New Atomicity-Exploiting Algorithms for Super-Resolution X-Ray Crystallography

Andrew E. Yagle
Department of EECS, The University of Michigan, Ann Arbor, MI 48109-2122

Abstract— **The X-ray crystallography problem is to reconstruct a crystalline structure from the Fourier magnitude of its diffracted scattering data. This has three major difficulties: (1) Only Fourier magnitude (not phase) data are known; (2) There is no support constraint (since the crystal is periodic); and (3) only low-wavenumber scattering data are available. But it also has two major advantages: (1) the crystal is sparse (atomicity) since it consists of isolated atoms; and (2) the crystal structure often has even symmetry. We exploit atomicity to show that the crystal can be reconstructed easily from only low wavenumber Fourier data. We also propose new algorithms for reconstruction of crystals with even or no symmetry from low-wavenumber Fourier magnitude data using two or one isomorphic replacements (4 algorithms). Small numerical examples illustrate the algorithms.**

Keywords— **X-ray crystallography, sparse images.**
Phone: 734-763-9810. Fax: 734-763-1503.
Email: aey@eecs.umich.edu. **EDICS: 2-REST.**

I. INTRODUCTION

A. Background

X-ray crystallography is the reconstruction of molecular structures from X-ray scattering data. An X-ray source generates a plane wave that probes the structure and is scattered and diffracted from it. The magnitude of the scattered field is measured. The goal is to reconstruct the structure from measurements of the Fourier magnitude of the scattered field.

Since scattering from individual atoms is very weak, crystal structures are used. The periodicity of crystal structures, which consist of many repetitions of a unit cell, means that scattering from each unit cell is reinforced across all the unit cells, resulting in a strong scattering signal, which creates an interference pattern of spots. However, since there is no lens for X-rays, only the Fourier magnitude of the scattered field can be measured; the phase must be determined separately (phase retrieval).

Each unit cell in turn consists of atoms, surrounded by a cloud of electrons and arranged in an unknown structure. Most of the unit cell can be modelled as empty space, with atoms represented as impulses in the electron density distribution (atomicity). In practice the empty space is filled with solvent, but its effects can be subtracted off (solvent leveling).

B. Approaches

One important approach to X-ray crystallography consists of *direct methods* developed by Karle & Hauptman [1,2], for which they won the Nobel prize in 1985. Non-negativity of the electron density manifests itself in the scattering data as a set of matrix determinant inequalities (derived below), and these determine inequalities for the unknown Fourier phases. Using probability density function models, the probability of an inequality determining the phase can be explicitly computed, and a model of the crystal built up. However, these methods are computationally intensive and require much human analysis, interpretation, and modelling. In practice, direct methods are only useful for small (less than 100 atoms per unit cell) molecules, since the resolution is limited.

Another approach used for large molecules (e.g., proteins) is *isomorphic replacement*. A heavy atom is inserted into each unit cell of the crystal structure, and scattering data for the crystal both before and after this insertion is measured. Performing this operation twice uniquely determines the phases, as shown below. However, this requires extensive laboratory work, time, and cost. Performing this operation only once, termed single isomorphic replacement (SIR), would reduce time and cost by almost half.

Still another approach is *anomalous dispersion*. The X-ray source is tuned to an absorption edge of a heavy atom, altering its scattering $o(x)$ and in fact making it complex (due to absorption and dispersion). This has an effect analogous to SIR. Performing it twice, termed multiple anomalous dispersion (MAD), again uniquely determines the phases. However, repeated bombardment with X-rays can "cook" the crystal, so single anomalous dispersion (SAD) is preferable to multiple anomalous dispersion (MAD).

For structures with more than about 500 atoms per unit cell, interatomic distances within the unit cell are less than the wavelength of the X-rays (about 1 Angstrom). This makes the problem much harder, since atomicity can no longer be used (atomicity is still true, but it does not appear in the data).

Most approaches use symmetries in the crystal structure. There are two types of symmetry:

- *Crystallographic symmetry* (CS) in which the structure is invariant to rotations preserving the lattice and thus apply throughout the crystal;
- *Noncrystallographic symmetry* (NCS) in which the structure is invariant to rotations interpolating the lattice and thus apply only locally.

NCS is discussed extensively in [3], in which the lattice interpolation is used as additional data in an iterative phase retrieval algorithm. It is not used in this paper, so it is not discussed further. CS, specifically even or centrosymmetric symmetry, is used in this paper and is discussed in more detail below.

C. Contribution of This Paper

This paper presents a new approach to the X-ray crystallography problem. It presents the following:

- A new superresolution algorithm that exploits atomicity (sparseness) directly to reconstruct atomic locations from low-wavenumber scattering data;
- Two new algorithms for reconstruction of symmetric and non-symmetric structures from SIR data. Both are explicit and noniterative (closed form);
- Sets of equations for reconstruction of symmetric and non-symmetric structures without using isomorphic replacement or anomalous dispersion. Solution is computationally intensive, but provide sufficient conditions for unique reconstruction.

This paper also presents new formulae for reconstruction using two isomorphic replacements that obviate the need for human-assisted structure modelling.

D. Paper Organization

This paper is organized as follows. Section II reviews diffraction imaging, and then specializes to X-ray crystallography. The presentation illustrates how the latter problem can be viewed as a special case of the former one. Section III presents a new algorithm for reconstruction of crystals from low-wavenumber scattering data that exploits sparsity but requires phase. Section IV presents new formulae for crystallography with two isomorphic replacements. Section V, the main contribution, presents new algorithms for single isomorphic replacement (SIR) for crystals with even symmetry or no symmetry; these algorithms can also be applied to single anomalous dispersion (SAD). Section VI presents equations for crystals with even or no symmetry without requiring either isomorphic replacement or anomalous dispersion. Although computationally intensive, these equations demonstrate sufficient conditions for uniqueness of reconstruction, which is itself of interest. Section VII is a table comparing terms in X-ray crystallography with equivalent terms used in signal processing Fourier analysis.

II. CRYSTAL DIFFRACTION IMAGING

The first subsection reviews general diffraction imaging in optics; the second specializes to crystals.

A. Diffraction Imaging

Consider an object $\{o(x), x \in \mathcal{R}^3\}$ known to be zero outside the sphere $|x| \leq R$ for some finite radius R (so $o(x)$ has compact support) illuminated with a plane wave $\delta(t - \vec{e_I} \cdot x/c)$ in a direction specified by the unit vector $\vec{e_I}$, travelling at wave speed c. Taking temporal Fourier transforms to replace time dependence with frequency dependence results in

$$\mathcal{F}_{t\to\omega}\{\delta(t - \vec{e_I} \cdot x/c)\} = \int \delta(t - \vec{e_I} \cdot x/c) e^{-i\omega t} dt$$
$$= e^{-i\omega(\vec{e_I}\cdot x)/c} = e^{-i2\pi(\vec{e_I}\cdot x)/\lambda} \quad \lambda = \text{wavelength} \quad (1)$$

Here wavelength replaces frequency over wave speed.

For each $x \in \mathcal{R}^3$, this plane wave is scattered by $o(x)$, producing a spherically-spreading scattered field. In the direction specified by the unit vector $\vec{e_S}$, the scattered field in the far field (large $|x|$) is

$$e^{-i\frac{2\pi}{\lambda}\vec{e_I}\cdot x} o(x) \frac{1}{4\pi|x|} e^{i\frac{2\pi}{\lambda}\vec{e_S}\cdot x}$$

We now make the *Born approximation*, which is that the scattered field is not further scattered by $o(x)$ at other values of x. This amounts to assuming that $|o(x)| << 1$, so that $o(x_1)o(x_2)$ is negligible vs. $o(x)$. This linearizes the problem, and allows us to state that the total measured field from all of $o(x)$ is

$$\underbrace{e^{-i\frac{2\pi}{\lambda}\vec{e_I}\cdot x}}_{\text{INCIDENT}} + \frac{1}{4\pi|x|} \underbrace{\int e^{-i\frac{2\pi}{\lambda}(\vec{e_I}-\vec{e_S})\cdot x} o(x) dx}_{\text{SCATTERED}} \quad (2)$$

The goal is to reconstruct $o(x)$ from this field. In the sequel we subtract off the incident plane wave and omit the geometric spreading factor $1/(4\pi|x|)$.

As the incident $\vec{e_I}$ and scattered $\vec{e_S}$ directions sweep over the unit sphere, the 3-D Fourier transform $O(k)$ of $o(x)$ is determined over a set of spheres of radius $\frac{2\pi}{\lambda}$ centered on another sphere of radius $\frac{2\pi}{\lambda}$. This is also the set of all spheres of radius $\frac{2\pi}{\lambda}$ that pass through the origin.

Some thought (picture a flyball governor spinning around all possible axes) shows that this sweeps over all wavenumbers with magnitudes $\leq \frac{4\pi}{\lambda}$. In fact, we can omit half of the incident or scattered directions and still recover $\{O(k), |k| \leq \frac{4\pi}{\lambda}\}$ (picture an anemometer spinning around all possible axes and

use reciprocity). Hence we can recover only a low-wavenumber-filtered version of $o(x)$; the shorter the wavelength λ, the higher the $o(x)$ cutoff wavenumber. This is important in the development to follow.

For more details see any paper on diffraction scattering or tomography. We have found the papers of A.J. Devaney to be particularly helpful to us.

B. X-Ray Crystallography

We now specialize to the case of $o(x)$ is a crystal:

- $o(x)$ is *periodic*: $o(x) = o(x+[L_x, L_y, L_z])$ for some lengths L_x, L_y, L_z. Each period of $o(x)$ is a unit cell;
- $o(x)$ is *atomic*: $o(x) = \sum_{n=1}^{M} o_n \delta(x - x_n)$ for some values and locations $\{(o_n, x_n), n = 1 \dots M\}$. o_n is proportional to the atomic number of the n^{th} atom;
- $o(x)$ is *non-negative*: $o(x) \geq 0$;
- $o(x)$ may be *(centro)symmetric*: $o(x) = o(-x)$.

More precisely $o(x)$ is sparse (mostly zero-valued) and its nonzero values specify the electron density, which is clustered around atomic nuclei. It thus may be collections of small regions rather than impulses.

These properties of $o(x)$ imply the following properties of its Fourier transform $O(k)$:

- $O(k)$ is *discrete* in wavenumber k (see below);
- $O(k)$ is *real-valued* if $o(x)$ is symmetric;
- Atomicity and non-negativity of $o(x)$ lead to more complicated properties of $O(k)$ (see below).

More precisely $O(k)$ can be written as

$$\sum_{i,j,k} O_{i,j,k} \delta(k_x - i\frac{2\pi}{L_x}) \delta(k_y - j\frac{2\pi}{L_y}) \delta(k_z - k\frac{2\pi}{L_z}). \quad (3)$$

More properly, the periodic $o(x)$ can be expanded in a 3-D Fourier series with Fourier coefficients $O_{i,j,k}$.

However, there are also consequences to the use of X-ray wavelengths, which have $\lambda \approx 1$ Angstrom:

- Only the $o(x)$ Fourier magnitude $|O(k)|$ can be measured, since there is no lens for imaging X-rays;
- Only a low-wavenumber-filtered version of $o(x)$ can be recovered, with resolution about 1 Angstrom.

We thus have the two problems of phase retrieval (recovering $\angle O(k)$) and superresolution or bandwidth extrapolation (recovering $O(k)$ for large $|k|$). There are well-known iterative algorithms for both of these problems that require a *support constraint*: $o(x) = 0$ for $|x| > R$ for some R. However, these algorithms cannot be applied here, since the periodicity of $o(x)$ implies there is no support constraint. It could be assumed that in each unit cell there is a bounding region in which the crystal is known to be empty of atoms, but this is seldom true in practice.

The problem is as follows: How to take advantage of sparsity and periodicity to recover the phase and high-wavenumber information. Note sparsity cannot be used as a support constraint: Although $o(x)$ is mostly zero, there are no regions in which $o(x)$ is *known* to be zero, so there is no *fixed* constraint.

In the next section, we present algorithms that solve the superresolution half of the problem.

III. SPARSE IMAGE SUPERRESOLUTION

A. Problem Formulation

We assume the Fourier transform $|O(k)| = 0$ for $|k_i| > K$ for some large K. Then we can sample $o(x)$ on a lattice x=$[i, j, k]\Delta$ where $\Delta = \frac{2\pi}{K}$. Alternatively, nonzero values of $o(x)$ may be constrained to lie only on such a lattice for some Δ; we do not distinguish these cases in the sequel. Then the spatial Fourier transform, which has become a Fourier series due to the periodicity of $o(x)$, now further becomes a 3-D discrete Fourier transform (DFT) with N=$KL/(2\pi)$:

$$O_{l,m,n} = \sum_{i,j,k} o_{i,j,k} e^{-2\pi\sqrt{-1}(il+jm+kn)/N} \quad (4)$$

where we take L_x=L_y=L_z=L for convenience and

$$O_{l,m,n} = O([l, m, n]\frac{2\pi}{L}); \quad o_{i,j,k} = o([i, j, k]\frac{2\pi}{K}) \quad (5)$$

Note that the DFT implicitly creates periodic extensions of both $o_{i,j,k}$ and $O_{i,j,k}$. This is quite proper since (showing only one of the three dimensions):

$$o(x) = o(n\frac{2\pi}{K}) = o(x + L) = o((n + N)\frac{2\pi}{K})$$
$$O(k) = O(n\frac{2\pi}{L}) = O(k + K) = O((n + N)\frac{2\pi}{L}) \quad (6)$$

since we have the relations

$$N\frac{2\pi}{K} = \frac{KL}{2\pi}\frac{2\pi}{K} = L; \quad N\frac{2\pi}{L} = \frac{KL}{2\pi}\frac{2\pi}{L} = K. \quad (7)$$

Note $O(k)$ can be made periodic without aliasing since $O(k)$ is assumed to be bandlimited.

$o_{i,j,k}$ is known to be nonzero only at M unknown locations. The goal is to reconstruct $o_{i,j,k}$ from the low-wavenumber data $\{O_{i,j,k}, |i|, |j|, |k| \leq M^{1/3}\}$.

B. Problem Solution

Let $s_{i,j,k}$ be the 3-D function having DFT $S_{i,j,k}$:

- $S_{i,j,k}$=0 unless $|i|, |j|, |k| \leq \frac{1}{2}M^{1/3}$ (not $M^{1/3}$);
- Support of $S_{i,j,k}$ is $\frac{1}{8}$ #known values of $O_{i,j,k}$;

- $s_{i,j,k}$ is the indicator function for nonzero $o_{i,j,k}$:

$$\begin{cases} s_{i,j,k} = 0 & \text{if } o_{i,j,k} \neq 0; \\ s_{i,j,k} \neq 0 & \text{if } o_{i,j,k} = 0 \end{cases}. \tag{8}$$

Then ($***$ denotes 3-D cyclic convolution)

$$o_{i,j,k}s_{i,j,k} = 0 \rightarrow O_{i,j,k} ** S_{i,j,k} = 0. \tag{9}$$

Since there are only M nonzero values of $o_{i,j,k}$, $s_{i,j,k}$ is uniquely determined, to an (irrelevant) scale factor.

The second of (9) can be written as a Toeplitz-nested-Toeplitz (TNT) linear system of equations for the unknown $S_{i,j,k}$. This matrix is block Toeplitz, and each block is itself block Toeplitz, and each block of the blocks is Toeplitz. This TNT structure is derived as a submatrix of a circulant-block-circulant matrix in the alternate derivation to follow below.

$S_{i,j,k}$ are the elements of the null vector of this TNT matrix constructed from the known $O_{i,j,k}$. Then an inverse 3-D DFT computes $s_{i,j,k}$, which is zero at the locations $\{(i_n, j_n, k_n), 1 \leq n \leq M\}$ of the nonzero $o_{i,j,k}$. Once the locations of the nonzero $o_{i,j,k}$ are known, their values can be computed by solving a second linear system of equations in M unknowns.

Finally, note that if $o_{i,j,k}$=0 or 1 (binary object), then we have Sayre's equation (compare to (9))

$$o_{i,j,k}o_{i,j,k} = o_{i,j,k} \rightarrow O_{i,j,k} ** O_{i,j,k} = O_{i,j,k}. \tag{10}$$

C. Matrix Derivation of Algorithm

We now present a second derivation of the above algorithm, for reasons noted below. We can write

$$O_{i_1-i_2,j_1-j_2,k_1-k_2} = \sum_{n=1}^{M} o_{i_n,j_n,k_n} e^{-\frac{2\pi}{N}\sqrt{-1}[i_n(i_1-i_2)+j_n(j_1-j_2)+k_n(k_1-k_2)]}. \tag{11}$$

The $(N^3 \times N^3)$ circulant-block-circulant matrix [C] having for its first row $O_{i,j,k}$ can then be factored as

$$[C] = D^H \cdot \text{DIAG}[o_{1,1,1} \cdots o_{N,N,N}]D \tag{12}$$

where D is the Kronecker product of the 1-D DFT matrix having $(n,k)^{th}$ element $e^{-i\frac{2\pi}{N}(n-1)(k-1)}$ with itself three times (this is the 3-D DFT matrix).

By assumption only M of the N^3 diagonal values $o_{i,j,k}$ are nonzero. Let F be the submatrix of D in which all rows but those corresponding to $\{(i_n, j_n, k_n), 1 \leq n \leq M\}$ have been deleted, and all columns but those corresponding to the support of $s_{i,j,k}$ (which is defined above) have been deleted. Then the (M+1)×(M+1) TNT submatrix $[O]$ of $[C]$ can be factored as

$$[O] = F^H \cdot \text{DIAG}[o_{(i_n,j_n,k_n)}]F. \tag{13}$$

This (M+1)×(M+1) matrix clearly has rank M, so it has a null vector $\vec{a} = [a_1 \ldots a_{M+1}]'$. Postmultiplying (13) by this null vector $\vec{a}$ gives

$$[O]\vec{a} = F^H \cdot \text{DIAG}[o_{i_n,j_n,k_n}]F\vec{a} = 0 \rightarrow F\vec{a} = 0. \tag{14}$$

Since F is a submatrix of the 3-D DFT matrix D, we can compute $D\vec{a}$ using a 3-D FFT and see which values are zero. The rows of D corresponding to those values are the rows of F, and this identifies the locations $\{i_n, j_n, k_n\}$ of nonzero $o_{i,j,k}$.

Although this derivation is more complicated than the first one, it makes three important points:

- The *conditioning* of the problem is determined primarily by the condition number of the matrix F. If a compact support constraint is used, F is extremely ill-conditioned. But if the M nonzero values o_{i_n,j_n,k_n} are (roughly) evenly spaced throughout the $N^3 \times N^3$ region, F will be reasonably well-conditioned;
- If there are in fact fewer than M nonzero values of $o_{i,j,k}$, the second derivation shows that the TNT matrix $[O]$ is rank-deficient by more than one; its rank is the actual number of nonzero $o_{i,j,k}$;
- Since the diagonal elements are $o_{i,j,k} \geq 0$, the matrix $[C]$ is positive semi-definite. Sylvester's criterion then states that all of the leading principal minors of $[C]$ are non-negative. These are the Karle-Hauptman determinants, which are used in direct methods.

D. Missing Low-wavenumber Data

In practice, very low-wavenumber data is difficult to obtain. The above procedure can easily be adapted to missing low-wavenumber data, as follows.

Suppose that $\{O_{i,j,k}, K_L \leq |i|, |j|, |k| \leq K_H\}$ is known for some bounds K_L and K_H. This forms a hollow cube in wavenumber. We merely alter the support of the indicator function $s_{i,j,k}$ to the similarly-shaped, but smaller, hollow cube

$$\left\{\frac{3}{4}K_L + \frac{1}{4}K_H \leq |i|, |j|, |k| \leq \frac{3}{4}K_H + \frac{1}{4}K_L\right\}. \tag{15}$$

Then (9) can still be written as a system of equations for reconstructing $s_{i,j,k}$ from the given $O_{i,j,k}$. This can easily be visualized by sliding around the support of $s_{i,j,k}$ inside the support of the given $O_{i,j,k}$. Then a 3-D DFT (which can be computed quickly using the FFT) still indicates the locations of non-zero values of $o_{i,j,k}$ with zero values of $s_{i,j,k}$.

Other topologies, such as a hollow sphere of $O_{i,j,k}$, can be accommodated similarly. The indicator function $s_{i,j,k}$ has a shape like that of the given $O_{i,j,k}$, so that it can slide around inside the support of the given $O_{i,j,k}$, generating a linear system of equations for the unknown $s_{i,j,k}$.

This shows that atomicity of the crystal can be exploited to reconstruct it exactly from the low-wavenumber diffraction data obtained using X-rays. However, this still requires phase retrieval from the given magnitude data. We address this next.

IV. TWO ISOMORPHIC REPLACEMENTS

A. Problem Formulation

In isomorphic replacement a single atom is introduced into the crystal at the same position in each unit cell. This is a complicated laboratory procedure that requires facilities, time and money. But it does allow phase data to be obtained from the magnitude of the diffraction data, as follows.

A single atom is introduced into the unit cell. Without loss of generality, let this atom be an impulse of unit area and its location be the origin. Then

$$|1 + O_{i,j,k}|^2 - |O_{i,j,k}|^2 = 1 + 2|O_{i,j,k}| \cos\theta_{i,j,k} \quad (16)$$

so $\cos\theta_{i,j,k}$ is determined from magnitude at each wavenumber observed. This still leaves a sign ambiguity in $\theta_{i,j,k}$. This is resolved by using a second isomorphic replacement at a different location, and solving two simultaneous equations. A more elegant presentation, which seems to be new, is as follows.

B. New Algorithm

Let the origin be defined as the *midpoint* of the two included atoms, which are now at positions $\pm x_o$. We observe $|O(k)|$ and $|O(k)+e^{\pm ik\cdot x_o}|$ (vectors k and x_o):

$$\begin{aligned} |O(k) + e^{-ik\cdot x_o}|^2 &= 1 + |O(k)|^2 - 2|O(k)|\cos(\theta_k + k\cdot x_o) \\ |O(k) + e^{+ik\cdot x_o}|^2 &= 1 + |O(k)|^2 - 2|O(k)|\cos(\theta_k - k\cdot x_o) \end{aligned} \quad (17)$$

Adding and subtracting gives the two equations

$$\begin{aligned} |O(k) + e^{-ikx_o}|^2 + |O(k) + e^{+ikx_o}|^2 = \\ 2 + 2|O(k)|^2 - 4|O(k)|\cos(k\cdot x_o)\cos(\theta_k) \\ |O(k) + e^{-ikx_o}|^2 - |O(k) + e^{+ikx_o}|^2 = \\ 4|O(k)|\sin(k\cdot x_o)\sin(\theta_k) \end{aligned} \quad (18)$$

using the two trigonometry identities

$$\begin{aligned} \cos(x-y) + \cos(x+y) &= 2\cos(x)\cos(y) \\ \cos(x-y) - \cos(x+y) &= 2\sin(x)\sin(y) \end{aligned} \quad (19)$$

Now we proceed as follows:

- Use the 1^{st} equation to determine $\cos(\theta_k)$ and the 2^{nd} to determine the sign of θ_k from $\sin(\theta_k)$, which has the same sign as θ_k. This uniquely specifies θ_k;
- This has the advantage that magnitude (squared) data are added, rather than subtracted, for a larger signal-to-noise ratio. On the other hand, subtracting can eliminate any common interference in the data;
- When $|O(k)| = 0$ the formulae break down, but then the phase is irrelevant anyways;
- If $\cos(k\cdot x_o)$ or $\sin(k\cdot x_o)$ are zero, the formulae also break down. Either $\cos(\theta_k)$ or $\sin(\theta_k)$ is determined, but not both, leaving a sign ambiguity.

The superresolution algorithm presented in the previous section can then be applied to the low-wavenumber magnitude and phase data to reconstruct the atomic structure exactly. This obviates the need for human-assisted visualization from low-wavenumber (blurred) electron density. This is a new algorithm for X-ray crystallography. But only one isomorphic replacement is needed, as we show next.

V. ONE ISOMORPHIC REPLACEMENT

The expense of isomorphic replacement means that algorithms that require only one are greatly preferable to algorithms requiring two. In this section, we present two new algorithms for X-ray crystallography that require only a single isomorphic replacement, and thus require phase retrieval as well as superresolution. One algorithm is for the (centro)symmetric objects, and the other is for non-symmetric objects.

A. (Centro)symmetric Objects

Suppose that the object is (centro)symmetric, so that $o(x) = o(-x)$. Then the object is invariant to 180^o rotation about three orthogonal axes. Note:

- Symmetry about each individual axis is sufficient but not necessary for overall (centro)symmetry;
- This is crystallographic symmetry, not lattice-interpolating non-crystallographic symmetry (ncs);
- (Centro)symmetric objects have phases=0 or π. This greatly simplifies the phase retrieval problem.

One heavy atom is inserted into the unit cell at location x_o. It's represented by $c\delta(x - x_o)$ for some constant c proportional to the atomic number of the atom. Translate the origin by x_o, so the location of the inserted atom is now the origin. We now observe

$$\begin{aligned} |c + e^{-ik\cdot x_o} O(k)|^2 = \\ [c + O(k)\cos(k\cdot x_o)]^2 + [O(k)\sin(k\cdot x_o)]^2 \end{aligned} \quad (20)$$

since $O(k)$ is real-valued. Then we can compute

$$c + O(k)\cos(k \cdot x_o) =$$

$$\sqrt{|c + e^{-ik \cdot x_o} O(k)|^2 - [O(k)\sin(k \cdot x_o)]^2} \qquad (21)$$

provided $c > |O(k)|$ so the positive square root is used. The 1^{st} term is observed and the 2^{nd} term is known since $|O(k)|^2 = O(k)^2$ since $O(k)$ is real-valued. Hence $O(k)$ is determined unless $\cos(k \cdot x_o) = 0$.

The condition $c > |O(k)|$ should not be confused with the original heavy atom condition $c >> |O(k)|$, in which case all of the phases were approximately zero. Here we only need that c is large enough that the two-fold phase ambiguity can be resolved. Note that if $x_o = 0$ (the point of inclusion is also the point of symmetry), the condition $c > |O(k)|$ means that $c + O(k) > 0$ so that $c\delta(x) + o(x)$ can be reconstructed directly from its low-wavenumber data using the superresolution algorithm, without the above formulae.

B. Numerical Example

We present a numerical example of this algorithm. An easy-to-follow small example is provided as part of the small example for non-symmetric objects below. For ease of visualization a 2-D problem is used.

624 non-zero values of a even binary function `X` were randomly distributed throughout a 375×375 region by thresholding and rounding Matlab's `rand` function. The threshold was determined by trial and error, and the signal added to its 2-D reversal. The squared magnitudes of the DFT of the function and the function with a unit inclusion at the origin were computed at the lowest 25×25 wavenumbers, centered at the origin. The above algorithm was used.

Fig. 1 shows the original image and Fig. 2 shows the reconstructed image. They can be seen to be identical. Fig. 3 shows the reconstruction without superresolution. Note the following about this:

- The minimum singular value of the 625×625 TNT matrix is 4×10^{-11}, so that Matlab's `null` can be used. The next-smallest singular value is 1.5×10^{-6};
- Typical values of the indicator function indicating non-zero locations of `X` are around 10^{-12}. Other values are around 0.01. Fig. 2 displays the reciprocal of the function$+10^{-6}$ (added for display purposes);
- The low-wavenumber reconstruction Fig. 3, which uses the correct phases and sets the unknown high wavenumbers to zero, shows the significance of the superresolution in specifying non-zero locations.

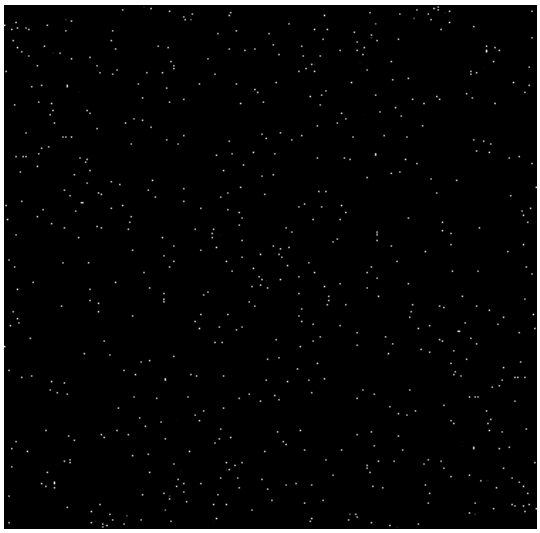

Fig. 1. Original image.

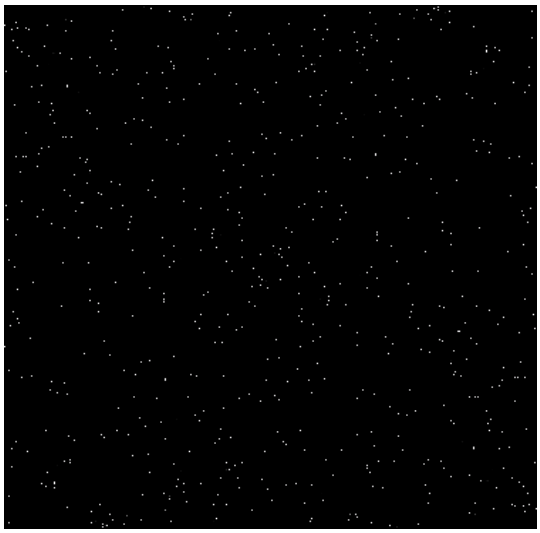

Fig. 2. Reconstructed image.

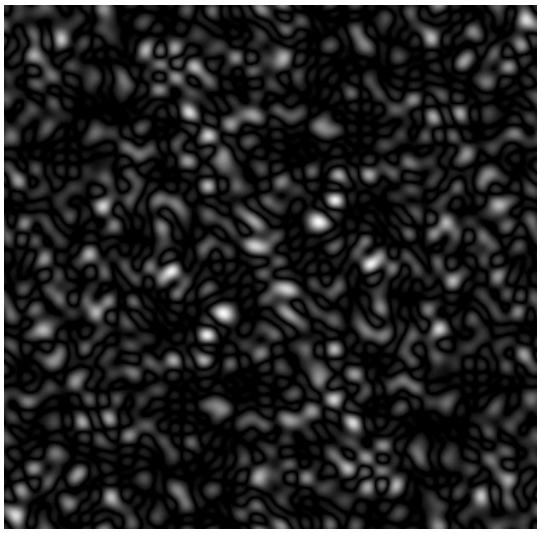

Fig. 3. Low-wavenumber image.

```
clear;rand('seed',1);X=rand(375,375);
X(X<0.997905)=0;X=X+fliplr(flipud(X));
X=round(X);I1=[188:375 1:187];
F=fftshift(fft2(X(I1,I1)));
Y1=abs(F+ones(375,375)).^2;
Y0=abs(F).^2;Y2=(Y1-Y0-ones(375,375))/2;
Y3=Y2(164:212,164:212);Y4=real(Y3(:));
TT=toeplitz(flipud(Y4(1:1201)),Y4(1201:2401)');
I2=[];for I=0:24;I2=[I2 [1:25]+49*I];end;
T=TT(I2,I2);S4=abs(ifft2(Y3,375,375));
N=null(T);S2=fft2(reshape(N,25,25),375,375);
I3=[189:375 1:188];
S3=1./(.000001+abs(S2(I3,I3)));
figure;imagesc(X),colormap(gray)
figure;imagesc(S3),colormap(gray)
figure;imagesc(S4),colormap(gray)
```

C. Non-Symmetric Objects

Now we no longer assume $o(x)$ has any symmetry. A single atom is inserted at the origin. Recall (16)

$$|1+O_{i,j,k}|^2-|O_{i,j,k}|^2=1+2\text{Real}[O_{i,j,k}] \qquad (22)$$

Define the even (symmetric) part $e_{i,j,k}$ and the odd (antisymmetric) part $x_{i,j,k}$ of the object $o_{i,j,k}$ as

$$\begin{aligned} e_{i,j,k} &= [o_{i,j,k}+o_{N-i,N-j,N-k}]/2 \\ x_{i,j,k} &= [o_{i,j,k}-o_{N-i,N-j,N-k}]/2 \\ o_{i,j,k} &= e_{i,j,k}+x_{i,j,k} \qquad . \end{aligned} \qquad (23)$$

The DFTs $E_{i,j,k}$ of $e_{i,j,k}$ and $X_{i,j,k}$ of $x_{i,j,k}$ are

$$\begin{aligned} E_{i,j,k} &= \quad \text{Real}[O_{i,j,k}] \\ X_{i,j,k} &= j\text{Imag}[O_{i,j,k}] \end{aligned} \qquad (24)$$

so that $E_{i,j,k}$ is known for low wavenumbers.

Since $o_{i,j,k}$ is sparse, $e_{i,j,k}$ and $x_{i,j,k}$ are both also sparse, with at most $2M$ nonzero values of each. Thus if $E_{i,j,k}$ is known for 2M+1 low wavenumbers:

- $e_{i,j,k}$ can be computed **using superresolution**;
- $x_{i,j,k}\neq 0$ only where $e_{i,j,k}\neq 0$ (same support);
- $X_{i,j,k}=\pm\sqrt{|O_{i,j,k}|^2-E_{i,j,k}^2}$ at low wavenumbers;
- Arbitrary-phase phase retrieval has become sign-ambiguity phase retrieval, with support constraint.

This phase retrieval problem could be solved using a variation of the hybrid input-output phase retrieval algorithm, in which the known support constraint is alternated with the known (to sign) low-wavenumber Fourier values. This algorithm must combine superresolution with phase retrieval, and may not converge. But we can solve the problem in closed form.

D. 1^{st} Sign Reconstruction Algorithm

The first algorithm requires

- More data: $\frac{1}{2}$M(M+1) non-redundant values;
- Less computation: Solving a $\frac{1}{2}$M(M+1)$\times\frac{1}{2}$M(M+1) linear system of equations.
- Direct computation of the odd part $x_{i,j,k}$ so that extrapolation of $X_{i,j,k}$ is not required.

Squaring each of the M equations

$$\sum_{n=1}^{M} x_{i_n,j_n,k_n}\sin(\frac{2\pi}{N}[i_nI+j_nJ+k_nK])=X_{I,J,K} \qquad (25)$$

yields a linear system of equations in the $\frac{1}{2}$M(M+1) unknowns $\{x_{i_m,j_m,k_m}x_{i_n,j_n,k_n}\}$ with the known $X_{I,J,K}^2$ as data. Arranging the solution into an $M\times M$ rank=1 matrix and computing its rank-one decomposition yields x_{i_m,j_m,k_m} to an overall sign.

$$o_{i,j,k}=e_{i,j,k}\pm x_{i,j,k}. \qquad (26)$$

The sign ambiguity here produces the reversal ambiguity inherent in phase retrieval of non-symmetric objects. Since this amounts to looking at the crystal from either of opposite sides, this is not an issue.

E. Numerical Example

We present a numerical example of this algorithm. Again a 2-D problem is used for visualization.

40 non-zero values of a binary function `X` were randomly distributed throughout a 128×128 region by thresholding and rounding Matlab's `rand` function. The threshold was determined by trial and error. The squared magnitudes of the DFT of the function and the function with a unit inclusion at the origin were computed at the lowest 41×41 wavenumbers, centered at the origin. The above algorithm was used.

Fig. 4 shows the original image and Fig. 5 shows the reconstructed image. They can be seen to be identical. Note the following about this example:

- $o_{i,k,j}$ has 40 non-zero values; $e_{i,k,j}$ and $x_{i,k,j}$ each has 80 non-zero values but only 40 *unknown* ones by symmetry; leading to $\frac{1}{2}$(40)(41)=820 unknowns;
- $|O_{i,k,j}|$ is known for 41×41 values, half of which are redundant leading to $\frac{1}{2}((41)^2$-1)=840 equations;
- $s_{i,j,k}$ has 9×9 support and it is computed from 17×17 lowest wavenumbers (17=9+9-1), leading to computing the null vector of an 81×81 TBT matrix;
- The supports of $e_{i,k,j}$ and hence $x_{i,k,j}$ are both determined from the DFT of $S_{i,k,j}$ as before;

• Numerical properties can be determined by running the following (very inefficient) Matlab program.

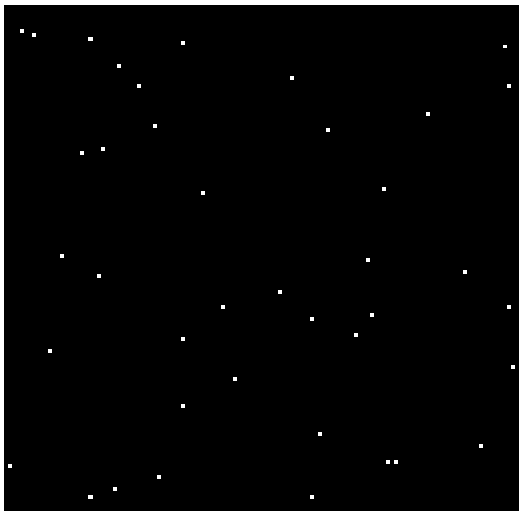

Fig. 4. Original image.

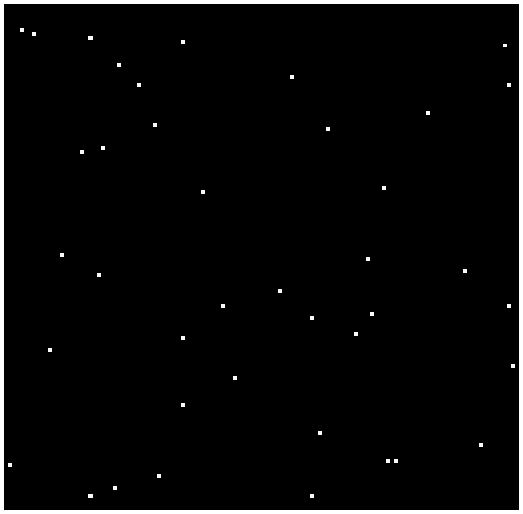

Fig. 5. Reconstructed image.

Matlab code used to generate this example:

```
clear;rand('seed',1);X=rand(128,128);
X=round(X-.49784);F=fft2(X,128,128);
Y=abs(F+1).^2-abs(F).^2-ones(128,128);
Y1=fftshift(Y);Y2=Y1(57:73,57:73);Y3=Y2(:);
TT=toeplitz(flipud(Y3(1:145)),Y3(145:289)');
I2=[];for I=0:8;I2=[I2 [1:9]+17*I];end;
T=TT(I2,I2);N=null(T);
S2=fft2(reshape(N,9,9),128,128);
%figure,imagesc(1./(.00001+abs(S2)))
[M,N]=find(abs(S2)<0.0001);
K11=[0:20 44:63]';K21=[0:20]';
F1=abs(F(K11+1,K21+1)).^2;F2=F1(:);
F3=Y(K11+1,K21+1).^2;F4=F3(:)/16;
K1=kron(ones(21,1),K11);
K2=kron(K21,ones(41,1));
AA=[];A=[];for I=1:40;
W=pi/64*((M(I)-1)*K1+(N(I)-1)*K2);
AA=[AA -sin(W)];end
for K=1:41*21;C=[];B=AA(K,:)'*AA(K,:);
for I=1:40;C=[C B(I,I)];
for J=I+1:40;C=[C 2*B(J,I)];
end;end;A=[A;C];end;XOHAT=A\(F2/4-F4);
XE=zeros(128,128);XO=zeros(128,128);
for I=1:40;
XO(M(I),N(I))=2*XOHAT(I);XE(M(I),N(I))=1/2;
XO(130-M(I),130-N(I))=-2*XOHAT(I);
XE(130-M(I),130-N(I))=1/2;end;XHAT=XE+XO;
figure,imagesc(X),colormap(gray)
figure,imagesc(XHAT),colormap(gray)
```


F. 2^{nd} Sign Reconstruction Algorithm

The second algorithm requires

• Less Fourier data: 125M non-redundant values;
• More computation: Solution of a $\frac{1}{2}$125M(125M+1) $\times\frac{1}{2}$125M(125M+1) linear system of equations;
• Direct computation of $X_{i,j,k}$ for small (i,j,k) so that extrapolation of $X_{i,j,k}$ is now needed.

$S_{i,j,k}$ is now computed from $E_{i,j,k}$ by writing

$$\begin{aligned} e_{i,j,k}s_{i,j,k} = 0 &\rightarrow E_{i,j,k} * * * S_{i,j,k} = 0 \\ x_{i,j,k}s_{i,j,k} = 0 &\rightarrow X_{i,j,k} * * * S_{i,j,k} = 0 \end{aligned} \qquad (27)$$

as a TNT linear system of equations. Now write the second equation as an underdetermined TNT linear system of equations in which $S_{i,j,k}$ is known and $X_{i,j,k}$ is unknown. But $X_{i,j,k}$ **is** known to a sign, so that we have an underdetermined linear system of equations for which each unknown is merely ± 1.

This system is solved as follows. Multiplying each equation by each unknown in succession, collecting terms, and equating squares of the unknowns to unity, we get a linear system of equations in products of the unknowns. This can then be solved, the solution arranged into a matrix, and a rank one decomposition computed to determine the unknowns. The signs are attached to the known $|O_{i,j,k}|$.

For this to work, we need the number of equations to be at least half the number of unknowns. After multiplying M equations by each of $2M$ unknowns, we have $2M^2$ equations in $\frac{1}{2}$2M(2M–1) unknowns, since $2M$ of these unknowns are known to be unity. Note that without that condition, the larger system is still underdetermined by M. Also note that multiplying each equation by every other equation would

yield an equivalent system of equations, since each equation is a linear combination of the unknowns.

For the X-ray crystallography problem, there is a problem: The number of equations is only one-eighth the number of unknowns, since

- $S_{i,j,k}$ has known $M^{1/3} \times M^{1/3} \times M^{1/3}$ values;
- $O_{i,j,k}$ has unknown $2M^{1/3} \times 2M^{1/3} \times 2M^{1/3}$ signs;
- So #equations is $M^{1/3} \times M^{1/3} \times M^{1/3}$.

But if we increase the number of known $|O_{i,j,k}|$, then

- $S_{i,j,k}$ has known $M^{1/3} \times M^{1/3} \times M^{1/3}$ values;
- $O_{i,j,k}$ has unknown $5M^{1/3} \times 5M^{1/3} \times 5M^{1/3}$ signs;
- So #equations is $4M^{1/3} \times 4M^{1/3} \times 4M^{1/3}$.

Since $5^3 < 2 \cdot 4^3$ we can now use this approach.

G. Example of Computing ± 1

We provide a simple example to illustrate all this. We wish to solve the underdetermined linear system

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 2 & 1 & 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} v \\ w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{28}$$

where each unknown equals ± 1, which is written as

$$v^2 = w^2 = x^2 = y^2 = z^2 = 1 \tag{29}$$

Multiplying both equations by the five unknowns and collecting terms gives the linear system of equations

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 2 & 2 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 & 0 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 1 & 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} vw \\ vx \\ vy \\ vz \\ wx \\ wy \\ wz \\ xy \\ xz \\ yz \end{bmatrix} = - \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \\ 1 \\ 2 \\ 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} \tag{30}$$

We solve this and arrange the solution into a matrix

$$\begin{bmatrix} v^2 & vw & vz & vy & vz \\ wv & w^2 & wx & wy & wz \\ xv & xw & x^2 & xy & xz \\ yv & yw & yx & y^2 & yz \\ zv & zw & zx & zy & z^2 \end{bmatrix} = \begin{bmatrix} v \\ w \\ x \\ y \\ z \end{bmatrix} [v, w, x, y, z] \tag{31}$$

from which we obtain

$$[v, w, x, y, z] = \pm [1, 1, -1, -1, 1] \tag{32}$$

H. Small Numerical Example

Applying this procedure to the previous numerical example would require the following:

- 40 non-zero values of $o_{i,j,k}$, implying as many as 80 nonzero values of $x_{i,j,k}$;
- $S_{i,j,k}$ has 9×9 (80+1 values) support as before;
- Fourier data for 29×29 lowest wavenumbers;
- We have 21^2=441 equations in 29^2=841 unknowns ± 1, where 29=21+9-1;
- Half of the equations and unknowns are redundant, leaving 220 equations in 420 unknowns;
- The linear system is then 87990×87990 where 87990=(420)(419)/2.

This is quite large, so we present a small illustrative example. We are given the following low-wavenumber squared magnitude data, with and without an inclusion at the origin (4×4 2-D DFT):

$$Y1 = \begin{bmatrix} * & * & * & * & * \\ * & 64 & 40 & 0 & * \\ * & 20 & 100 & 20 & * \\ * & 0 & 40 & 64 & * \\ * & * & * & * & * \end{bmatrix} \tag{33}$$

$$Y0 = \begin{bmatrix} * & * & * & * & * \\ * & 81 & 29 & 1 & * \\ * & 29 & 81 & 29 & * \\ * & 1 & 29 & 81 & * \\ * & * & * & * & * \end{bmatrix} \tag{34}$$

- DC (wavenumber=0) is at the center
- $*$ denotes an unknown magnitude value
- Top and bottom rows are unknown and identical
- Leftmost and rightmost columns are likewise
- The object has only three nonzero values

The goal is to reconstruct the object.

For a problem of this size (27) become

$$\begin{bmatrix} E_{0,0} & E_{-1,0} & E_{0,-1} & E_{-1,-1} \\ E_{1,0} & E_{0,0} & E_{1,-1} & E_{0,-1} \\ E_{0,1} & E_{-1,1} & E_{0,0} & E_{-1,0} \\ E_{1,1} & E_{0,1} & E_{1,0} & E_{0,0} \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{35}$$

for computing the indicator function $\begin{bmatrix} w & y \\ x & z \end{bmatrix}$ of the

even part $e_{i,j}$ and

$$\begin{bmatrix} z & y & 0 & x & w & 0 & 0 & 0 & 0 \\ 0 & z & y & 0 & x & w & 0 & 0 & 0 \\ 0 & 0 & 0 & z & y & 0 & x & w & 0 \\ 0 & 0 & 0 & 0 & z & y & 0 & x & w \end{bmatrix} \begin{bmatrix} E_{-1,-1} \\ E_{0,-1} \\ E_{1,-1} \\ E_{-1,0} \\ E_{0,0} \\ E_{1,0} \\ E_{-1,1} \\ E_{0,1} \\ E_{1,1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{36}$$

for computing the signs of the DFT of the odd part $X_{i,j}$ ($|X_{i,j}|$ is known). Note the TNT structure.

The Fourier transform of $e_{i,j}$ is

$$\frac{1}{2}[Y1 - Y0 - 1] = \begin{bmatrix} * & * & * & * & * \\ * & -9 & 5 & -1 & * \\ * & -5 & 9 & -5 & * \\ * & -1 & 5 & -9 & * \\ * & * & * & * & * \end{bmatrix}. \tag{37}$$

The support of $e_{i,j}$ is found by solving the TNT linear system of equations

$$\begin{bmatrix} 9 & -5 & 5 & -9 \\ -5 & 9 & -1 & 5 \\ 5 & -1 & 9 & -5 \\ -9 & 5 & -5 & 9 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{38}$$

which has the solution

$$\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{39}$$

The squared magnitude of the 2-D DFT is

$$\text{DFT} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 2 & 1 & 0 \\ 1 & 2 & 1 & 0 & 1 \\ 2 & 1 & 0 & 1 & 2 \end{bmatrix} \tag{40}$$

The zeros indicate positions of the nonzero $e_{i,j}$. $e_{i,j}$ is then found by solving a linear system and is

$$e_{i,j} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{41}$$

One of the indicated nonzero values of $e_{i,j}$ is in fact zero; this only happens because the problem is so small. Recall that the top and bottom rows are identical to clarify the image symmetry.

The DFT of the odd part $x_{i,j}$ is

$$X^2_{i,j} = -\begin{bmatrix} * & * & * & * & * \\ * & 81-81 & 29-25 & 1-1 & * \\ * & 29-25 & 81-81 & 29-25 & * \\ * & 1-1 & 29-25 & 81-81 & * \\ * & * & * & * & * \end{bmatrix}. \tag{42}$$

We still need the signs of $jX_{i,j}$. $X_{i,j} * * * S_{i,j}$=0 can be solved for $X_{i,j}$, yielding

$$X_{i,j} = j \begin{bmatrix} * & * & * & * & * \\ * & 0 & -2 & 0 & * \\ * & -2 & 0 & 2 & * \\ * & 0 & 2 & 0 & * \\ * & * & * & * & * \end{bmatrix} \tag{43}$$

$x_{i,j}$ is the inverse DFT of this. The final answer is

$$o_{i,j} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{44}$$

VI. ZERO ISOMORPHIC REPLACEMENTS

We now consider the case of no isomorphic replacement. This is the purest form of the X-ray crystallography problem, and its solution saves much laboratory time and money. Direct methods work for crystals with less than about 100 atoms. It is not known whether the problem has a unique solution.

Here we present new equations for (centro)symmetric and non-symmetric objects. These equations at the very least establish sufficient conditions for uniqueness, although we do not claim they are necessary conditions as well.

A. (Centro)symmetric Objects

First we note that if we are fortunate enough to have a heavy atom at the origin (point of symmetry), then the DFT values are all real and non-negative, so we just use the superresolution algorithm.

In general, we must recover the signs of $O_{i,j,k}$ from $|O_{i,j,k}|$. In principle we could try all possible combinations of signs and see which one yields a singular TNT matrix, since there is only a finite number of possibilities. Generically, only one choice will work, since $O_{0,0,0} > 0$.

This does establish a sufficient (and likely necessary) condition for unique reconstruction: Knowledge of $\{|O_{i,j,k}|, |i|, |j|, |k| \leq M^{1/3}\}$, and a procedure for computing it. A *practical* procedure is not known (although see below).

B. Non-Symmetric Objects

The problem is to compute the Fourier phases $\theta_{i,j,k}$ from the magnitudes $|O_{i,j,k}|$, for low-wavenumber data only, using the sparsity condition. In principle this can be done as follows.

Let $n_{i,j,k}$ (rather than $s_{i,j,k}$) be the indicator function for nonzero values of $e_{i,j,k}$ and $x_{i,j,k}$, which are nonzero for (at most) $(2M)^3$ (not 2M) locations. Let $c_{i,j,k} = \cos(\theta_{i,j,k})$ and $s_{i,j,k} = \sin(\theta_{i,j,k})$. Then

$$\begin{bmatrix} |O_{i,j,k}|c_{i,j,k} & \dots \\ \vdots & \ddots \end{bmatrix} \begin{bmatrix} n_{0,0,0} \\ \vdots \\ n_{M,M,M} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \tag{45}$$

$$\begin{bmatrix} |O_{i,j,k}|s_{i,j,k} & \dots \\ \vdots & \ddots \end{bmatrix} \begin{bmatrix} n_{0,0,0} \\ \vdots \\ n_{M,M,M} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \tag{46}$$

$$c_{i,j,k}^2 + s_{i,j,k}^2 = 1 \tag{47}$$

This is an overdetermined system of simultaneous quadratic equations in the unknowns $c_{i,j,k}$, $s_{i,j,k}$ and $n_{i,j,k}$. Generically, it will have as its only solution the correct values of the unknowns, to sign ambiguities. These sign ambiguities create the well-known reversal ($o_{i,j,k}$ vs $o_{N-i,N-j,N-k}$) and sign ($\pm o_{i,j,k}$) ambiguities of phase retrieval. Of course, $o_{i,j,k} \geq 0$.

This establishes a sufficient condition for uniqueness. Of course, solving this system of simultaneous quadratic equations is another matter. Continuation or homotopy methods, in which an easily-solved system is perturbed towards the desired system, with the solution updated using gradient methods at each step, can work, but require much time and computation. On the other hand, existing algorithms for X-ray crystallography also require much time and computation, as well as extensive human interaction.

Another approach would be to form the TNT matrix of $O_{i,j,k}$ using initial guesses for the phases, and gradually perturb the TNT matrix towards singularity using structured total least squares. This would require a large amount of time and computation.

Note that for symmetric objects we have $s_{i,j,k}$=0 and $c_{i,j,k}$=$\pm$1. More importantly, $o_{i,j,k}$ and $e_{i,j,k}$ have the same supports, so the number of data points $|O_{i,j,k}|$ need only exceed the number of nonzero $o_{i,j,k}$ by a factor of 8, not 16.

VII. GLOSSARY

X-RAY CRYSTAL.	SIGNAL PROC.
Electron density	image or object
Crystal structure	space-periodic
Unit cell lengths	spatial periods
Atomicity	object sparsity
$P\bar{1}$ group	even symmetry
Reciprocal space	Fourier domain
Structure factors	$\mathcal{F}\{\text{object}\}$
Structure amplitude	$\lvert\mathcal{F}\{\text{object}\}\rvert$
Patterson map	autocorrelation
Karle-Hauptman determinants$\geq$0	Circulant matrix pos.semi-definite
Sayre equation	$\mathcal{F}\{o(x)(o(x)-1)\}$
Isomorphous replacement	Inserting atoms into the object
Anomalous dispersion	Vary λ excite heavy atoms
Crystallographic symmetry	Rotation invariant within the lattice
Noncrystallographic symmetry	Rotation invariant extending lattice
Born approx.	Linearization

References

[1] J. Karle and H. Hauptman, "The phases and magnitudes of the structure factors," *Acta Crystal.* 3, 181-187, 1950.

[2] M.M. Woolfson, "Direct methods in crystallography," *Rep. Prog. Phys.* 34, 369-434, 1971.

[3] R.P. Millane, "Phase retrieval in crystallography and optics," *J. Opt. Soc. Am. A* 7, 394-411, 1990.