Experimental demonstration of coupled multi-peak Bragg coherent diffraction imaging with genetic algorithms

Matthew J. Wilkin, Siddharth Maddali, Stephan O. Hruszkewycz, Anastasios Pateras, Richard L. Sandberg, Ross Harder, Wonsuk Cha, Robert M. Suter, and Anthony D. Rollett

Department of Materials Science and Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA
Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA
Brigham Young University, Provo, Utah 84602, USA
Advanced Photon Source, Argonne National Laboratory, Argonne, Illinois 60439, USA

(Received 3 March 2021; revised 14 May 2021; accepted 20 May 2021; published 1 June 2021)

Bragg coherent diffraction imaging has the potential to provide significant insight into the structure-properties relationship for crystalline materials by imaging, with nanoscale resolution, three-dimensional strain fields within individual grains and nanoparticles. The capability of present-day synchrotrons to locate and measure a multiplicity of Bragg reflections from a single grain makes it possible to recover the full strain tensor with nanometer resolution. Recent methods for coupling reconstructions from several peaks to determine the strain tensor have been developed and applied to synthetic data, but have not been applied to experimental data. Here, using a coupled genetic reconstruction algorithm, we reconstruct an experimental data set and demonstrate improvements in the ability to resolve vector-valued displacement fields internal to the particle as compared to what is achieved with a noncoupled approach. The coupled approach developed in this work was also validated on simulated data sets. In both simulated and experimental data, reconstructions from our coupled Bragg peak algorithm show improvements over the noncoupled independent reconstruction method of 5% in terms of accuracy and 53% in terms of consistency.

DOI: 10.1103/PhysRevB.103.214103

I. INTRODUCTION

Illuminating the structure-property relationship in crystalline materials remains a primary objective for the crystallographic community. To this end, many x-ray and electron diffraction techniques have been developed to probe fundamental mechanisms at the nanoscale [1–3]. To reveal strain inhomogeneity and defect fields within crystals, several such methods require the recovery of phase information, which is lost during the measurement of diffracted intensity on a detector. Various computational phase-retrieval techniques have been developed to efficiently solve the general “phase-problem” for scattering from finite-sized objects [4].

Among the techniques to benefit from progress in these fields is Bragg coherent diffraction imaging (BCDI), an experimental tool for nondestructive imaging of the shape and displacement fields of a three-dimensional (3D) crystal from a coherent Bragg peak [2]. Recent work demonstrated its viability for the study of dislocation dynamics, stacking faults, and twinning at the nanoscale through reconstruction of a single Bragg peak [5–12]. Tools such as Laue diffraction [13] and high energy diffraction microscopy (HEDM) [14,15] have shown potential to locate neighboring grains and nanoparticles of interest on the length scales required for BCDI opening the door for multi-grain experiments. With the enhanced x-ray brightness from upcoming fourth-generation synchrotrons, this technique will be capable of achieving significant improvements in throughput and spatial resolution, with the potential to help validate theoretical models, such as plasticity, on atomistic and mesoscopic length scales [16–18].

The phase retrieval process applied to a BCDI data set measured at a single Bragg reflection returns a 3D image sensitive to a single component of the lattice distortion field, from which the corresponding strain projection can be calculated [19,20]. Three or more BCDI data sets from noncoplanar Bragg reflections contain enough information to determine the entire displacement vector field, and by extension, the full six-component strain tensor [21,22]. Though this has typically been accomplished through independent reconstruction of each Bragg peak, recently developed phase retrieval methods proposed coupling multiple BCDI data sets [23,24], sharing information among peaks within the image reconstruction algorithm. While reconstructing Bragg peaks independently is more straightforward in practice, coupling reconstructions from multiple Bragg peaks from the same crystal has the potential to significantly reduce the error of the retrieved strain field and crystal shape by increasing the number of constraints on the reconstructed object. These methods have proven effective in the reconstruction of simulated data sets, but as yet, none have been used on experimental data, which typically present more challenges than simulated data [25]. Here we demonstrate that a coupled Bragg peak phase retrieval approach can successfully be applied to reconstruct displacement vector fields in three dimensions from experimental data and that this approach provides improvements as compared to the noncoupled independent reconstruction approach in terms.
of robustness and error, marking a significant step in moving the method towards practical use.

II. BACKGROUND: CONVENTIONAL PHASE RETRIEVAL

In BCDI, a compact single crystal is illuminated by a nominally monochromatic and spatially coherent x-ray beam. As the crystal is rotated through one of its Bragg conditions, corresponding to reciprocal lattice vector \( \mathbf{G}_{hkkl} \), the 3D coherent diffraction pattern is sampled in a series of area detector images with the detector in the optical far-field or Fraunhofer regime [20]. In this limit, the complex scattered wave amplitude \( \tilde{\psi}(\mathbf{q}) \), with \( \mathbf{q} = \mathbf{Q} - \mathbf{G}_{hkkl} \) being the deviation from the nominal Bragg condition in reciprocal space, can be approximated as the Fourier transform (FT) of the exit wave field at the location of the particle.

The real-to-reciprocal space relationship can be denoted in shorthand as

\[
\psi(\mathbf{q}) = \text{FT}[\psi(\mathbf{x})].
\]

Given that BCDI measurements consist of pixelated detector images measured at regular angular intervals, the fields \( \psi(\mathbf{x}) \) and \( \tilde{\psi}(\mathbf{q}) \) are discrete and span a finite range in \( \mathbf{x} \) and \( \mathbf{q} \), and we designate this discretization (or voxelization) of these spaces with the subscript \( n \), (i.e., \( x_n \)). Thus the we define the discrete complex-valued sample exit field corresponding to a given Bragg peak as

\[
\psi(x_n) = \rho(x_n) e^{2\pi i \phi(x_n)}. \tag{2}
\]

The amplitude \( \rho(x_n) \) corresponds to the electron density, and the phase \( \phi(x_n) \) is a projection of the atomic displacement \( \mathbf{u}(x_n) \) with respect to the unperturbed lattice along the scattering vector \( \mathbf{G}_{hkkl} \):

\[
\phi_{hkkl}(x_n) = \mathbf{u}(x_n) \cdot \mathbf{G}_{hkkl}. \tag{3}
\]

The measured intensity of a Bragg peak is related to \( \tilde{\psi}(\mathbf{q}) \) by

\[
I(\mathbf{q}) = |\tilde{\psi}(\mathbf{q})|^2. \tag{4}
\]

To find the phase and amplitude of \( \psi(\mathbf{x}) \), it is necessary to recover the lost phases of \( \psi(\mathbf{q}) \), which is accomplished via iterative phase retrieval. In this work, we utilize phase retrieval algorithms that are broadly used in the BCDI field, namely error reduction (ER) [26] and hybrid input-output (HIO) [27]. The approach entails making an initial guess for \( \psi(\mathbf{x}) \) and utilizing a combination of ER and HIO to subsequently improve the guess of \( \psi(\mathbf{x}) \) by iteratively transforming the object between real space and Fourier space, enforcing constraints in each domain until the solution converges with respect to the error metric \( \chi^2 \), given as

\[
\chi^2 = \sum_{n=1}^{N} (\sqrt{T_n} - |\hat{\psi}_n|)^2, \tag{5}
\]

where \( n \) denotes a sampled point in reciprocal space and \( N \) is the total number of sampled points.

In the iterative ER and HIO routines, constraints are applied in both reciprocal and real spaces. The Fourier space constraint consists of setting the forward-modeled Fourier space amplitude \( |\hat{\psi}_n| \) equal to the measured amplitude \( \sqrt{T_n} \).

The real space constraint is applied by cropping the real space modulus [4] by a support function \( S \)

\[
\psi(x) = 0 \quad \text{if} \quad x \notin S. \tag{6}
\]

The support function is allowed to shrink over the course of the phase retrieval procedure by a process called “shrinkwrap” [28], wherein the object amplitude in real space, \( |\psi(x)| \) is convolved with a Gaussian \( f \),

\[
\xi(x) = (f \ast |\psi(x)|), \tag{7}
\]

and thresholded, according to

\[
S = \begin{cases} 1, & \text{if} \quad \xi(x) > (th \times \xi_{\text{max}}), \\ 0, & \text{otherwise}, \end{cases} \tag{8}
\]

where \( th \) is the threshold value as a fraction of the maximum value in \( \xi(x) \). Together, HIO, ER, and shrinkwrap make up the foundation of the methods presented here, together with a strategy for coupling information from multiple Bragg peaks during the reconstruction process.

III. COUPLED PHASE RETRIEVAL WITH GENETIC ALGORITHMS

The coupled phase retrieval (CPR) approach we developed and deployed on simulated and experimental data is described in this section. The strategy is composed of two major components: 1) an adaptation of the BCDI ER/HIO phase retrieval strategy in which a mechanism of coupling information from multiple Bragg peaks is introduced, 2) and the utilization of multiple instances of this modified ER/HIO strategy within the framework of a genetic algorithm, which is known to provide robustness against stagnation in the context of experimental data.

A. Coupled phase retrieval with ER/HIO

A flow chart of the CPR algorithm utilized here is pictured in Fig. 1(a). A CPR reconstruction consists of \( J \) parallel constituent reconstructions (hereafter called “constituents”) of \( J \) Bragg peaks which have been interpolated onto a uniform grid and rotated into a common laboratory frame using the approach defined by Maddali et al. [20]. A sequence of fixed point iteration steps consisting of ER and HIO (hereafter called a “recipe”) is used for each constituent. Every 20 iterations, the phase and amplitude values from each constituent \( \phi_j(x) \) and \( \rho_j(x) \), are combined to calculate universal values \( \overline{\phi}(x) \) and \( \overline{\rho}(x) \), which are defined below. These values are then used to seed the new phase and amplitude values \( \hat{\phi}_j(x) \) and \( \hat{\rho}_j(x) \) for the next 20 iterations of each constituent.

During the combination step, the universal amplitude \( \overline{\rho}(x) \) is calculated by averaging the \( \hat{\rho}_j(x) \) fields returned from each peak according to

\[
\overline{\rho}(x) = \frac{\sum_{j=1}^{J} \hat{\rho}_j(x)}{J}. \tag{9}
\]

where \( \hat{\rho}_j(x) \) is the real space amplitude for constituent \( j \), normalized by its sum over all voxels. At specified iterations, shrinkwrap [Eq. (8)] is applied to this universal amplitude \( \overline{\rho}(x) \) [rather than to each constituent amplitude \( \rho_j(x) \)], and the universal support \( S \) is updated. This new support function is
then used as the new real-space constraint for each constituent at all successive ER and HIO steps until shrinkwrap is applied again. This support refinement process effectively selects the “fittest” individual is bred with all others, updating their genotypes and creating a new generation of individuals. This process proceeds until the fittest individual sees negligible improvement from generation to generation. The work by Ulvestad et al. describes a genetic algorithm in which each individual is a reconstruction of the same Bragg peak, with \( \psi(x) \) being the genotype. The initial population of individuals in the first generation are seeded with a random-phase starting guess and a cube support, and phase retrieval is applied to each. The individual with the lowest \( \chi^2 \), Eq. (5), after phasing is selected as the “fittest” and bred with all other individuals according to

\[
\psi_m^{i+1} = \sqrt{\psi_m^i(x) * \psi_m^0(x)},
\]

where the subscript \( f \) denotes the fittest individual, the subscript \( m \) denotes one of the other individual reconstructions within generation \( i \). Then, \( \psi_m^{i+1} \) is used as the initial guess for individual \( m \) and another round of phase retrieval is applied. This process is repeated for a specified number of generations, after which, the fittest individual is chosen as the final reconstruction. The state-of-the-art for independent phase retrieval (IPR) of multiple Bragg peaks consists of reconstructing each peak using this GA and finding the common \( u(x) \) and \( \bar{\rho}(x) \) from the final reconstructions using Eqs. (9) and (10), respectively. We will call this approach “IPR-GA.”

In this work, we modify the above GA routine so that each individual is a CPR reconstruction and breeding is performed on the quantities \( \bar{\rho}(x) \) and \( u(x) \). We will call this approach “CPR-GA.” A flow chart showing the CPR-GA is shown in Fig. 1(b) for three individuals bred for three generations. For the first generation in this example, each of the three individuals is seeded with a random-phase starting guess and a cube object support on all its constituents. The CPR routine is executed on each individual, yielding three amplitudes \( \bar{\rho}(x) \) and displacement fields \( u(x) \). The individual with the lowest \( \chi^2 \) (summed across all constituents) is selected as the fittest, and its amplitude and displacement are bred with all other individuals to seed the next generation of CPR reconstructions. Thus, the amplitudes and displacement fields of the \((i+1)\) generation of individual \( m \) are defined by the

\[
\chi^2_{\rho_j} = \sum_j \chi^2_{\rho_j} \quad \text{(summed across all constituents)}
\]
breeding functions
\[
\hat{\rho}^{i+1}_m(\mathbf{x}) = \sqrt{\hat{\rho}^i_m(\mathbf{x}) \ast \hat{\rho}^i_f(\mathbf{x})}
\]
(12)
and
\[
u^{i+1}_m(\mathbf{x}) = \frac{\mathbf{u}^i_m(\mathbf{x}) + \mathbf{u}^i_f(\mathbf{x})}{2}.
\]
(13)
This pattern repeats for a specified number of generations until \(\chi^2\) converges.

IV. METHODS

The CPR-GA was validated using both simulated and experimental data sets. For the simulated data set, a 3D shape function of a nominally 400-nm gold nanoparticle was extracted from a grain growth simulation performed with the SPARKS Monte Carlo Potts model [32] and decorated with a \(\mathbf{u}(\mathbf{x})\) field. Four Bragg reflections, \((1, 1, 1), (1, 1, \bar{1}), (1, \bar{1}, 1),\) and \((\bar{1}, 1, 1)\), were forward-modeled and used for validation. The peak intensity was scaled to a max photon count of \(10^7\) and Poisson noise was then added on a pixel-by-pixel basis.

This data set was reconstructed using the IPR-GA and the new CPR-GA. For the IPR-GA, five generations of eight individuals were used, with the population being culled to the four best individuals after two generations. For each generation, all individuals were reconstructed using the following phase retrieval recipe, which consisted of:

1. 100 iterations of ER, update support after iterations 60 and 100;
2. 60 iterations of HIO;
3. 80 iterations of ER, update support after iterations 40;
4. 60 iterations of HIO;
5. 100 iterations of ER, update support after 20 iterations;

At the end of each generation, the displacement was calculated using the least-squares approach in Eq. (10) with the fittest individuals from each peak.

For the CPR-GA, the recipe above was used for all constituents to evolve a single generation for each individual. For each CPR reconstruction in the CPR-GA, the universal values \(\mathbf{u}(\mathbf{x})\) and \(\hat{\rho}(\mathbf{x})\) were updated every 20 iterations during the recipe, and, at specified iterations, the support update operation was applied to the global support \(\mathbf{S}\). Five generations of eight individuals were used for the CPR-GA, with the population being culled to the four best individuals after two generations.

An experimental data set was collected at the 34-ID-C beamline at the Advanced Photon Source (APS). Using Laue diffraction, four reflections, \((1, \bar{1}, \bar{1}), (1, 1, \bar{1}), (1, \bar{1}, 1),\) and \((2, 0, 0)\), were found for a Au nanocrystal de-wetted onto a Si(100) substrate [33]. BCDI measurements were taken for a Si(100) substrate [33]. BCDI measurements were taken for a specified number of generations until \(\chi^2\) converges.

A plot of confidence metrics averaged over all 40 runs is shown in Fig. 3. The CPR-GA reconstruction shows lower \(\chi^2\) in both \(\chi^2_{\text{tot}}\) and \(\chi^2_{\text{tot}}\) showing an improvement of 3.87 \(\times\) \(10^{-2}\) in \(\chi^2_{\text{tot}}\) and 6.66 \(\times\) \(10^{-2}\) in \(\chi^2_{\text{tot}}\) or 39.2% and 96.7%, respectively. An \(L\) map of the best reconstruction from each method can be seen in Fig. 3. The CPR-GA reconstruction shows lower \(L\) at nearly every voxel.

V. RESULTS AND DISCUSSION

A. Simulated data

To assess and compare the overall efficacy of the CPR-GA approach versus the IPR-GA approach over repeated instantiations, the simulated data sets were reconstructed 40 times using the IPR-GA and 40 times using the CPR-GA, each with different random starting guesses. In each of the 40 reconstruction runs, five generations of the GA were performed. A plot of confidence metrics averaged over all 40 runs is shown at each generation (Fig. 2). The plotted metrics for each method are the average \(L\) and \(\chi^2\) values at each generation normalized by the sum over all generations, \(L_{\text{tot}}\) and \(\chi^2_{\text{tot}}\), respectively.

By the end of five generations, the CPR-GA performed better in both \(\chi^2_{\text{tot}}\) and \(L_{\text{tot}}\) showing an improvement of 3.87 \(\times\) \(10^{-2}\) in \(\chi^2_{\text{tot}}\) and 6.66 \(\times\) \(10^{-2}\) in \(L_{\text{tot}}\) or 39.2% and 96.7%, respectively. An \(L\) map of the best reconstruction from each method can be seen in Fig. 3. The CPR-GA reconstruction shows lower \(L\) at nearly every voxel.

B. Experimental data

The experimental data sets were also reconstructed 40 times using both algorithms. An example CPR-GA reconstruction can be seen in Fig. 4. This result is physically plausible, as it is expected that the Au nanocrystal would be
FIG. 2. Comparison of the average $\chi^2$ and $L_{\text{tot}}$ over 40 reconstructions of the simulated (top) and experimental (bottom) data sets using the independent and coupled routines. It is clear that the coupled routine performs favourably compared to the independent routine for both data sets. Spurious errors for the $L_{\text{tot}}$ values for the independent reconstruction reflect the fact that $L_{\text{tot}}$ is unconstrained in this method.

FIG. 3. $L(x)$ map for IPR-GA (top) and CPR-GA (bottom) for the simulated data set. The best reconstructions were chosen for each method. Cross sections in the $x$-$y$, $x$-$z$, and $y$-$z$ planes are shown. The CPR-GA reconstruction shows lower $L$ values at nearly every voxel, suggesting good agreement between its constituents. $L$ values have been capped at $10^{-6}$ to visualize the values in the coupled reconstruction.
faceted and have a flat surface where it touches the Si substrate. The strain fields were calculated by differentiating the displacement field with respect to the sample-frame directions $x$, $y$, and $z$, given as

$$
\epsilon = \begin{bmatrix}
\frac{du_x}{dx} & \frac{1}{2}\left(\frac{du_x}{dx} + \frac{du_y}{dy}\right) & \frac{1}{2}\left(\frac{du_x}{dx} + \frac{du_z}{dz}\right) \\
\frac{du_y}{dy} & 0 & \frac{1}{2}\left(\frac{du_y}{dy} + \frac{du_z}{dz}\right) \\
\frac{du_z}{dz} & 0 & 0
\end{bmatrix}.
$$ (15)

As seen in Fig. 5, the CPR-GA reconstruction shows lower $L$ values than the IPR-GA reconstruction throughout the crystal volume. A comparison of the error for each reconstruction method can be seen in Fig. 2. For the experimental data, the CPR-GA again performed better than the independent routine in both metrics, showing improvement of $4.1 \times 10^{-3}$ in $\chi^2_{\text{tot}}$ and $5.09 \times 10^{-2}$ in $\chi^2_{\text{RMS}}$, improvements of 5.07%, and 53.23% respectively. While the $\chi^2_{\text{tot}}$ value converges for the

FIG. 4. CPR-GA reconstruction of experimental Au nanocrystal on Si substrate. The strong facets and existence of a flat surface on the $x$-$z$ plane agree with expectations.

FIG. 5. $L(x)$ map for IPR-GA (top) and CPR-GA (bottom) for the experimental data set. The best reconstructions were chosen for each method. Cross sections in the $x$-$y$, $x$-$z$, and $y$-$z$ planes are shown. The CPR-GA reconstruction shows lower $L$ values than the IPR-GA, suggesting good agreement between its constituents. $L$ values have been capped at $4 \times 10^{-5}$ to compare smaller values in the interior of the crystal.
IPR-GA, the $\frac{L_{\text{det}}}{L_{\text{tot}}}$ value actually increases gradually, suggesting that the constituents drift further apart with each generation. Conversely, the $\frac{L_{\text{det}}}{L_{\text{tot}}}$ for the CPR-GA stays constant as $\frac{L_{\text{det}}}{L_{\text{tot}}}$ converges. This demonstrates that, by combining reconstructions every 20 iterations, the CPR-GA effectively reduces $\frac{L_{\text{det}}}{L_{\text{tot}}}$ by sharing information from all Bragg peak measurements during phase retrieval, while the IPR-GA does not. Though less pronounced than for the simulated data, the improvements in both $\frac{L_{\text{det}}}{L_{\text{tot}}}$ and $\frac{L_{\text{det}}}{L_{\text{tot}}}$ for the experimental data demonstrate that the CPR-GA provides a more self-consistent estimate of $u(x)$ than the IPR-GA.

C. Discussion

Several CPR routines already exist in the literature, and recent methods developed by Newton et al. [23] and Gao et al. [24] have sought to design phase retrieval approaches that utilize multiple Bragg peak data sets in different ways than the routine presented here. Newton et al. used a sequence of HIO steps interleaved with least-squares corrections to the displacement field $u(x)$ similar to 10, and Gao et al. intermittingly update $u(x)$ using an iterative procedure inspired by the bisection method. What has not yet been addressed by these works is the applicability of CPR to experimental data. This is important to consider given that experimental BCDI data of multiple Bragg peaks inevitably contains artifacts from factors such as particle drift, imperfect beam/sample alignment, background scattering, and so on, which are not present in simulated data sets and which may adversely effect the performance of CPR strategies.

To handle such imperfections in the data, GA approaches for independent phase-retrieval (like the IPR-GA in this work) are typically necessary to produce the best reconstructions [21], and as such, should be the benchmark by which any CPR approach is judged. In performing such a comparison for simulated and experimental data, we demonstrated that using a GA with a coupled routine is able to outperform an independent GA, contributing to the advancement of this family of techniques. Given the nascent state of research in this field, it can be imagined that more sophisticated CPR strategies will be developed, and the use of genetic algorithms as implemented here promises to improve the robustness of such strategies in the context of experimental reconstructions.

VI. CONCLUSION

Coupled multipeak phase retrieval promises to aid considerably in the recovery of 3D strain fields for BCDI. By adapting a genetic algorithm for reconstructing a single Bragg peak to our coupled algorithm, we were able to achieve displacement field reconstructions that were more self-consistent in terms of both simulated and experimental data sets. Though not alone in the coupled reconstruction domain, by demonstrating the approach on experimental data, we establish the viability of such approaches for future of coherent x-ray experiments.

ACKNOWLEDGMENTS

The authors would like to thank Reejoo Pokharel for enlightening discussions. Joseph Pauza is thanked for providing a grain with a realistic shape based on a SPARKS simulation [32]. R.L.S. was supported by funding from the Department of Physics and Astronomy and the College of Physical and Mathematical Sciences at Brigham Young University. Work at CMU was supported by the US Department of Energy, Office of Science, Basic Energy Sciences, under Award No. DE-SC0019096. The diffraction data were obtained at the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. The work of S.M. and S.O.H. (contributing to the design and demonstration of the multi-Bragg-peak phase retrieval algorithm) was supported by the US Department of Energy, Office of Science, Basic Energy Sciences, Materials Science and Engineering Division.


