Phase retrieval of diffraction from highly strained crystals

Marcus C. Newton*
Advanced Technology Institute, University of Surrey, Guilford, Surrey, GU2 7XH, United Kingdom

Ross Harder
Argonne National Laboratory, Argonne, Illinois 60439, USA

Xiaojing Huang, Gang Xiong, and Ian K. Robinson
London Centre for Nanotechnology, University College London, WC1H 0AH, United Kingdom

(Received 18 August 2010; published 20 October 2010)

An important application of phase retrieval methods is to invert coherent x-ray diffraction measurements to obtain real-space images of nanoscale crystals. The phase information is currently recovered from reciprocal-space amplitude measurements by the application of iterative projective algorithms that solve the nonlinear and nonconvex optimization problem. Various algorithms have been developed each of which apply constraints in real and reciprocal space on the reconstructed object. In general, these methods rely on experimental data that is oversampled above the Nyquist frequency. To date, support-based methods have worked well, but are less successful for highly strained structures, defined as those which contain (real-space) phase information outside the range of ±π/2. As a direct result the acquired experimental data is, in general, inadvertently subsampled below the Nyquist frequency. In recent years, a new theory of “compressive sensing” has emerged, which dictates that an appropriately subsampled (or compressed) signal can be recovered exactly through iterative reconstruction and various routes to minimizing the ℓ₁ norm or total variation in that signal. This has proven effective in solving several classes of convex optimization problems. Here we report on a “density-modification” phase reconstruction algorithm that applies the principles of compressive sensing to solve the nonconvex phase retrieval problem for highly strained crystalline materials. The application of a nonlinear operator in real-space minimizes the ℓ₁ norm of the amplitude by a promotion-penalization (or “propenal”) operation that confines the density bandwidth. This was found to significantly aid in the reconstruction of highly strained nanocrystals. We show how this method is able to successfully reconstruct phase information that otherwise could not be recovered.

DOI: 10.1103/PhysRevB.82.165436 PACS number(s): 87.59.-e, 03.75.Hh, 71.15.Dx, 73.22.-f

I. INTRODUCTION

In a number of disciplines including x-ray crystallography,1 sparse-aperture imagery,2 and astronomy,3 phase information is lost as only intensity measurements can be made. It is therefore necessary to reconstruct the phase information subject to some assumed constraints imposed on the object. Various approaches have been taken to solve the classical phase retrieval problem and have benefited from advances in optimization theory. All these methods operate by the application of iterative projective algorithms that converge toward the optimal solution by applying constraints to the objective function at each iteration.4–8 A well-developed and widely used method is the application of a real-space support region (a subset of the function space) within which the object is confined. Here the objective function is optimized within some local region instead of the entire function space. This often results in an appreciable real-space solution in fewer projections as the distance of each projection is confined. The main disadvantage is that this method may not find the true globally optimal solution.

Other methods include solvent flattening9 and flipping in which the complex objective function is confined to a subset of the complex plane. Both methods can however struggle to reproducibly reconstruct the objective function and several attempts are often required to verify the result. The Sayre equation is an example of a nonlinear real-space constraint in which the real-space amplitude is projected onto its squared value. For high-resolution structures, it forms the basis of a powerful set of crystallographic “direct methods.”

The Shannon-sampling theorem dictates that to correctly reconstruct a signal containing component frequencies no greater than B from a series of discrete samples, the sampling interval must be greater than or equal to 1/2B. Sampling at a lower frequency than this will lead to aliasing within the reconstructed signal as ambiguities arise between harmonic Fourier components.

The theorem due to Candés et al.10 however, states that under certain conditions it is possible to reconstruct a sparse signal from a limited number of samples provided that a normalization constraint is applied to the object. The ℓ₁-norm plays a central role in this theorem and is defined as the sum of the absolute value of the components of a signal. Specifically, let ρ₀(x) ∈ ℂᴺ be the original function on discrete set S with Fourier transform ˆρ₀(k), where x ∈ ℤᴺ. Then, exact reconstruction is possible by the reverse Fourier transform if all the Fourier components are known. Suppose, however, that only a subset of components ˆρ₀|Ω are known and randomly sampled on some set Ω. The theorem predicts that it is then possible to recover ρ₀(x) exactly from ˆρ₀(k), where k ∈ Ω and |Ω| < |S| = N. We do so by minimizing the ℓ₁ norm of the recovered signal ρ such that...
method and continue by showing the results of simulated and experimental data for our algorithm.

II. ITERATIVE PHASE RETRIEVAL

When coherent x-ray diffraction measurements are performed phase information is lost and instead we obtain only the reciprocal-space scattering intensity distribution $I(q)$. This is closely approximated by the square of the modulus of the Fourier transform of the object function $\hat{\rho}_0(q)$ such that

$$I(q) = |\hat{\rho}_0(q)|^2 = |\hat{\rho}_0(q)e^{i\phi(q)}|^2.$$  

In general, we reconstruct the object with complex electron density $\rho(r)$ and retrieve the phase information by application of a number of constraints in direct and Fourier space. Reconstruction then proceeds by traversing back and forth between direct and Fourier space while applying these constraints at each turn. A successful approach is to define a subset or support region $S$ in direct space for which the amplitude of the object density is unrestricted. Outside this region the amplitude is either set to zero or minimized in some way. The Fourier space constraint, enforced by the operator $P_{|q|}$, requires that the amplitude of the object is equal to that of the original measurement on some set $M$, a subset of the Fourier space

$$P_{|q|}|\hat{\rho}(q)|e^{i\phi(q)} = |\hat{\rho}_0(q)|e^{i\phi_0(q)}, \quad \forall q \in M.$$  

Various algorithms that expand on these basic constraints have been proposed and are distinguished by their choice of real-space constraint. Particular examples are the hybrid input-output (HIO), hybrid projection-reflection (HPR), and the relaxed averaged alternating reflections (RAAR) algorithms. They all however consist of projections onto the support $S$ and modulus $M$ set. A well-known example is the HIO algorithm which can be written as

$$\rho^{(n+1)}(r) = \begin{cases} P_M\rho^{(n)}(r) & \forall r \in S \\ (I - \beta P_M)\rho^{(n)}(r) & \text{otherwise}, \end{cases}$$  

where $n$ is the iterate, $P_M = F^{-1}P_{|q|}F$ is the projection operator, $I$ is the identity, and $\beta \in [0,1]$ is a relaxation parameter. The direct and inverse Fourier transforms are represented by $F$ and $F^{-1}$, respectively.

III. NORMALIZED DENSITY PHASE RETRIEVAL

As the modulus projection operator $P_M$ [Eq. (3)] is non-convex, the phase retrieval problem is also in general non-convex preventing the application of direct optimization methods such as primal-dual algorithms. For this reason, the $\ell_1$-norm minimization constraint cannot be applied directly to the phase retrieval problem. Instead the algorithm that we propose here promotes $\ell_1$-norm minimization of the real-space density $\rho$ by the application of a promotion-penalization (or propenal) projection operator $P_{\text{ND}}^{(n)}$. The projection operator is defined as follows:
Similarly, those in the neighborhood of zero at the current iteration. Amplitude values that fall in the function with equal full width half maximum.

The original curves are shown with open circles ($\bigcirc$). The $P_{\text{ND}}$ operator is applied to each curve $m$ times where $m=\{1, 10, 100\}$ for symbols ($\times$), (+), and (■), respectively. [(a) and (c)] When $\nu_u$ and $\nu_l$ are sufficiently small (<0.01, here), each function converges toward a square function with equal full width half maximum. [(b) and (d)] Otherwise, a wider square function is obtained.

\[
P_{\text{ND}}^{(m)} \rho(\mathbf{r}) = \begin{cases} 
0 & |\rho(\mathbf{r})| \leq \nu_l |\rho|_{\text{max}} \\
|\rho(\mathbf{r})|_{\text{max}} & |\rho(\mathbf{r})| \geq (1 - \nu_u) |\rho|_{\text{max}} \\
\rho(\mathbf{r}) & \text{otherwise} 
\end{cases}
\]

where $\nu_u$ and $\nu_l$ are upper and lower thresholds such that $(\nu_u + \nu_l) \in [0, 1]$ and $|\rho|_{\text{max}}$ is the maximum amplitude of $\rho(\mathbf{r})$ at the current iteration. Amplitude values that fall in the neighborhood of $|\rho|_{\text{max}}$ [i.e., $|\rho| > (1 - \nu_u) |\rho|_{\text{max}}$] are set to this value. Similarly, those in the neighborhood of zero [i.e., $|\rho| < \nu_l |\rho|_{\text{max}}$] are set to zero. The remaining points are scaled to maintain a smooth distribution. This operator is judiciously applied $m$ times per iterate $n$ and after any additional real-space constraints are applied such as those defined by algorithms mentioned above. In doing so, we form a new family of density normalizing algorithms. For example, the normalized density operator [Eq. (5)] combined with the HIO operators [Eq. (4)] yield the normalized density hybrid input-output (NDHIO) algorithm.

Figure 2 shows the operator applied to normal Gaussian and Lorentzian functions. The $P_{\text{ND}}$ operator [Eq. (5)] is applied to each curve $m$ times where $m=\{1, 10, 100\}$ for symbols ($\times$), (+), and (■), respectively. The original curves are shown with open circles ($\bigcirc$). In the limit as the number of iterations $m$ tends to $\infty$, each function tends to a square function as the amplitude becomes uniform. When $\nu_u$ and $\nu_l$ are sufficiently small (<0.01, here), each function converges with a full width half maximum equal to the original function [Figs. 2(a) and 2(c)]. Otherwise, a divergent square function is obtained [Figs. 2(b) and 2(d)]. In the following we show how the application of this additional constraint can facilitate the reconstruction of highly strained materials.

**IV. SIMULATED PHASE RECONSTRUCTION OF HIGHLY STRAINED OBJECT**

We first defined a complex tapered hexagonal prism-shaped object with uniform density on a function space of $128 \times 128 \times 128$ voxel units. The dimensions of the object are 80 units in length and facet-to-facet widths of 40 and 20 units at either end. Employing a tapered object with reduced symmetry aids in identifying twinning in the phase reconstruction and avoids aliasing within the Fourier transform array. We then gave that object complex phase information $\phi$ as described by the following function:
Here wrapping occurs when the phase is mapped onto the complex plane. A relaxation parameter of $\beta = 0.9$ was used for the HIO algorithm. For the case of the NDHIO algorithm the parameter $m$ started at 0 for the first 50 iterations. It was then incremented from 1 to 10 in 30 iteration intervals after which the algorithm continued until 500 iterations were completed at this value. $r_u$ and $r_v$ take values of 0.1 and 0.027, respectively.

Reconstructions using the HIO algorithm are shown in Figs. 3(b), 3(e), and 3(h). The resulting object bears little resemblance to the original object in terms of both amplitude and phase. The objects amplitude varies significantly from point to point. This “pebbling effect” is the familiar result of failed reconstruction either from bad quality data, insufficiently oversampled data, or those associated with large strain. Similar results were obtained using the HPR and RAAR algorithms. Thus far we have been unable to adequately reconstruct objects containing phase variations extending beyond $\pm \pi/2$ using these more traditional algorithms for reasons outlined above.

Amplitude reconstruction using the NDHIO algorithm is shown in Fig. 3(c). A clear morphological resemblance can be seen between this and the original image. The dimensions of the reconstructed object closely match those of the original object. Cross sections of the phase information are also shown in Figs. 3(f) and 3(i). The recovered phase information also resembles that of the original image with an arbitrary phase offset. This implies that the NDHIO algorithm converges toward the correct solution. Few noisy voxels appear in the support region away from the object. This is a property of the $P_{NDHIO}^m$ operator that, in the limiting case, will set to zero voxels below $r_u|\hat{\rho}|_{max}$.

Figure 4 shows the convergence of the error metric for both the HIO and NDHIO algorithms. The error metric is defined as, $\Sigma_q |\hat{\rho}(q) - \tilde{\rho}_0(q)|^2/\Sigma_q |\tilde{\rho}_0(q)|^2$. Both algorithms converge within few iterations to below 0.1. The HIO algorithm, however, appears to oscillate around $\sim 0.05$. This effect persists up to 3000 iterations (not shown in figure) and little improvement in the reconstruction is observed. The

\[
\phi(x_i) = \phi_0 \prod_{i=1}^{i=3} \cos \left( \frac{\pi x_i}{L_i} \right),
\]

where $\phi_0$ is the maximum phase value, $x_i$ is the ordinate, and $L_i$ is the length of the object along the orthogonal axis $i$. The object was then Fourier transformed and all Fourier space phase information was removed. Phase reconstruction of the object was then attempted using the amplitude Fourier space information and random real-space phase information as a starting point. For values of $|\phi_0|$ greater than $\pi$, phase wrapping occurs when the phase is mapped onto the complex plane. Here $\phi_0=5.0$ radians. Figure 3 shows the amplitude (a) and phase [(d) and (g)] of the original object. Phase wraps are visible in the phase between the red and blue regions.

Phase recovery was carried out with both the HIO and NDHIO algorithms. A rectangular support region $S$ of $80 \times 80 \times 110$ units was used in each reconstruction. In each case, a total of 500 iterations was used for phase reconstruction. A relaxation parameter of $\beta = 0.9$ was used for the HIO algorithm.
NDHIO algorithm converges more smoothly toward an acceptable solution within 500 iterations. Small peaks can be observed at periodic intervals and result from changes in the value of \( m \).

V. COHERENT X-RAY DIFFRACTION MEASUREMENTS

Coherent x-ray diffraction measurements were performed on a nanometer scale strip of silicon-on-insulator (SOI). The strips of silicon were patterned from bulk material using electron-beam lithographic processing which utilized a reactive ion etching (RIE) step to remove excess silicon. The \( \text{SF}_6 \) RIE conditions used cause a strong undercutting of the wire which is seen in the resulting images (Fig. 6). Samples prepared in this manner were found to possess a strong-phase structure that might be related to defects such as dislocations at intervals along the length of the strip. Coherent x-ray diffraction data were measured at beamline 34ID-C at the Advanced Photon Source. A 0.3°-wide rocking scan was made through the off-specular \{111\} Bragg peak, recording a stack of two-dimensional slices of diffracted intensity onto a direct-detection charge coupled device (data not shown). Reconstruction was then attempted using the HIO and NDHIO algorithms. The SOI strip is approximately 2300 × 700 × 600 nm\(^3\) and occupies 90 × 46 × 23 voxels of data. Reconstruction was carried out using a shrink-wrap support method described in detail in Ref. 1. This began with 100 iterations of HIO after which two shrink-wrap loops consisting of ten iteration of HIO followed by two iterations of the error-reduction (ER) algorithm were performed. Finally 50 iterations of HIO and 50 iterations of phase-only ER (Ref. 17) were applied. For normalized density reconstruction, 300 iterations of the NDHIO algorithm (\( \nu_u=0.005 \), \( \nu_l=0.0005 \), and \( m=1–30 \)) were applied after the HIO shrink-wrap stage followed by 50 iterations phase-only ER. To avoid amplification of noise during normalization, a median noise filter with kernel size of 5 × 5 × 5 units was applied to the real-space amplitude at ten iteration intervals each time \( m \) was incremented during the NDHIO iterations (see Fig. 5).

Figure 6 shows the reconstructed image of the SOI strip obtained using the HIO and the NDHIO algorithms with the length of the amplitude and phase information portrayed in the usual way: the amplitude as an “isosurface” and phase as a color “cut-plane” contour map. A 75% (50%) isosurface is shown for the HIO (NDHIO) algorithms, respectively. The HIO algorithm fails to completely reconstruct the amplitude of the object for reasons described above and is akin to the effects seen in Fig. 3(b). More severely, this causes incorrect phase information to be obtained. This effect is clearly seen in Fig. 6(c) where phase ripples follow variations in amplitude of the object. Reconstruction with the NDHIO algorithm results in improved uniformity in the amplitude while maintaining the correct morphology [Fig. 6(b)]. This results in-phase information that is unhindered by artifacts in the objects amplitude [Fig. 6(d)], in contrast to the HIO reconstruction.

FIG. 6. (Color online) Amplitude and phase reconstruction of SOI. Reconstruction using the HIO [(a) and (c)] and NDHIO algorithm of Eq. (5) [(b) and (d)]. Missing amplitude due to phase wrapping occurs in the HIO reconstruction (circled in red). This influences the resulting phase information yielding incorrect phase reconstruction. This is alleviated when using the NDHIO algorithm which correctly reconstructs the objects amplitude. Isometric projections of the amplitude are shown in (i) while side views are shown in (ii).
VI. SUMMARY

In summary, the density-modification function we propose here minimizes the $\ell_1$ normalized density and therefore seeks a solution with minimum variation in amplitude (as opposed to the minimum-energy solution). This allows for iterative reconstruction of phase discontinuities common in highly strained materials. This method was combined with more conventional algorithms and resulted in a significantly improved amplitude reconstruction leading to the accurate reconstruction of phase information. This work therefore represents a step toward understanding the failure mode of iterative algorithms when applied to strong-phase structures. Implementation is comparatively simple, in relation to convex-optimization methods such as the primal dual and interior point, yet is able to provide high-quality reconstruction and also has the advantage of being less computationally demanding. Using data obtained from third- and fourth-generation synchrotron facilities, we envisage this method to be useful for phase reconstruction of strained periodic crystals.

ACKNOWLEDGMENTS

This work was partly supported by European Research Council under “Advanced” Grant No. 227711. The experimental work was carried out at APS beamline 34-ID-C, built with funds from the U.S. National Science Foundation under Grant No. DMR-9724294 and operated by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.

---

* m.newton@surrey.ac.uk