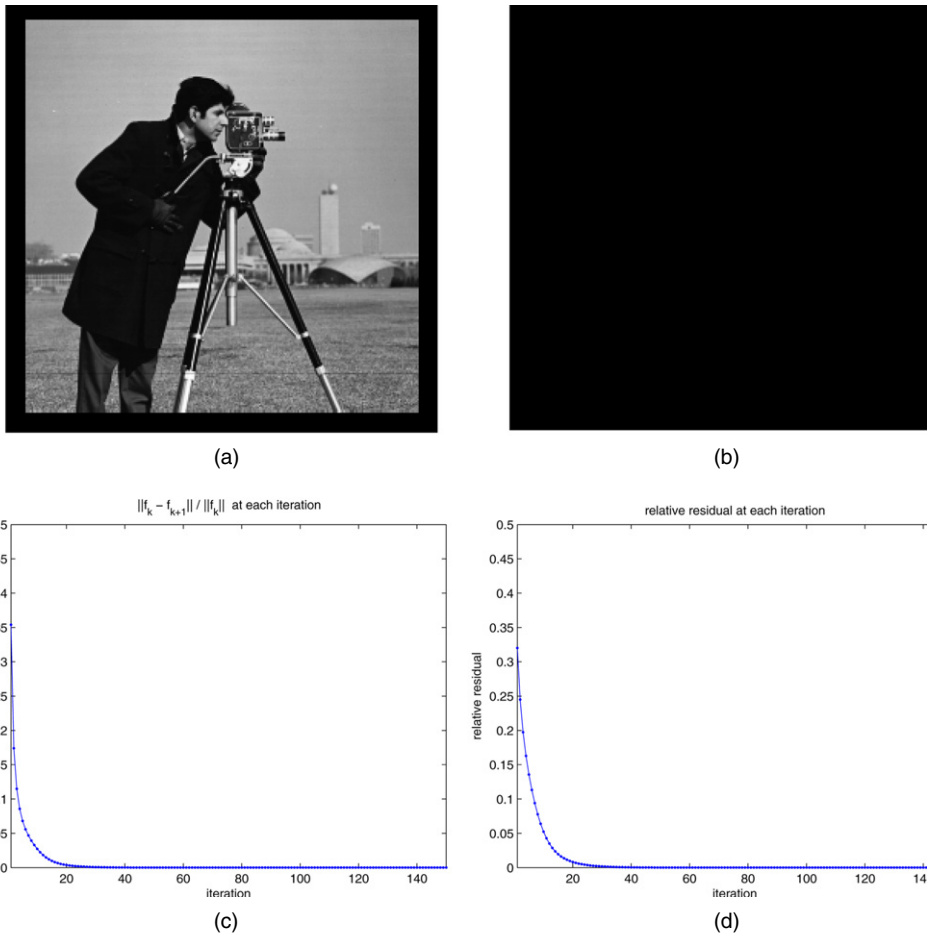
8

Figure 2. ER reconstruction with random phase illumination: (a) recovered image, (b) difference between the true and recovered images, (c) relative change $\|f_{k+1} - f_k\|/\|f_k\|$ and (d) relative residual $\||\widetilde{F}| - |\mathbf{\Phi}\Lambda f_k|\|/\|\widetilde{F}\|$ versus number of iterations.

Proof. The proof is similar to that for theorem 4(ii).

For (9) with any $\mathbf{m} \neq 0$ the $[K/2]$ independently distributed random variables $g(\mathbf{n})$ corresponding to $[K/2]$ nonoverlapping pairs of points $\{\mathbf{n}, \mathbf{n} + \mathbf{m}\}$ satisfy $0 < a \leqslant |g(\mathbf{n})| \leqslant b$ with probability less than $p^{-[K/2]}$ for any θ. Hence the probability that $g(\mathbf{n})$ with $\mathbf{m} \neq 0$ satisfy the magnitude constraint at K or more points is at most $|\mathcal{N}|p^{-[K/2]}$.

For (10) with any $\mathbf{m}$, $|g(\mathbf{n}_0)| = |f(\mathbf{n}_0)|$ at $\mathbf{n}_0 = (\mathbf{N} + \mathbf{m})/2$ and hence $g(\mathbf{n}_0)$ satisfies the magnitude constraint with probability 1. For $\mathbf{n} \neq \mathbf{n}_0$, there is at most probability p for $g(\mathbf{n})$ to satisfy the magnitude constraint. By independence, the $[(K-1)/2]$ independently distributed r.v.s $g(\mathbf{n})$ corresponding to nonoverlapping pairs of points $\{\mathbf{n}, \mathbf{n} + \mathbf{m}\}, \mathbf{n} \neq \mathbf{n}_0$, satisfy the magnitude constraint with probability at most $p^{[(K-1)/2]}$. Hence the probability that $g(\mathbf{n})$ given by (10) with arbitrary $\mathbf{m}$ satisfy the magnitude constraint at K or more points is at most $|\mathcal{N}|p^{[(K-1)/2]}$.

The global phase factor is clearly undetermined. $\qquad\square$

As in theorem 3 case (ii) the magnitude constraint here, however, is not convex.
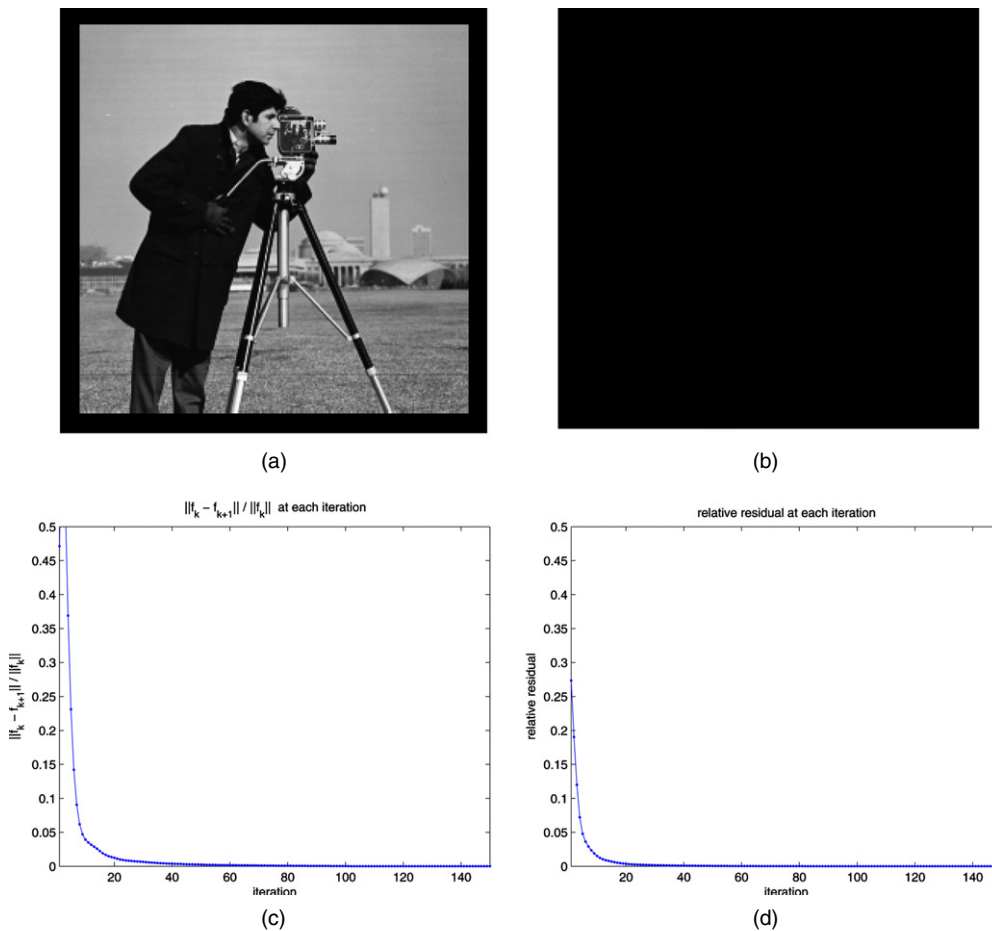
(a)

(b)

(c)

(d)

Figure 3. HIO reconstruction with random phase illumination: (a) recovered image, (b) difference between the true and recovered images, (c) relative change and (d) relative residual versus number of iterations. The final 50 iterations are ER.

4.4. Complex objects without constraint

For general complex-valued objects without any constraint, we consider two sets of Fourier magnitude data produced with two independent random illuminations and obtain almost sure uniqueness modulo global phase.

Theorem 6. *Let* $\{f(\mathbf{n})\}$ *be a finite complex-valued array whose support has rank* $\geqslant 2$. *Let* $\{\lambda_1(\mathbf{n})\}$ *and* $\{\lambda_2(\mathbf{n})\}$ *be two independent arrays of r.v.s satisfying the assumptions in theorem 2.*

Then with probability 1 $f(\mathbf{n})$ *is uniquely determined, up to a global phase, by the Fourier magnitude measurements on* $\mathcal{M}$ *with two illuminations* λ_1 *and* λ_2.

If the second illumination λ_2 *is deterministic and results in an irreducible z-transform while* λ_1 *is random as above, then the same conclusion holds.*
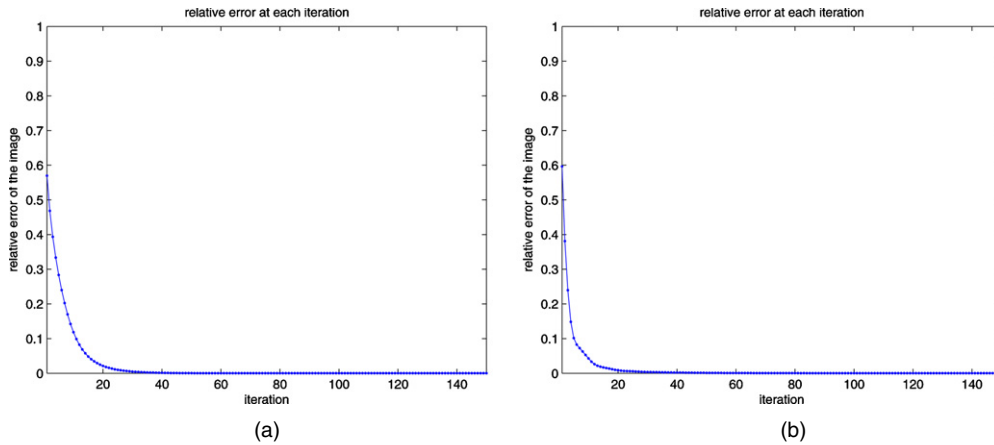
Figure 4. Relative error $\|f_k - f\|/\|f\|$ with (a) ER and (b) HIO versus number of iterations.

Proof. Let $g(\mathbf{n})$ be another array that vanishes outside $\mathcal{N}$ and produces the same data. By theoremes 1 and 2 and remark 1

$$g(\mathbf{n}) = \begin{cases} \mathrm{e}^{\mathrm{i}\theta_i} f(\mathbf{n} + \mathbf{m}_i)\lambda_i(\mathbf{n} + \mathbf{m}_i)/\lambda_i(\mathbf{n}) \\ \mathrm{e}^{\mathrm{i}\theta_i} f^*(\mathbf{N} - \mathbf{n} + \mathbf{m}_i)\lambda_i^*(\mathbf{N} - \mathbf{n} + \mathbf{m}_i)/\lambda_i(\mathbf{n}), \end{cases} \tag{11}$$

for some $\mathbf{m}_i \in \mathbb{Z}^d$, $\theta_i \in \mathbb{R}$, $i = 1, 2$.

Four scenarios of ambiguity exist but because of the independence of $\lambda_1(\mathbf{n})$, $\lambda_2(\mathbf{n})$ none can arise.

First of all, if

$$g(\mathbf{n}) = \mathrm{e}^{\mathrm{i}\theta_i} f(\mathbf{n} + \mathbf{m}_i)\lambda_i(\mathbf{n} + \mathbf{m}_i)/\lambda_i(\mathbf{n}), \quad i = 1, 2,$$

then

$$\mathrm{e}^{\mathrm{i}\theta_1} f(\mathbf{n} + \mathbf{m}_1)\lambda_1(\mathbf{n} + \mathbf{m}_1)/\lambda_1(\mathbf{n}) = \mathrm{e}^{\mathrm{i}\theta_2} f(\mathbf{n} + \mathbf{m}_2)\lambda_2(\mathbf{n} + \mathbf{m}_2)/\lambda_2(\mathbf{n}).$$

This almost surely cannot occur unless $\mathbf{m}_1 = \mathbf{m}_2 = \mathbf{0}$, $\theta_1 = \theta_2$ in which case g equals f up to a global phase factor.

The other possibilities can be similarly ruled out:

$$\begin{aligned} g(\mathbf{n}) &= \mathrm{e}^{\mathrm{i}\theta_1} f(\mathbf{n} + \mathbf{m}_1)\lambda_1(\mathbf{n} + \mathbf{m}_1)/\lambda_1(\mathbf{n}) \\ &= \mathrm{e}^{\mathrm{i}\theta_2} f^*(\mathbf{N} - \mathbf{n} + \mathbf{m}_2)\lambda_2^*(\mathbf{N} - \mathbf{n} + \mathbf{m}_2)/\lambda_2(\mathbf{n}) \end{aligned}$$

and

$$g(\mathbf{n}) = \mathrm{e}^{\mathrm{i}\theta_i} f^*(\mathbf{N} - \mathbf{n} + \mathbf{m}_i)\lambda_i^*(\mathbf{N} - \mathbf{n} + \mathbf{m}_i)/\lambda_i(\mathbf{n}), \quad i = 1, 2, \tag{12}$$

for any $\mathbf{m}_i, \theta_i, i = 1, 2$.

The same argument above applies to the case of deterministic λ_2 if the resulting z-transform is irreducible. □

5. Numerical examples

Our previous numerical study [9] and the following numerical examples give a glimpse of how the quality and efficiency of reconstruction can be improved by random illumination.

We test the case of random *phase* illumination on a real, positive 269×269 image consisting of the original 256×256 Cameraman in the middle, surrounded by a black margin
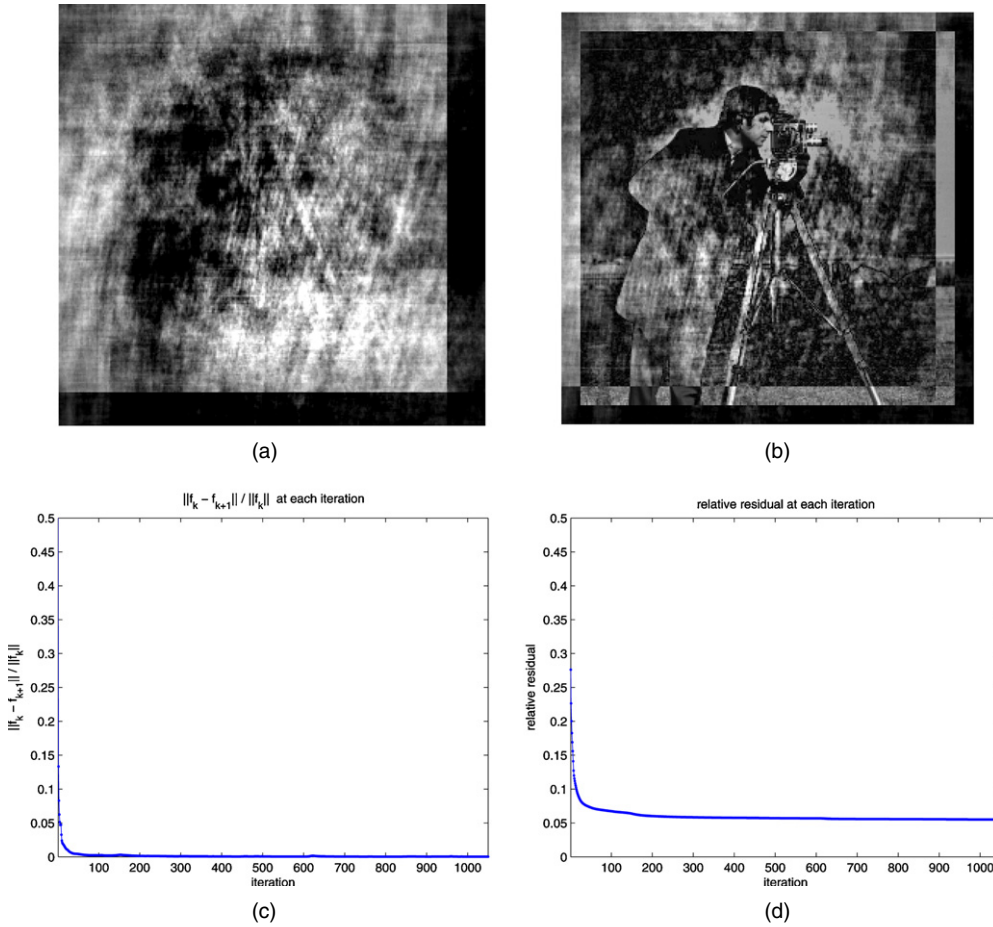
Figure 5. ER reconstruction with uniform illumination: (a) recovered image, (b) difference between the true and recovered images, (c) relative change and (d) relative residual versus number of iterations.

(zero padding) of 13 pixels in width (figure 2(a)). We synthesize and sample the Fourier magnitudes at the Nyquist rate (remark 1) and implement the standard Error Reduction (ER) and Hybrid–Input–Output (HIO) algorithms in the framework of the oversampling method [23, 24]. By corollary 1, absolute uniqueness holds with a random phase illumination.

Let $\boldsymbol{\Phi}$ and Λ be the Fourier transform and the diagonal matrix $\text{diag}[\lambda(\mathbf{n})]$ representing the illumination. For the uniform illumination $\Lambda = \mathbf{I}$. The ER and HIO algorithms are described below.

ER algorithm

Input: Fourier magnitude data $\{\widetilde{F}(\mathbf{w})\}$, initial guess f_0.
Iterations:

- Update Fourier phase: $G_k = \boldsymbol{\Phi}\Lambda f_k = |G_k(\mathbf{w})|e^{i\theta_k(\mathbf{w})}$.
- Update Fourier magnitude: $G'_k(\mathbf{w}) = |\widetilde{F}(\mathbf{w})|e^{i\theta_k(\mathbf{w})}$.
- Impose object constraint $f_{k+1}(\mathbf{n}) = \begin{cases} f'_k(\mathbf{n}) \text{ if } f'_k(\mathbf{n}) = \Lambda^{-1}\boldsymbol{\Phi}^* G'_k(\mathbf{n}) \geqslant 0 \\ 0 \text{ otherwise} \end{cases}$.
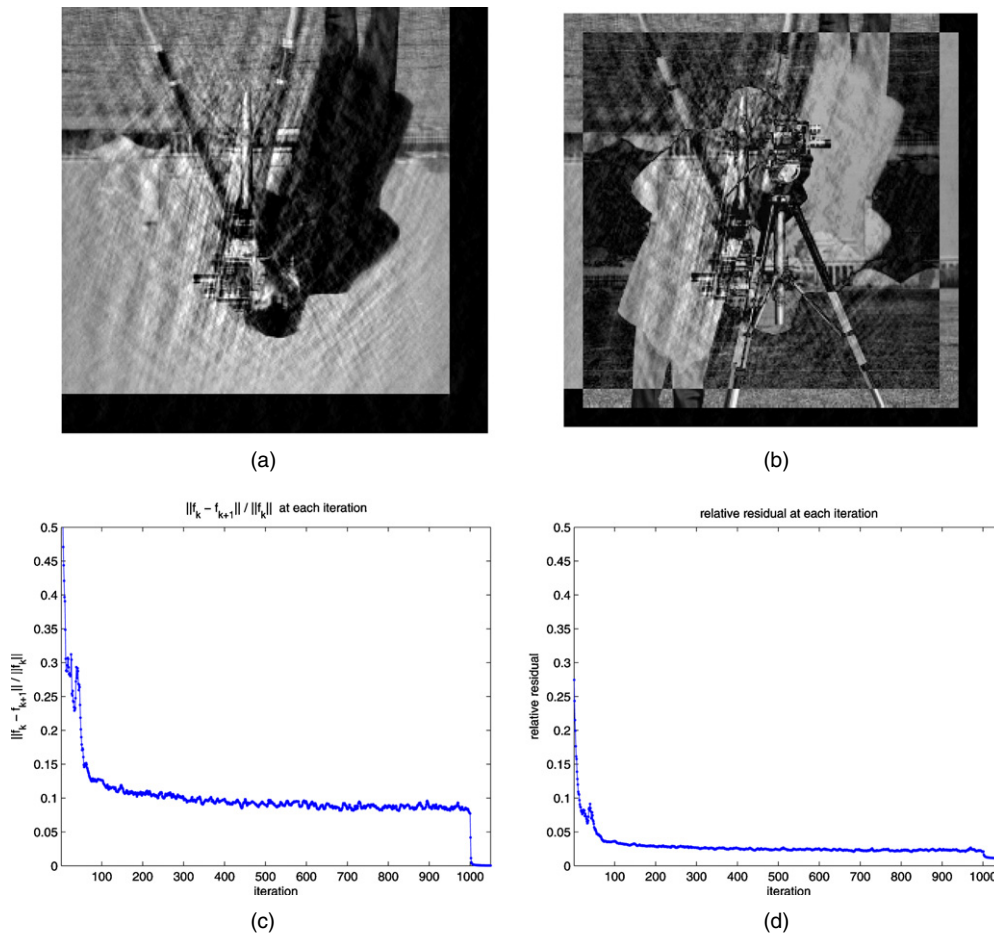

Figure 6. HIO reconstruction: (a) recovered image, (b) difference between the true and recovered images, (c) relative change and (d) relative residual versus number of iterations. The final 50 iterations are ER, causing a dip in the relative change and residual.

In the original version of HIO [12], the hard thresholding is replaced by

$$f_{k+1}(\mathbf{n}) = \begin{cases} f'_k(\mathbf{n}) & \text{if } \mathrm{f}'_k(\mathbf{n}) = \Lambda^{-1}\boldsymbol{\Phi}^*\mathrm{G}'_k(\mathbf{n}) \geqslant 0, \\ f_k(\mathbf{n}) - \beta f'_k(\mathbf{n}) & \text{otherwise} \end{cases},$$

where the feedback parameter $\beta = 0.9$ is used in the simulations. ER has the desirable property that the residual $\||F| - |G_k|\|$ is reduced after each iteration under either uniform or random phase illumination [12, 10]. When absolute uniqueness holds, a vanishing residual then implies a vanishing reconstruction error.

Figure 2 shows the results of ER reconstruction with random phase illumination. The ER iteration converges to the true image after 40 iterations. For HIO reconstruction, we apply 50 ER iterations after 100 HIO iterations as suggested in [21]. HIO has essentially the same performance as ER (figures 3(a) and (b)). The relative residual curve, figure 3(d), and the relative error curve, figure 4, however, indicate a small improvement by HIO. The close proximity between the vanishing residual curve and the vanishing error curve for ER and HIO reflects the absolute uniqueness under random illumination.
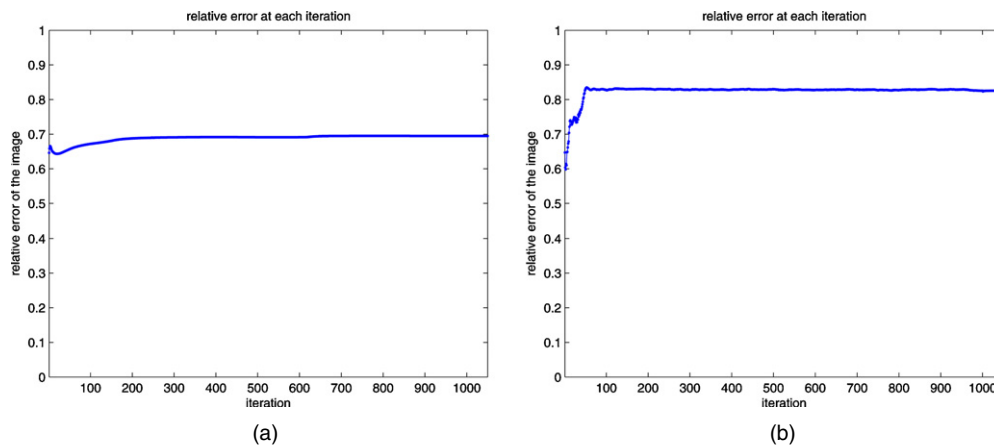
Figure 7. Relative error with (a) ER and (b) HIO versus number of iterations.

With uniform illumination, ER produces a poor result (figure 5(a)), resulting a 72.8% error (figure 5(b)) after more than 1000 iterations. The relative change curve, figure 5(c), indicates stagnation or convergence to a fixed point after 100 iterations and the relative residual plot, figure 5(d), shows non-convergence to the true image. For HIO reconstruction, we augment it with 50 ER iterations at the end of 1000 HIO iterations. While HIO improves the performance of ER but still leads to a shifted, inverted image which is also severely distorted (figures 6(a) and (b)). The ripples and stripes in figure 6(a) are a well-known artifact of HIO reconstruction [12, 14]. As expected, HIO reduces the residual and does not stagnate as much as ER (figures 6(c) and (d)) but its error is greater than that of ER due to the interferences from shifted and twin images present under the uniform illumination (figure 7).

To summarize, under a random phase illumination, the problems of stagnation and error disappear and phasing with ER/HIO achieves accurate, high-quality recovery. These experiments confirm our belief that a central barrier to stable and accurate phasing by the standard methods is the lack of absolute uniqueness.

6. Conclusions

In conclusion, we have proposed random illumination to address the uniqueness problem of phase retrieval. For general random illumination we have proved almost sure irreducibility for *any* complex-valued object whose support has rank $\geqslant 2$ (theorem 2). We have proved the almost sure uniqueness, up to a global phase, under the two-point assumption (theorem 3). The absolute uniqueness is then enforced by the positivity constraint (corollary 1). Under the tight sector constraint, we have proved the absolute uniqueness with probability exponentially close to unity as the object sparsity increases (theorem 4). Under the magnitude constraint, we have proved uniqueness up to a global phase with probability exponentially close to unity (theorem 5). For general complex-valued objects without any constraint, we have established almost sure uniqueness modulo global phase with two independent illuminations (theorem 6).

Numerical experiments reveal that phasing with random illumination drastically reduces the reconstruction error, the number of Fourier magnitude data and removes the stagnation problem commonly associated with the ER and HIO algorithms. Enforcement of absolute

uniqueness therefore appears to have a profound effect on the performance of the standard phasing algorithms.

Systematic and detailed study of phasing in the presence of (additive or multiplicative) noise with low-resolution random illuminations and sub-Nyquist sampling rates will be presented in the forthcoming paper [10].

Acknowledgments

Appendix. Proof of theorem 2

Our argument is based on [19, 25] and can be extended to the case of more than two independent variables. For simplicity of notation, we present the proof for the case of two independent variables.

Proof. First we state an elementary result from algebraic geometry (see, e.g., [32], page 65).

Proposition 1. *If a homogeneous polynomial* $P(z_0, z_1, z_2)$ *of (total) degree* $\delta \geqslant 2$ *is irreducible, then* $P^\flat(z_1, z_2) \equiv P(1, z_1, z_2)$ *is also irreducible with degree* δ.

For a polynomial $Q(z_1, z_2)$ of degree δ, the expression

$$Q^\sharp(z_0, z_1, z_2) = z_0^\delta Q\left(\frac{z_1}{z_0}, \frac{z_2}{z_0}\right) \tag{A.1}$$

defines a homogeneous polynomial of degree δ with the property $Q(z_1, z_2) = Q^\sharp(1, z_1, z_2)$. The process from Q to $Q^\sharp$ is called homogenization while the reverse process is called dehomogenization. Homogenization, in conjunction with proposition 1, is a useful tool for studying the question of irreducibility.

Let $\Sigma \subset \mathbb{N}^2$ be a given support set satisfying the assumptions of theorem 2. We now show that almost all homogeneous polynomials of 3 variables with the support

$$\Sigma^\sharp = \{(\delta - n_1 - n_2, n_1, n_2) : \mathbf{n} = (n_1, n_2) \in \Sigma\} \tag{A.2}$$

are irreducible.

We represent the set of all homogeneous polynomials of degree δ by the projective space $\mathbb{P}^{\nu_\delta}$ of dimension $\nu_\delta = \binom{2+\delta}{\delta} - 1$. Each homogeneous coordinate of $\mathbb{P}^{\nu_\delta}$ represents a monomial of degree δ. The homogeneous polynomials supported on $\Sigma^\sharp$ are represented by the projective subspace

$$\mathbb{X} = \{P \in \mathbb{P}^{\nu_\delta} : p(\mathbf{n}) = 0, \forall \mathbf{n} \notin \Sigma^\sharp\},$$

where $\{p(\mathbf{n})\}$ are the coefficients of P. Clearly $\mathbb{X}$ is isomorphic to the projective space $\mathbb{P}^s$. Let $\mathbb{Y} \subset \mathbb{X}$ denote the set of reducible homogeneous polynomials supported on $\Sigma^\sharp$. We claim (cf [29], page 47)

Proposition 2. $\mathbb{Y}$ *is a closed subset of* $\mathbb{X}$ *in the Zariski topology.*

15

A subset of a projective space is closed in the Zariski topology if and only if it is an algebraic variety, i.e. the common zero set

$$\{F_1 = F_2 = \cdots = F_m = 0\}$$

of a finite number of homogeneous polynomials $F_1, \ldots, F_m$ of the homogeneous coordinates of the projective space. The Zariski topology is much cruder than the metric topology. Indeed, a Zariski closed set is either the whole space or a measure-zero, nowhere-dense closed set in the metric topology as stated in the following (see, e.g., [20], page 115).

Proposition 3. *Any Zariski closed proper subset of a (real or complex) projective variety has measure zero with respect to the standard measure on the projective variety.*

Proof of proposition 2. Let the projective spaces $\mathbb{P}^{\nu_j}$ and $\mathbb{P}^{\nu_{\delta-j}}$ represent the homogeneous polynomials of degree j and $\delta - j$, respectively, where $\nu_j = \binom{2+j}{j} - 1$ and $\nu_{\delta-j} = \binom{2+\delta-j}{\delta-j} - 1$. Let $\mathbb{Y}_j \subset \mathbb{Y}$ be the set of points corresponding to polynomials supported on $\Sigma^\sharp$ that split into factors of degree j and $\delta - j$. Clearly $\mathbb{Y} = \cup_{j=1}^{\delta-1} \mathbb{Y}_j$ and we need only prove that each $\mathbb{Y}_j$ is Zariski closed.

Now the multiplication of two polynomials of degree j and $\delta - j$ determines a regular (i.e. polynomial) mapping

$$\Phi : \mathbb{P}^{\nu_j} \times \mathbb{P}^{\nu_{\delta-j}} \longrightarrow \mathbb{P}^{\nu_\delta}$$

in the following way. Let $G(\mathbf{z})$ and $H(\mathbf{z})$ be homogeneous polynomials of degrees j and $\delta - j$, respectively. Let $\{g(\mathbf{n})\}$ and $\{h(\mathbf{n})\}$ be the coefficients of G and H, respectively. Then the coefficients of the image point $\Phi(G, H)$ are given by

$$\left\{ \sum_{|\mathbf{n}|=\delta-j} g(\mathbf{m} - \mathbf{n})h(\mathbf{n}) : |\mathbf{m}| = m_0 + m_1 + m_2 = \delta \right\}. \tag{A.3}$$

In other words Φ is bilinear in $\{g(\mathbf{n})\}$ and $\{h(\mathbf{n})\}$ and thus is regular. Clearly we have

$$\mathbb{Y}_j = \Phi(\mathbb{P}^{\nu_j} \times \mathbb{P}^{\nu_{\delta-j}}) \cap \mathbb{X}.$$

Since the product of projective spaces is a projective variety and the image of a projective variety under a regular mapping is Zariski closed [29], $\mathbb{Y}_j$ is a Zariski closed subset of $\mathbb{X}$. $\quad\square$

Let $\Sigma = \{\mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_S\}$ and let the ensemble of polynomials corresponding to $\{\lambda(\mathbf{n})f(\mathbf{n})\}$ be identified with

$$\prod_{\mathbf{n}\in\Sigma} f(\mathbf{n})\mathcal{V}(\mathbf{n}) = (f(\mathbf{n}_1)\mathcal{V}(\mathbf{n}_1)) \times (f(\mathbf{n}_2)\mathcal{V}(\mathbf{n}_2)) \times \cdots \times (f(\mathbf{n}_S)\mathcal{V}(\mathbf{n}_S)) \subseteq \mathbb{R}^{2S},$$

where

$$f\mathcal{V} = \{(xf_1 - yf_2, xf_2 + yf_1) \in \mathbb{R}^2 : f_1 = \Re(f), f_2 = \Im(f), (x, y) \in \mathcal{V}\}.$$

Note that $f\mathcal{V}$ is a real algebraic variety in $\mathbb{R}^2$ if $\mathcal{V}$ is also. In a similar vein, we now treat complex projective space $\mathbb{X}$ (resp. variety $\mathbb{Y}$) as real projective space (resp. variety) of double dimensions and, by homogenization, we embed $\prod_{\mathbf{n}\in\Sigma} f(\mathbf{n})\mathcal{V}(\mathbf{n})$ in $\mathbb{X}$ and denote the resulting projective variety as $\mathbb{V} \subset \mathbb{X}$.

Clearly $\mathbb{Y} \cap \mathbb{V}$ is a real algebraic subvariety in $\mathbb{V}$. To show $\mathbb{Y} \cap \mathbb{V}$ is a measure-zero subset of $\mathbb{V}$ we only need to show that $\mathbb{Y} \cap \mathbb{V} \subsetneq \mathbb{V}$ in view of proposition 2 and 3.

Following the suggestion in [19], we now prove

Proposition 4. $\mathbb{Y} \cap \mathbb{V} \subsetneq \mathbb{V}$.

Proof of proposition 4. It suffices to find *one* irreducible polynomial in $\mathbb{V}$.

The argument is based on two observations. First the polynomial

$$F(x, y, z) = ax^r + by^r + cz^r \tag{A.4}$$

is irreducible for any positive integer r and any nonzero coefficients a, b, c. This follows from the fact that the criticality equations $F_x = F_y = F_z = 0$ have no solution in $\mathbb{P}^2$ and thus the algebraic variety $F = 0$ is non-singular.

Secondly, for any Σ satisfying the assumptions of theorem 2, there exists a set $T \subset \Sigma^\sharp$ of three points which can be transformed into $\{(r, 0, 0), (0, r, 0), (0, 0, r)\}$, the support of (A.4), under a rational map.

We separate the analysis of the second observation into two cases.

Case 1. $(0, 0) \in \Sigma$. Then there are at least two other points, say (m, n), (p, q), belonging to Σ. Without loss of generality, we assume $p + q = \delta$. Because Σ has rank 2, $mq - np \neq 0$.

We look for the rational mapping

$$z_1 = x^{k_{11}} y^{k_{21}} z^{k_{31}}, \quad z_2 = x^{k_{12}} y^{k_{22}} z^{k_{32}}, \quad z_0 = x^{k_{13}} y^{k_{23}} z^{k_{33}} \tag{A.5}$$

with $k_{ij} \in \mathbb{Z}$ that maps the polynomial

$$P(z_0, z_1, z_2) = cz_0^\delta + az_0^{\delta-m-n} z_1^m z_2^n + bz_1^p z_2^q$$

to $F(x, y, z)$. This amounts to a linear transformation from the set of independent vectors

$$(m, n, \delta - m - n), \quad (p, q, 0), \quad (0, 0, \delta)$$

to the set $\{(r, 0, 0), (0, r, 0), (0, 0, r)\}$. This transformation can be accomplished by the following matrix:

$$r \begin{pmatrix} m & p & 0 \\ n & q & 0 \\ \delta - m - n & 0 & \delta \end{pmatrix}^{-1} = \frac{r}{\delta(mq - np)} \begin{pmatrix} q\delta & -n\delta & -q(\delta - m - n) \\ -p\delta & m\delta & p(\delta - m - n) \\ 0 & 0 & mq - np \end{pmatrix}, \tag{A.6}$$

where the divisor is nonzero. To ensure integer entries in (A.6) we set

$$r = \delta(mq - np)$$

and obtain the transformation matrix

$$\begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} = \begin{pmatrix} q\delta & -n\delta & -q(\delta - m - n) \\ -p\delta & m\delta & p(\delta - m - n) \\ 0 & 0 & mq - np \end{pmatrix}.$$

Case 2. $(m, 0), (0, n) \in \Sigma$ for some positive integers m, n. Then there is at least another point $(p, q) \in \Sigma$ such that $(m, 0), (0, n), (p, q)$ are not collinear, which means $mn - np - mq \neq 0$.

Suppose $p + q = \delta$. Consider the polynomial

$$P(z_0, z_1, z_2) = az_0^{\delta-m} z_1^m + bz_0^{\delta-n} z_2^n + cz_1^p z_2^q.$$

By the same analysis above the form (A.4) can be achieved by the transformation matrix

$$r \begin{pmatrix} m & 0 & p \\ 0 & n & q \\ \delta - m & \delta - n & 0 \end{pmatrix}^{-1} = \frac{r}{\delta(mn - mq - np)} \begin{pmatrix} -q(\delta - n) & q(\delta - m) & -n(\delta - m) \\ p(\delta - n) & -p(\delta - m) & -m(\delta - n) \\ -pn & -mq & mn \end{pmatrix},$$

which has integer entries if r is a multiple of $\delta(mn - mq - np)$. With the choice

$$r = \delta(mn - mq - np)$$

the transformation matrix becomes

$$\begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} = \begin{pmatrix} -q(\delta - n) & q(\delta - m) & -n(\delta - m) \\ p(\delta - n) & -p(\delta - m) & -m(\delta - n) \\ -pn & -mq & mn \end{pmatrix}.$$

Suppose $n = \delta$. Consider the polynomial

$$P(z_0, z_1, z_2) = a z_0^{n-m} z_1^m + b z_2^n + c z_0^{n-p-q} z_1^p z_2^q.$$

The form (A.4) can be achieved by the transformation matrix

$$r \begin{pmatrix} m & 0 & p \\ 0 & n & q \\ n-m & 0 & n-p-q \end{pmatrix}^{-1}$$

$$= \frac{r}{n(mn - mq - np)} \begin{pmatrix} n(n-p-q) & q(n-m) & n(m-n) \\ 0 & mn - mq - np & 0 \\ -pn & -mq & mn \end{pmatrix}.$$

With

$$r = n(mn - mq - np)$$

the transformation matrix becomes

$$\begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} = \begin{pmatrix} n(n-p-q) & q(n-m) & n(m-n) \\ 0 & mn - mq - np & 0 \\ -pn & -mq & mn \end{pmatrix}.$$

Suppose $m = \delta$. Consider the polynomial

$$P(z_0, z_1, z_2) = a z_1^m + b z_0^{m-n} z_2^n + c z_0^{m-p-q} z_1^p z_2^q.$$

The form (A.4) can be achieved by the transformation matrix

$$r \begin{pmatrix} m & 0 & p \\ 0 & n & q \\ 0 & m-n & m-p-q \end{pmatrix}^{-1}$$

$$= \frac{r}{m(mn - mq - np)} \begin{pmatrix} mn - mq - np & 0 & 0 \\ p(m-n) & m(m-p-q) & m(n-m) \\ -pn & -mq & mn \end{pmatrix}.$$

With

$$r = m(mn - mq - np)$$

the transformation matrix becomes

$$\begin{pmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{pmatrix} = \begin{pmatrix} mn - mq - np & 0 & 0 \\ p(m-n) & m(m-p-q) & m(m-n) \\ -pn & -mq & mn \end{pmatrix}.$$

To conclude the proof of proposition 4, in any above case, if the polynomial $P(z_0, z_1, z_2)$ is reducible (i.e. has a non-monomial factor), then we can write $P = P_1 P_2$ and

$$F(x, y, z) = P_1(z_0(x, y, z), z_1(x, y, z), z_2(x, y, z)) P_2(z_0(x, y, z), z_1(x, y, z), z_2(x, y, z)), \quad \text{(A.7)}$$

where P_1, P_2 are non-monomial factors. Let l be the lowest (possibly negative) power in x, y, z of $P_i(z_0(x, y, z), z_1(x, y, z), z_2(x, y, z))$, $i = 1, 2$. If $l \geqslant 0$, then the factorization (A.7)

implies that $F(x, y, z)$ has a non-monomial factor. If $l < 0$, then the factorization (A.7) implies that $(xyz)^{-l}F(x, y, z)$ has a non-monomial factor. Either case contradicts the fact that F is irreducible. So P is irreducible. The proof of proposition 4 is complete. $\qquad\square$

Continuing the proof of theorem 2, we have from propositions 3 and 4 that $\mathbb{Y} \cap \mathbb{V}$ is a measure-zero subset of $\mathbb{V}$. By dehomogenization and proposition 1 reducible polynomials of a fixed support Σ under the assumptions of theorem 2 comprise a measure-zero subset of all polynomials of the same support. The proof of theorem 2 is now complete.

References

[1] Almoro P F and Hanson S G 2008 Random phase plate for wavefront sensing via phase retrieval and a volume speckle field *Appl. Opt.* **47** 2979–87
[2] Almoro P F, Pedrini G, Gundu P N, Osten W and Hanson S G 2011 Enhanced wavefront reconstruction by random phase modulation with a phase diffuser *Opt. Laser Eng.* **49** 252–7
[3] Bates R H T 1982 Fourier phase problems are uniquely soluble in more than one dimension: I. Underlying theory *Optik* **61** 247–62
[4] Bauschke H H, Combettes P L and Luke D R 2002 Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization *J. Opt. Soc. Am.* A **19** 1334–45
[5] Bräuer R, Wojak U, Wyrowski F and Bryngdahl O 1991 Digital diffusers for optical holography *Opt. Lett.* **16** 1427–9
[6] Bruck Yu M and Sodin L G 1979 On the ambiguity of the image reconstruction problem *Opt. Commun.* **30** 304–8
[7] Falldorf C, Agour M, Kopylow C V and Bergmann R B 2010 Phase retrieval by means of a spatial light modulator in the Fourier domain of an imaging system *Appl. Opt.* **49** 1826–30
[8] Fannjiang A 2011 Exact localization and super-resolution with noisy data and random illumination *Inverse Problems* **27** 065012
[9] Fannjiang A and Liao W 2011 Compressed sensing phase retrieval *Proc. IEEE Asilomar Conf. on Signals, Systems and Computers (Pacific Grove, CA, 6–9 Nov.)* pp 735–8
[10] Fannjiang A and Liao W 2012 Phase retrieval with random phase illumination, submitted
[11] Fenimore E E 1980 Coded aperture imaging: the modulation transfer function for uniformly redundant arrays *Appl. Opt.* **19** 2465–71
[12] Fienup J R 1982 Phase retrieval algorithms: a comparison *Appl. Opt.* **21** 2758–69
[13] Fienup J R 1987 Reconstruction of a complex-valued object from the modulus of its Fourier transform using a support constraint *J. Opt. Soc. Am.* A **4** 118–23
[14] Fienup J R and Wackerman C C 1986 Phase-retrieval stagnation problems and solutions *J. Opt. Soc. Am.* A **3** 1897–907
[15] Gottesman S R and Fenimore E E 1989 New family of binary arrays for coded aperture imaging *Appl. Opt.* **28** 4344–52
[16] Hayes M 1982 The reconstruction of a multidimensional sequence from the phase or magnitude of its Fourier transform *IEEE Trans. Acoust. Speech Signal Process.* **30** 140–54
[17] Hayes M H and McClellan J H 1982 Reducible polynomials in more than one variable *Proc. IEEE* **70** 197–8
[18] He H 2006 Simple constraint for phase retrieval with high efficiency *J. Opt. Soc. Am.* A **23** 550–6
[19] Kuperberg G 2012 Personal communication
[20] Landsberg J M 2012 *Tensors: Geometry and Applications* (Providence, RI: American Mathematical Society)
[21] Marchesini S 2007 A unified evaluation of iterative projection algorithms for phase retrieval *Rev. Sci. Instrum.* **78** 011301
[22] Miao J, Ishikawa T, Shen Q and Earnest T 2008 Extending x-ray crystallography to allow the imaging of non-crystalline materials, cells and single protein complexes *Annu. Rev. Phys. Chem.* **59** 387–410
[23] Miao J and Sayre D 2000 On possible extensions of x-ray crystallography through diffraction-pattern oversampling *Acta Cryst.* A **56** 596–605
[24] Miao J, Sayre D and Chapman H N 1998 Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects *J. Opt. Soc. Am.* A **15** 1662–9
[25] Osserman B personal communication
[26] Pitts T A and Greenleaf J F 2003 Fresnel transform phase retrieval from magnitude *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **50** 1035–45

[27] Romberg J 2009 Compressive sensing by random convolution *SIAM J. Imaging Sci.* **2** 1098–128
[28] Rothenberg J E 1997 Improved beam smoothing with SSD using generalized phase modulation *Proc. SPIE* **3047** 713
[29] Shafarevich I R 1977 *Basic Algebraic Geometry* (Berlin: Springer)
[30] Stark H 1987 *Image Recovery: Theory and Applications* (New York: Academic)
[31] Sylman D, Mico V, Garcia J and Zalevsky Z 2010 Random angular coding for super-resolved imaging *Appl. Opt.* **49** 4874–82
[32] Ueno K 1997 *Introduction to Algebraic Geometry* (Providence, RI: American Mathematical Society)