NS190 – Research Paper

Nanoscale imaging: TXM and diffractive methods

A discussion on the use of coherent diffraction of X-rays in the observation of bulk heterojunctions

Jordi Antoja Lleonart (S2820048)
Top Master Programme in Nanoscience
Cohort 2014-2016
Supervisor: Ra'anan Tobey

Abstract

In this paper, the issue of imaging the active layer of organic solar cells is presented. Afterwards, a brief introduction is made on electron microscopy, which is a current, routine method for direct nanoscale imaging. X-ray microscopy is then used as a vehicle to quickly introduce some important concepts in the field, such as coherence, generation, and focussing of X-rays. Then, after a reminder of basic X-ray diffraction, potential alternatives to electron microscopy are presented, starting with holography-based methods, along with some of their representative experimental results. Once their advantages and limitations have been exposed, the more involved iterative phase retrieval methods are introduced, and their strong points are highlighted. After reviewing some more practical cases, the specific case of ptychography is described in more detail. The general algorithm is explained, and a few practical uses are listed. In the conclusions the main features of each diffractive method are summarised, and the procedure with the most potential is chosen.

List of Contents

Common abbreviations

AC: Autocorrelation
BHJ: Bulk heterojunction
CDI: Coherent diffractive imaging
EDX: Energy dispersive X-ray (analysis)
EELS: Electron energy loss spectroscopy
FEL: Free electron laser
FT: Fourier transform
FTH: Fourier transform holography
FZP: Fresnel zone plate
OPV: Organic photovoltaics
SEM: Scanning electron microscope
SNR: Signal-to-noise ratio
(S)TEM: (Scanning) Transmission electron microscope
(S)TXM: (Scanning) Transmission X-ray microscope
URA: Uniformly redundant array
XR: X-rays

Introduction

X-Ray (XR) diffraction nowadays is extensively known as a tool for the imaging of crystalline samples, and so the imaging of non-periodic objects is usually assumed to be out of its reach. In this paper, an attempt will be made to disprove this assumption.

The use of XR diffraction on periodic systems has been extensively studied, as have been related techniques, such as XR-reflectivity. On the other hand, XR coherent diffractive imaging for non-periodic samples is less well-known. This is not to say that it is not already being used. The most famous current example is the work by Chapman on proteins which are difficult to crystallize. This paper will go through this case and many others.

What may be even more shocking is that diffractive imaging has been around for quite some time now. It was actually several decades ago that Fienup demonstrated a way to get around the crystallographic phase problem[i] for complex-valued objects, and there has been steady progress in the field from then on.

Nowadays, there are several techniques that make use of coherent XR to obtain real-space images of a wide variety of objects. Some of the most representative will be discussed here, the first one being XR holography and, more conveniently, Fourier Transform (FT) XR holography (visible light FT holography was shown to work back in the sixties[ii]).

The second one, which carries a heavier computational load, is generally referred to as iterative phase retrieval; a myriad of different algorithms and experimental setups exist in this context, and some of them will be reviewed. Then, the discussion will move on to the imaging of extended samples, which is not as trivial as one would initially think.

It is important to keep in mind that there already are well-known, routinely used techniques in the field of nanoscale imaging which do *not* rely on XR to work. These include a wide variety of scanning probe microscopies (which will not be discussed here) and electron microscopy.

Thus, it is also important to judge if the above mentioned XR methods are actually better than the existing alternatives. To do that, the discussion will be driven by the need to obtain good-resolution images of a specific sample type: the bulk heterojunction (BHJ) in organic photovoltaic (OPV) devices.

The area of OPV works towards new cell designs that use cheap, light and easily processible materials, partly striving to reduce their initial costs. The base design for some of the devices with the largest efficiencies[iii] consists, in short, of a 100 to 200 nm-thick heterojunction between two electrodes, one of which is usually transparent. The BHJ is composed of a mixture of donor and acceptor materials, which are usually a conjugated polymer and a fullerene derivative, respectively.

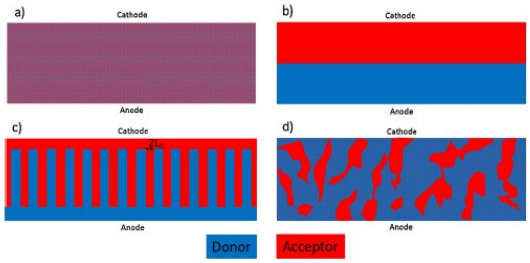

Figure 1: Side view of some BHJ types in Ref. 4. In a) the donor and acceptor are intimately blended, whereas in b) they are clearly separated. c) illustrates a design with large interfacial area (note that the domain size is close to the exciton diffusion length). d) shows what a real BHJ may look like.

The nature of the heterojunction is responsible for a great part of the overall device performance[iv]. It has been determined that, because of the excitons' limited lifetime, only the regions within a limited distance of the donor-acceptor interface will contribute to the current generation. This distance is assumed to be of the order of the exciton diffusion length, or a few nanometers. A design has to be chosen that maximizes the area of this interface, and this goes through obtaining domain sizes below the 10nm mark.

Imaging at resolutions of a few nm is not unusual for electron microscopy. This approach, though, faces several limitations when attempting to image organic, thick and/or hydrated specimens. In this context, it will be seen that XR can work with such samples. This paper starts by discussing the XR analogs of TEM and STEM (regular and Scanning Transmission Electron Microscopes, respectively), which nevertheless show very good resolution on

test objects[v], and then moves on to the less straightforward areas of holography, coherent diffractive imaging (CDI) and ptychography.

Prelude: Electron Microscopy:

One of the motivations for using electron over visible light microscopes is the extremely lower diffraction limit that accelerated elecrons have. We may first have a look at the equation for the Rayleigh criterion, which sets this very fundamental limit to the resolution of conventional radiation-based microscopies.

$$r = \frac{0.61 \cdot \lambda}{\sin \alpha}$$

Formula 1: One of the forms of the Rayleigh criterion, where r is the minimum separation necessary to resolve two separate objects. α is the half-opening angle of the optics, and λ is the wavelength being used. Note that r=0.61λ is the lower resolution bound reached by maximizing the aperture angle.

Thus, assuming that everything else works perfectly, the diffraction limit is reached, and the smallest detail that can be resolved according to the Rayleigh criterion is usually of the order of the wavelength being used. Thus, according to the wave-particle duality of electrons, it can be worked out that an electron accelerated at 100 kV has a *de Broglie* wavelength of about 3.70 pm, which sets diffraction-limited resolution potentially below typical interatomic distances in a solid. For comparison, visible light has wavelengths in the 390 to 700 nm range.

In practice, some impressive sub-nm resolutions have been achieved in TEM and STEM. In STEM, this may be used in conjunction with Electron Energy Loss Spectroscopy (EELS), Energy Dispersive X-Ray (EDX) analysis or time-resolved cathodoluminiscence to obtain detailed information on the local properties at different positions in the (preferably inorganic) sample[vi].

Furthermore, regular TEM has shown tomography capabilities that allow 3D imaging of reasonably thin samples. SEM (Scanning Electron Microscope), which is usually limited to imaging surface topography, works in an analogous way when combined with sample cutting by FIB (Focused Ion Beam)[vii].

In spite of all these strengths, electron microscopy has some issues in the imaging of these organic systems; due to the strong interaction of the electrons with matter, the images are limited to very thin, dry specimens (or to their surface) and these techniques usually involve heavy damage to the sample by the electron beam. As a result, it can be argued that other imaging methods would be welcome in the field.

I: Transmission X-Ray Microscopies:

· Overview:

An interesting alternative to electron microscopy, and one that circumvents many of its limitations, is XR transmission microscopy. A transmission XR microscope (TXM) functions in a way that is very much analogous to a regular TEM, in that it consists of a (XR) source, a set of optics to focus the beam onto the sample and a module that records the transmitted intensity downstream from the sample holder. These microscopes will serve as a vehicle to introduce some important concepts that will be relevant later on.

A valid question to ask is whether or not is there any actual upside to using TXM rather than TEM or visible light transmission. We should keep in mind that soft XR radiation usually has wavelengths in the range of 2-4 nm, and hard XR have wavelengths of the order of a single Ångstrom. This means that, although the potential resolution for TXM and related techniques is much better than that of visible

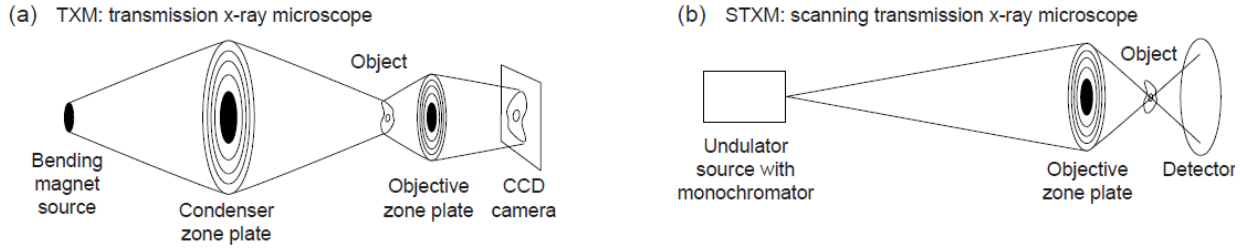

Figure 2: From Ref. 8, usual setup for the two main types of XR transmission microscopes. These geometries are very similar to the ones used in diffractive methods later on.

light, it is not as extreme as that of TEM.

In spite of this fact, there are notable advantages to using XR rather than electrons as well. The main strong point of TXM is the higher penetration depth of XR. TEM usually faces heavy challenges with samples over 100 nm in thickness, whereas TXM can handle imaging of several-microns-thick specimens. Other advantages include lower sample damage (except when very bright XR sources are used) and the ability to tune the XR energy. This way, it is possible to have it be somewhere between the K absorption edges of carbon (290 eV) and oxygen (540 eV), for instance. This energy range is referred to as the *water window* because the imaging so performed has strong (absorption and phase) contrast for organic materials, but very low contrast for water, allowing the use of hydrated organic and biological samples.

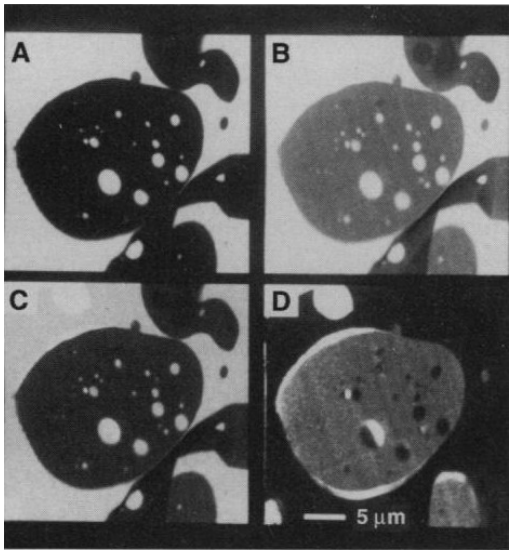

Figure 3: From Ref. 9, absorption-contrasted STXM images of a sample containing polypropylene and a styrene-acrylonitrile (50% weight of each) copolymer. From A to D, the radiation energies used are 285.5, 286.2, 286.8 and 287.9 eV. PP and the copolymer, respectively clear and dark in micrograph A, are clearly phase separated.

There are more examples of the versatility of XR transmission[viii]. Assuming that the energy of the incoming photons can be selected with a resolution of about 0.2 eV (in EELS, the resolution can be below 0.1 eV, according to Ref. 6), it is possible to scan the beam energy while registering the absorption. This results in so-called *spectromicroscopy*, where different elements or compounds have different signatures (absorption spectra), and therefore, to some extent, composition maps of the sample can be obtained[ix]. Tomography (for 3D imaging) can be obtained by registering several images of the sample at different angles, as long as the extended exposure can be withstood.

· **Coherence of X-rays[x]:**

Before proceeding any further, some definitions need to be established about the general properties of an XR beam, namely transverse and longitudinal (also refered to, respectively, as *spatial* and *temporal*) coherence. These can be estimated by working out the relevant coherence length values, as follows. A large enough value of these is crucial for diffraction experiments.

$$\Lambda_t \leqslant \frac{\lambda}{\epsilon} \qquad\qquad \Lambda_l = \frac{\lambda^2}{2\,\delta\lambda}$$

Formula 2: Two out of the many forms of the coherence lengths. λ is the radiation wavelength (mean) and δλ is the wavelength standard deviation. ε is the beam FWHM divergence, i.e. the angle that relates the beam diameter to the distance from its source.

Broadly speaking, these lengths define the volume of a beam that oscillates at a reasonably constant phase. We can then see that this volume is larger, and so coherence is better when the beam is properly collimated (i.e. has low divergence) and monochromatic (so it has low wavelength spread). For instance, transverse coherence can be improved by a proper choice of aperture, and longitudinal coherence, with monochromators.

Coherence will become especially relevant in the case of diffractive imaging, since not meeting certain coherence values can result in loss of information. Although there is some freedom to it, it is generally agreed that the transverse coherence should be over twice the sample width[xi]. This has to do with the concept of oversampling, which will be dealt with later, but for now it suffices to explain that the wave phase has to be maintained not just along the sample, but that of its autocorrelation function, which is twice as large. Analogously, the beam divergence should be below half of the diffraction angle; the reasoning behind it is very similar. Longitudinal coherence requirements are more intuitive; it needs to be larger than the maximum optical path difference between any two points in the object.

· Machinery - Production, control and detection of X-rays:

So to speak, the "handling" of XR is quite different from that of regular, visible light in classical microscopy. Beginning with the source of the radiation, there is a wide range of devices that may be used to obtain a primary beam, depending on the requirements and the budget of the experiment. The simplest ones are vacuum tubes, which consist of two metal electrodes which are separated by a vacuum. Applying a high voltage between them releases electrons from the cathode, which are accelerated towards the anode; when they strike it, they generate XR radiation through a variety of electron relaxation processes in the metal. Their efficiency is rather small, and the emission spectra that can be obtained are far from monochromatic, and so not very coherent.

Monochromaticity is a requirement for many applications of XR reported here. Therefore, quite a bit of effort is usually put into modifying the spectrum of whatever source is being used so that it becomes as narrow as possible. To do that, it is not unusual to place a filter (for instance, nickel or zirconium) in the way of the beam to take out the undesired frequencies. Another routinely used monochromator consists of a single crystal placed at a specific Bragg angle relative to the beam. This results in the several wavelengths that compose the beam to scatter in slightly different angles; using a narrow slit, a small range of them can be selected. However, using this kind of monochromators involves that the primary beam, i.e. the most intense one, is lost, and only the diffracted, lower intensity beam can be used.

$$2 \cdot d_{hkl} \cdot \sin(\Theta_{hkl}) = n \cdot \lambda$$

Formula 3: Bragg's law. XR of different wavelengths will have different reflection angles (Θ) for the same diffraction plane (hkl).

At this stage, it is necessary to introduce different routes to the generation of high-quality XR beams. It should be noted that most elements that modify the beam in any way (filters, mirrors, zone plates, etc.) usually have intrinsically low efficiencies. This is usually enough for XR crystal diffraction, as in these techniques the intensity is concentrated at the Bragg peaks and the Signal to Noise ratio (SNR) is large enough. In the case of CDI, it will be seen that this does not happen and a bright, coherent source is needed.

High-harmonic generation of XR deserves a special mention here, as it can yield very coherent radiation in tabletop (i.e. low-scale) setups. However, its lower brightness means that prohibitively long exposure times are sometimes required[xii]. A more practical, if expensive, solution, is currently found in synchrotron facilities and XR Free Electron Lasers (XFELs).

In summary, a synchrotron consists of a system in vacuum where electrons go around a storage ring at high speed. Initially are emitted from a pulsed source and initially go through a linear accelerator. Then they are transferred to a booster ring, where they are accelerated before entering the storage ring; they circulate it in small packets, or *bunches*, and usually stay there for hours until they are lost (for instance, upon impact with the ring's walls). The electrons' trajectory is controlled via external magnetic structures. These can be bending magnets, which help the beam change direction, or wigglers and undulators, which make it oscillate perpendicularly to its propagation direction. In any case, the electrons going through these structures emit electromagnetic radiation. Ultimately, with proper control, coherent radiation of a wide frequency range, as well as different polarizations, can be achieved.

A more recent type of structure is the Free Electron Laser (FEL). This resembles an elongated linear accelerator with a long undulator section at the end. Without going too much into detail, this increased length results in the emitted radiation on the back side of a bunch stimulating further emission at the same frequency down the line. This is known as Self-Amplified Spontaneous Emission (SASE), and it results in higher coherence and peak brightness that those from a regular synchrotron.

However one chooses to do it, once a sufficiently monochromatic beam has been obtained, focusing it is the next challenge. For visible optics, regular lenses work properly. However, because XR interact weakly with them, a different paradigm is required. In this context, the most extensively used alternative are Fresnel Zone Plates (FZP). They consist of a series of concentric rings (called *zones*) of appropriate thickness and spacing, which decrease away from the center, sometimes including a central beam stop to take away the unwanted, nonfocused primary beam. When about a hundred zones are present, a FZP behaves like a regular lens over a

wide energy range.

$$r=\frac{k_1\cdot\lambda}{NA_{ZP}} \quad \text{or, if spherical aberration is low} \quad r=2\cdot k_1\Delta_r$$

Formula 4: From Ref. 5, formula for the smallest resolution that can be achieved at the first order focal spot with a FZP. k_1 is an experiment-dependent constant, NA_{ZP} is the numerical aperture of the FZP and Δr is the width of the outermost zone.

When such a construction is illuminated with, for instance, XR plane waves perpendicular to the zone plane, the waves are diffracted. This happens in such a way that constructive interference occurs at specific points of the optical axis. These points are referred to as n-th order focal spots; for a theoretical FZP with a realistic geometry it was calculated[xiii] that between 10% and 20% of the incident light would actually be focused in the first order spot, with increasingly smaller amounts being focused in the higher order spots, and the rest of it being wasted. This diffraction efficiency actually depends mostly on the material the zones are made of. Moreover, it is higher if the zones are not opaque, but introduce a phase shift in the radiation instead. In any case, with a source of enough brilliance efficiency is less of an issue. This brings up the real challenge for FZP use in high-resolution techniques, such as STXM.

That is, the need to fabricate zones as thick and narrow as possible. The thickness ensures better efficiency and, as shown in Formula 4, this is needed to improve the FZP resolution. Indeed, clever nanopatterning techniques(REF Weilun Chao)[v,xiv] have succeeded in generating zones with these geometries, but the diffraction limit still seems to be out of reach. It will be seen later that zone width does not set a resolution limit for CDI and ptychography.

The last element worth mentioning at this stage is the detector itself. Just as optical microscopes usually resort to cameras to record images, in XR microscopy and related techniques a widely used detector type is the CCD (Charge-Coupled Device). This kind of detector roughly consists of a 2D array of individual MOS capacitors, such that when a photon strikes any of them, charge photogeneration occurs with some probability. These charges can then be transferred outwards, and read as an output voltage, which can in turn be eventually converted into a histogram of photon counts per pixel. This is usually interpreted as the image itself. Note that, as described, this image will not contain straightforward information about the photons' wavelengths. When it is not necessary to detect intensities over a whole area simultaneously, a simpler point detector may be used.

II: XR Diffraction in general:

A comprehensive introduction to XR physics can be found in *Elements of Modern X-ray Physics*[xv]. Here we will summarize just a few concepts relevant to diffractive imaging. Starting with some formulae, for a nonperiodic distribution of atoms we can define a structure factor such as:

$$F(\vec{Q})=\sum_j f_j(\vec{Q})\cdot e^{i\vec{Q}\vec{r}_j} \quad \text{where} \quad \vec{Q}=\frac{4\pi\sin(\Theta)}{\lambda}$$

Formula 5: Structure factor for a finite atom distribution. In the formula, $\boldsymbol{Q}$ is the wavevector transfer in reciprocal space; Θ is the diffraction angle and λ is the radiation wavelength. For each atom in the system, $\boldsymbol{r}_j$ is the position in real space, and f_j is the form factor. At $\boldsymbol{Q}$=0, f_j equals the number of electrons in atom j.

In the far field region, the intensity that will be scattered by this atomic distribution at an angle Θ_0 (and therefore wavevector transfer $\mathbf{Q_0}$) is proportional to $F^*(\mathbf{Q_0})\cdot F(\mathbf{Q_0})$. Now, due to the reasonably weak interaction of XR with matter, the diffracted beams are usually considered as small perturbations of the primary, intense beam. This is known as the kinematical approximation, and it is justified by the experimental observation that, indeed, the diffracted intensities are much lower than the nondiffracted one. These low intensities pose a problem for recording diffraction patterns in nonperiodic systems. Classically, this problem has been solved by building periodic arrays (crystals) of the systems at hand. If this can be achieved, we can write:

$$F(\vec{Q})=\sum_j f_j(\vec{Q})\cdot e^{i\vec{Q}\vec{r}_j}\cdot\sum_k e^{i\vec{Q}\vec{R}_k}$$

Formula 6: Equation for a periodic distribution of atoms. $\boldsymbol{R_k}$ is the position of each unit cell of the crystal.

Where the new term is called *lattice sum.* Now, the first sum runs over all the atoms in a unit cell, and the second sum runs over all unit cells in the crystal. Thus, we see that, as a consequence of the periodicity in the system, the scattered waves from a

huge number of unit cells may constructively interfere at some special **Q** vectors. These correspond to Bragg angles (see Formula 3) and result in intensity peaks in a diffraction pattern. Close to (and at) these peaks, the SNR is large enough to be measured easily in a regular lab.

At any rate, if the F(**Q**) values could be obtained for several diffraction angles, it would in principle be possible to build a nonlinear equation system using Formula 5; for an N-atom system, there would be 3N unknowns (that is, their coordinates in direct space), so merely obtaining F for 3N different **Q** should lead to a unique solution. However, the so-called *phase problem* needs to be solved. This problem stems from the fact that F is a complex number, whereas the measured diffracted intensities are real numbers proportional, as explained above, to the the squared modulus of F. Therefore, its phase is lost in the measurement; several strategies have been developed in order to get around this problem.

In crystalline systems, the solution to the phase problem usually involves the well-established use of direct methods for single crystals; other methods are more suited when working with powder samples[xvi]. However, there are objects of interest, such as some types of nuclear proteins, which cannot be crystallized. For these systems, a relatively recent approach is the use of high brilliance sources such as the ones described previously; in that approach, the primary beam is very intense, and so are the diffracted beams, thus resulting in measurable diffracted intensities, even without resorting to lattice sums. The way these intensities are used will be detailed later on.

As a small intermezzo, it is necessary to add that a major consequence of using such an intense beam (i.e. large electric field) is that the sample usually suffers irreversible damage within very short exposure times. The way this is sometimes handled in CDI is using very short pulses rather than a continuous beam; if the pulse is shorter than the photoionization timescales (tens of ps) this can lead to the measurement of a diffraction pattern before the sample is damaged.

Especially when using biological objects, freezing the sample seems to be the method of choice in several XR microscopies. However, using femtosecond pulses[xvii] has shown some of the best results in the area of CDI. Thus, the discussion on the sample damage will be kept to a minimum by assuming that either the beam intensity is not too large, or the pulses available are conveniently intense and short; that is, many of the cases described from now on rely on either synchrotron or FEL radiation to work.

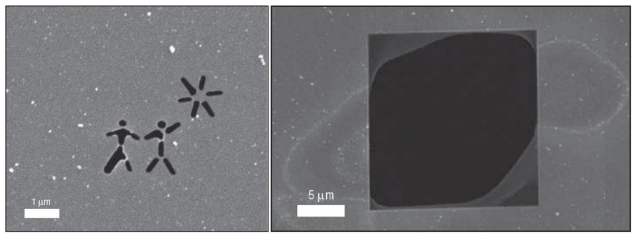

Figure 4: (Composed from figures in Ref. 17) SEM image of a structure carved into silicon nitride, before (left) and after (right) a 25fs FEL pulse. Around $4 \cdot 10^{13}$ W/cm^2 were deposited in the sample during that time. The dark square in the right image shows the position where the structure was before the pulse, illustrating the damage sustained.

III: Coherent diffractive methods:

So far, we have discussed diffraction as it is used in the field of structure determination. The absence of Bragg peaks in nonperiodic systems requires new approaches to the phase problem. However, the experimental setups are usually indistinguishable from the regular TXM geometry or, in later cases, STXM.

When dealing with a (nonperiodic) structure as large as a normal BHJ, it is not reasonable to model it as a collection of atoms. Instead, the usual procedure is to picture the object under observation as a 2D array of pixels. A density value (f) is assigned to each of them. It will be seen that, in some simple cases, this value can be restricted to always be a real number, but in general it is complex.

· **Holography:**

The first, deceptively simple approach to the phase problem consists in not formally solving the phase problem at all, and instead using a clever trick to achieve a reasonable image of the sample. Here it may be useful to introduce a small diagram that is usually well-known to crystallographers. Keep in mind that we work with object density (i.e. pixel value) but the scheme works similarly with electron density or, to be precise, the form factor.

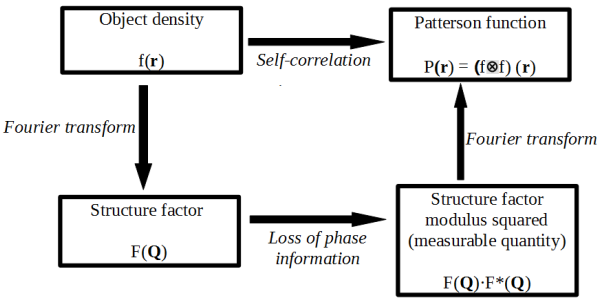

Figure 5: A quick guide to the phase problem. Note that Fourier transforms and reverse Fourier transforms are used interchangeably.

It now becomes clear that the measured data can be Fourier-transformed (FT) in order to obtain the Patterson function (more accurately, its modulus, but for now we will stick to real-valued functions for simplicity). This function can and has been used in the past to solve the phase problem for crystals; peaks in P(**r**) correspond to interatomic spacings and have height proportional to the product of the form factor of the two atoms involved. Thus, if a crystal unit cell contains a heavy atom, P(**r**) will show sharp peaks and the Patterson method can yield good results. In our case, however, it has a different use. As it is the autocorrelation (AC, i.e. convolution with itself) of the object density, this function still carries a good deal of information about the object.

This information can be accessed through a proper experimental setup and very little computational cost. In fact, the first demonstrations of the technique were carried out in the late 1940s purely through optical means by Dennis Gabor, who would go on to win the 1971 Nobel Prize in Physics for his work in holography.

$$AC(\vec{r}) = \int_{-\infty}^{+\infty} f(\vec{x}) \cdot f(\vec{x}+\vec{r}) \cdot d\vec{x}$$

Formula 7: AC general expression for a real-valued density function. If the density function is complex, the AC function may be complex as well. Note that, by construction, it will have its highest peak at, and will be centrosymmetric around zero shift (r=0).

In general, if one performs the AC of a regular object, the result cannot be easily traced back to the original object. The main idea in holography is using a reference wave to *phase* the diffracted waves from the object. In practice, this is commonly done by adding a hole in the same plane the object is in. This is equivalent to placing a second object, ideally not too close to the first one, and with a narrow and large density distribution. This can be illustrated with an example in one dimension, as seen in Figure 6. It is not difficult to imagine that this also works with 2D arrays of pixels.

As seen in the figure, the AC obtained from the regular object does not reveal much about the original object. However, upon addition of the reference, new peaks appear. These contributions stem from the correlation between the object itself and the reference, which in this case resembles a Delta function. Close analysis reveals that these peaks are, in the ideal case, symmetry analogs of the original object. It is intuitive that, as these features arise from cross-correlation, they will resemble the object down to a resolution that is limited by the reference aperture size. This limit is not fundamental, and it has been seen that adding iterative phase retrieval (which will be introduced later) to the procedure allows for resolutions even

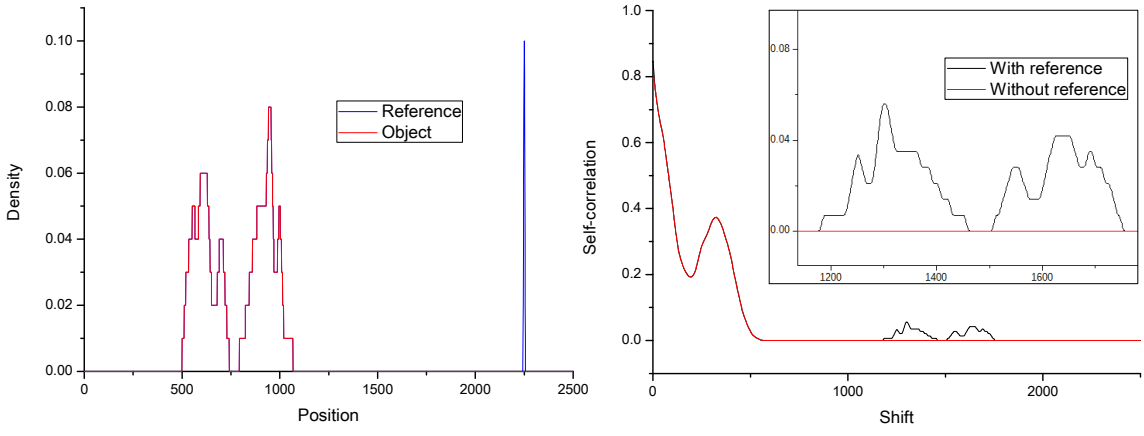

Figure 6: Self-correlation in practice. To the left, original object and reference; to the right, AC function. Inset: close-up on the peaks that result from adding the reference.

closer to the diffractive limit.

Thus, the real issue becomes obtaining the AC function through diffractive methods. If one shines plane waves of some radiation into the object-reference construct, a lens (or a FZP) needs to be placed downstream in order to produce the hologram out of the diffracted waves. Alternatively, if one can achieve diffracted spherical waves of the same sphericity (as shown in Ref. 2) that eliminates the need for a lens. This is the case that shows the best analogy to the previous XRD scheme.

Then again, the hologram (which, in the end, is just a diffraction pattern) has to be Fourier-transformed in order to yield the AC function. In the past, the plate with the hologram had to be placed in a separate beam and FT was performed optically with a lens, as in Figure 7; the work by Baez[xviii] gives further information on that method. Nowadays the hologram is usually recorded in a CCD and the FT is carried out computationally, which is relatively faster and less material-consuming. This is called FTH (Fourier Transform Holography).

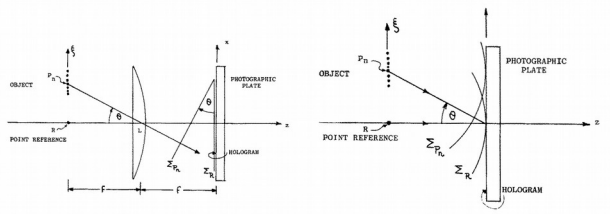

Figure 7: From Ref. 2, setups for holography experiments. To the left, holography with a lens; the object and the photographic plate are at the lens' focal planes. To the right, lensless holography.

· FTH in practice:

Now that the principles of holography have been set, a number of more recent experimental results can be presented. Domain imaging will be shortly discussed (albeit for magnetic samples). However, the first few cases show some practical problems with FTH, and several attempts to solve them.

For instance, the work of Schlotter *et al.*[xix] addresses the reference size dicotomy; on one hand, the resolution accessible through FTH is, as mentioned, strongly limited by the size of the reference aperture. On the other, if the aperture is too small, it will not transmit enough intensity and the SNR on the AC function will be too low for imaging. The solution found by these researchers is to add several, properly spaced, identical reference apertures around the object. This generates several repetitions of the object in the AC image, which can in turn be superimposed to generate a higher SNR image (by {Number of apertures}$^{1/2}$ following Poisson statistics). This approach constitutes an elegant alternative to using higher primary intensities, and so reduces sample damage.

A similar experiment by Stadler *et al*. involved a gold structure in the shape of the letter P instead[xx]. It was noted that, due to a beam stop being used to block the primary XR beam, intensities diffracted at low spatial frequencies were lost, and the eventual AC function appeared distorted. On a separate note, it was demonstrated that the phase differences on the AC function could be used to provide a good estimate for the object height (i.e. thickness), which showed agreement with AFM measurements.

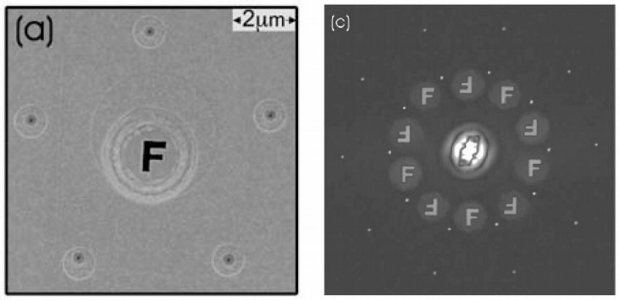

Figure 8: (Composed from figures in Ref. 19) Left: SEM image of the sample. A letter F, as well as five reference holes, have been carved with a Focused Ion Beam into a gold film. Right: FTH reconstruction; note the central symmetry.

Using spread out groups of pinholes still remains, however, an inefficient method. This is because, as made obvious by Figure 8, they have to be sufficiently spaced in order to avoid overlap in the AC image. A different approach involves fabricating a FZP in place of a simple hole[xxi]. This allows fine selection of the flux through the reference by careful selection of zone sizes; likewise, the FZP's focal length and spot size (which is effectively the reference size) can be easily chosen. Admittedly, the fact that the reference beam is now focused needs to be acounted for, and this requires a conceptually more complicated procedure overall.

In a different line Marchesini *et al.* proposed using coded apertures[xxii], which show promise for large resolution *and* SNR. A coded aperture may be thought of as a complicated distribution of pinholes, but a very deliberate one. They initially found use in

XR telescopes, where the spatial distribution of multiple XR sources could be determined from the "shadow" cast by such an aperture. Further development resulted in Uniformly Redundant Arrays (URAs).

Other than that, they are used in the same way as a single point reference; of course, the correlation between an object and the URA cannot be interpreted "by eye", but knowing the exact shape of the URA allows for image reconstruction, down to a resolution which is usually limited by a characteristic size in the array. That is, in turn, determined by the smallest feature that can be achieved by nanolithography, and in particular Ref. 22 mentions URAs with around 25nm feature sizes.

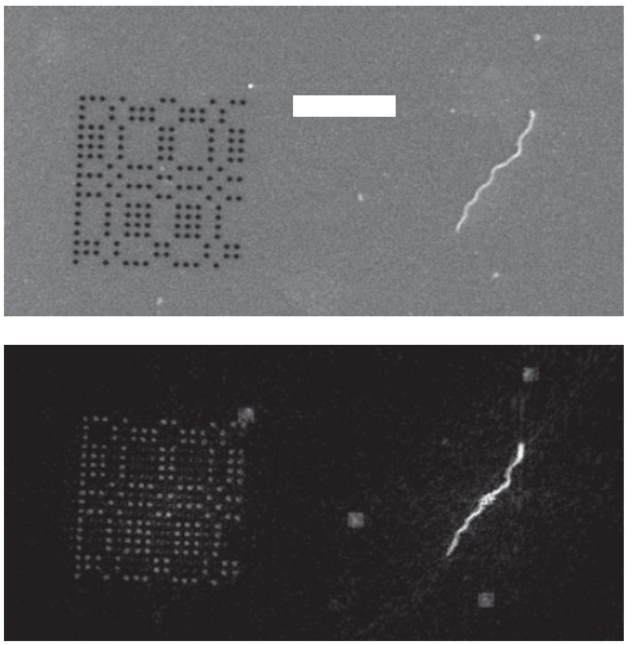

Figure 9: (From Ref. 22) Top: a SEM image of the URA beside a Spiroplasma *cell. For scale, the diameter of the holes is around 150nm, and the white bar is 4μm in length. Bottom: reconstructed image after FTH and iterative phase retrieval. It is worth noting that the hologram was measured after a single 15fs pulse from the FLASH FEL in Hamburg, and that this pulse destroyed both the array and the sample.*

A more recent way to improve SNR is to use extended references instead of pinholes or pinhole arrays. In favourable cases, if the shape of a large reference is known, or closely estimated, one may computationally recover the object[xxiii]. This method is, in fact, quite resilient, and tolerates some inaccuracy in the previous knowledge that one has about the reference aperture. The reconstruction procedure itself is a bit more involved than the one used for point references, and will not be discussed here, but it appears to be flexible enough to work with a wide range of apertures, including URAs.

Moving on from test objects, there is a particular, conceptually simpler case of FTH that deserves special attention. In the paper by Eisebitt *et al.*[xxiv], circularly-polarized XR from a synchrotron source are used to image a magnetic multilayer Co/Pt circular sample. The wavelength is set to the L_3 absorption edge of Co, and a single reference hole is carved via focused ion beam.

The results are nothing short of spectacular. Not only can the out-of-plane magnetic domains be mapped down to a resolution of about 50nm, but the image obtained is virtually identical to that recorded from STXM. Inverting the beam helicity also inverts the image contrast, confirming that the magnetic domains are, in fact, what makes up the image.

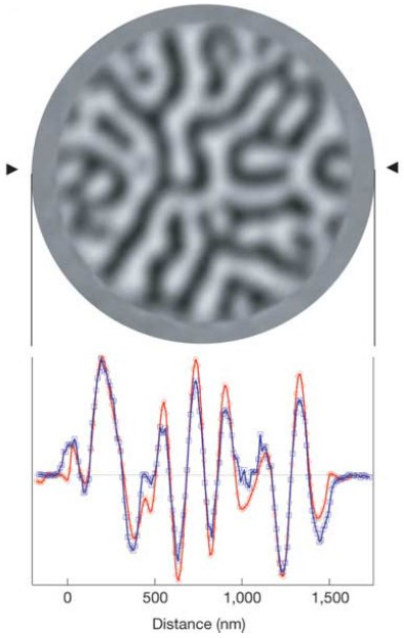

Figure 10: From Ref. 24, image of the two types of magnetic domains within the sample. The plot below illustrates image contrast along the diameter of the STXM (blue) and FTH (red) images, highlighting their similarity.

It goes without saying that the method this paper is, debatably, still far from the idea of imaging organic domains within a BHJ active layer. The reason it is being brought up here is that it conveys the message that, if a proper contrast mechanism can be realized for such a sample, FTH constitutes a reliable

approach to domain imaging.

As a closing note for FTH, it becomes clear that it is not a method that can, at present time, solve the BHJ imaging problem. It may initially be argued that good quality references, such as FZPs or URAs, take a good amount of work to build alongside the mask. This is especially troublesome if a FEL source is to be used, as the whole construct will be destroyed in the experiment. Thus, using these complicated apertures is not the cheapest option for routine imaging.

It follows that its limitation, and a recurring theme at this point, is that the more simple reference apertures are not yet reaching sizes below those of a domain in an ideal BHJ. Thus, even though some authors still praise holography for its simplicity and good results, it is clear that the efforts described so far are still not useful in the case at hand. Eventually, it becomes evident that holography *alone* is probably not enough for the routine imaging of samples such as a BHJ. This is why, in order to overcome these limitations, one may resort to iterative phase retrieval.

· **Iterative phase retrieval:**

Holography is not the only way one can determine the shape of small objects through XR diffraction patterns. Iterative phase retrieval, often referred to simply as CDI, does not rely on the AC function to obtain details of the sample, but it uses a much more fundamental mechanism instead. This is somewhat advantageous because, as mentioned, the resolution of these methods will not be limited by nanofabrication capabilities anymore. Instead, the resolution will be determined by the experimental setup (camera size and position, etc.) down to the diffraction limit. Many of the main concepts relevant in CDI are detailed in Reference 1, and will now be summarized.

A good way to see how CDI works is to go over the structure factor formula (Formula 5) again. It is easy to see that, for any $\mathbf{Q}$ value, $F(\mathbf{Q})$ will be a sum of j terms, where j is the number of pixels in the object. For several wavevector transfers, $\mathbf{Q_1}$ to $\mathbf{Q_m}$, there will, of course, also exist several structure factors, $F(\mathbf{Q_1})$ to $F(\mathbf{Q_m})$. This is not new, but it is now possible to envision this set of $F(\mathbf{Q})$ values as a system of equations; the unknowns in this system are the $f(\mathbf{x})$ values, of which there are precisely as many as there are pixels in the object.

In the paper by Miao *et al*. (Ref. 25), the object was approximated, depending on dimensionality, as a line, square or cube with a side of N pixels. Thus, the number of pixels within the object will be N, N^2 or N^3, respectively, and the number of unknowns in the system will be 2N, $2N^2$ or $2N^3$ as the density in each pixel has a real and imaginary part (or modulus and phase). Furthermore, they remarked that if the object density is real-valued, the diffraction pattern will be centrosymmetric; this is known as Friedel's law, and one of its consequences is that the number of unknowns will be lower (N, N^2 and N^3 respectively). In any case, it is obvious that one needs *at the very least* as many equations as unknown values one wishes to find, which results in a logical introduction of the concept of *oversampling*. That is, for a complex 3D object with N pixels per side, one needs to record at least $2N^3$ values of $F(\mathbf{Q})$. For comparison, in regular crystallography one would record $F(\mathbf{Q})$ only at the Bragg peaks, which are half as numerous *per dimension*; this means that such an object would require measuring $F(\mathbf{Q})$ at eight times as many points in the non-periodic case as it would in the periodic one.

Of course, this is the bare minimum sampling that one needs for the system to be solvable. Furthermore, one still needs to confront the fact that $F(\mathbf{Q})$ itself cannot be obtained from far-field diffraction patterns. One oft cited algorithm to solve this issue was devised by Gerchberg and Saxton, and taken further by J. R. Fienup (Ref. 1). We have seen that, owing to the phase problem, the diffraction pattern alone cannot be used in a straightforward way to obtain the density function of the object. This is why, in this algorithm, one needs to choose an empty region which does not scatter any radiation. This region is usually around the object or, in more technical terms, outside of the support.

In the simplest forms of the algorithm by Fienup, one starts the procedure with some initial map of the object density, $f(\mathbf{r})$; that image may very well be the one resulting from a previous FTH procedure, or a microscope observation.

The first step is to Fourier transform this density in order to obtain the structure factor function (modulus and phase of $F(\mathbf{Q})$). Only its phase will be used later on, so it is common to just start the process from a random set of phases instead. If this is the case, usually the process is repeated from tens of different initial phase sets; obtaining similar results from all of

them is regarded as a proof of consistency.

Either way, once the starting phases are obtained, they are paired with the structure factor moduli from the diffraction pattern. This yields a new F'(**Q**) function, which can in turn be Fourier transformed back and yield a new density function, f'(**r**). This function can then be used as the starting density for the second iteration and so on. In the interest of fast convergence, in the most basic procedures, all points outside of the support have their starting density reset to zero. This is the case of the so-called *error-reduction* algorithm. In the more complex *hybrid input-output* case , the starting density outside of the support includes a mix of the input and output densities from the previous iteration (see Formula 8). This is becomes more clear with Figure 11. In any case, the steps are repeated until the output density functions in two consecutive iterations are arbitrarily similar, in which case we say that convergence has been achieved.

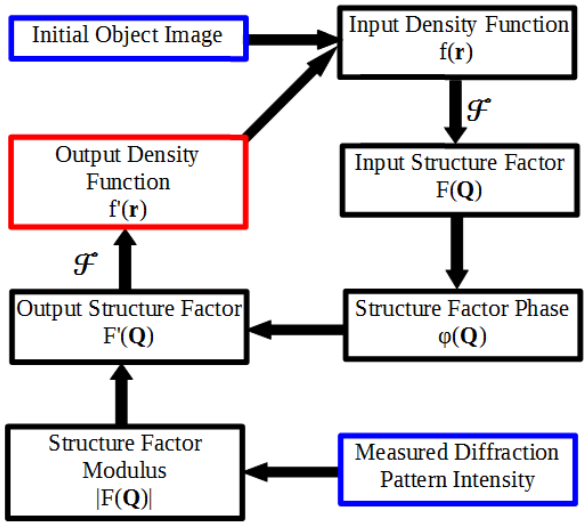

Figure 11: Overview of Fienup's original algorithm. The blue squares indicate the starting knowledge about the object, whereas the red square highlights the eventual result of the iterations.

A great number of uses of this algorithm can be found in the literature. For instance, the work by Miao, Sayre and Chapman shows a direct, elegant application of the hybrid input-output case[xxv], and so it will be reviewed a bit more extensively.

In order to satisfy the need for oversampling, they aid their algorithm through the use of what they call internal and external constraints. Conceptually, these constraints are just a way to feed the algorithm with information that could not otherwise be found within the diffraction pattern data. This is not so different from chemical composition or space group information that is required by many crystalline structure determination routines. The ultimate goal in both cases is to ensure that the solution is unique and that it can be found quickly.

The internal constraints are a delicate issue, because they rely on making very fundamental assumptions about the object. The most trivial of those is the so-called *positivity* constraint, where it is assumed that the object density is a real, positive number at all times. If one assumes this, the system becomes greatly simplified. This constraint is not suitable if one knows that the object density is complex-valued. However, even in that case, it is usually true that both the real and imaginary parts of the density are positive, so one may impose this analogous complex positivity constraint. This is the preferred one in the literature.

The case of external constraints is a more straightforward one; by choosing the region of space that constitutes the support, one then sets all the pixels outside of this region to have zero density. Thus, the number of unknowns in the F-system decreases dramatically.

In a nutshell, Miao and his colleagues define an oversampling ratio σ, between the number of total pixels over unknown (i.e. nonzero) pixels. It is reasonable that for $\sigma>2$ the equation system *might* start to be determinate. A way to look at it is that, for a ratio of exactly 2, the number of unknowns in the system matches the number of equations in the best case scenario. In real cases, a ratio larger than 2 is needed.

The ways all of these constraints are enforced are very similar throughout the literature. A robust example, which is used in Reference 25, involves splitting the output density array into two sets of pixels at the end of every iteration. Set S contains pixels within the support whose current density value obeys all internal constraints. The density of these pixels is copied into the next input density array without any changes. Set S^c contains all the other pixels in the image. Because these pixels do not fulfill one or more of the constraints, they receive a different treatment, shown in Formula 8.

$$f_{j+1}(\mathbf{x}) = \begin{cases} f'_j(\mathbf{x}) & \mathbf{x},\ f'_j(\mathbf{x}) \in S \\ f_j(\mathbf{x}) - \beta f'_j(\mathbf{x}) & \mathbf{x},\ f'_j(\mathbf{x}) \notin S \end{cases}$$

Formula 8: From Ref. 25, generation of the input density array for iteration j+1 (hybrid input-output algorithm). Pixels in S use the output density from iteration j directly, similarly to the error reduction case, but all the other pixels use a linear combination of input and output densities ($0.5 < \beta < 1$).

As a result of this treatment, the density of pixels outside the support will reportedly tend to vanish after a number of iterations. Likewise, for non-positive pixels within the support, the imaginary part of the density will steadily become positive. Thus, the constraints are not necessarily met directly after every iteration, but in principle they will generally be fulfilled by the end of the procedure.

· Practical cases of phase retrieval:

Now that the first approach to iterative phase retrieval has been explained, it is time to see it in action. As a disclaimer, it must be taken into account that, in the real world, diffraction patterns have noise associated to them, which means that the F-system cannot, in general, be solved perfectly; a close solution is the best that can be achieved.

When proposing new algorithms or modifications to existing ones in the literature, it is not unusual to report computer simulations rather than actual diffraction experiments. In these, the (known) density of a test object is used to calculate its diffraction pattern, sometimes adding noise to the photon count, and the phase retrieval starts from there.

In Reference 25, a complex-valued test object was used in such a way. In order to test the effect of changing the size of the support, several runs were performed with five different σ values, all greater than 2. That is, the support size was tuned in order to see the effect it had on the quality of the recovered image.

As one would expect, using smaller supports started over from a lot of known, zero-valued pixels, making the reconstruction process inherently easier. This was also shown with the blank pixels *around* the image, which is the usual way supports are used. However, Figure 12 here has them *within* the image; this is deliberate, as it stresses the fact that their position on the image is not important as long as one defines the support region accordingly.

Saying that the process was "easier", however, is quite a vague way of describing the effect of the support size. In particular, it would be nice to have a way to quantitatively determine the progress of any given reconstruction process. The literature contains a number of ways to do this, but most are similar to the *object-domain error metric*, shown in Fomula 9.

$$E_j = \left(\frac{\sum_{x \notin s} |f'_j(\mathbf{x})|^2}{\sum_{x \in s} |f'_j(\mathbf{x})|^2} \right)^{1/2}$$

Formula 9: For the j-th iteration of the phase retrieval, object-domain error metric by Fienup. It was later referred to just as "reconstruction error" by Miao and colleagues.

This E_j is a real, positive number that can be calculated at every iteration. We already know that at any point, the object image can be divided into two sets of pixels; those that fulfil all the constraints, both internal and external, fall into set S. The rest fall into set S^c. The numerator within E_j is nothing more than the square sum of the densities of pixels within S^c. Analogously, the denominator is the same sum, but running over the S set instead.

This all boils down to E_j being a measure of the amount of reconstructed density that obeys the support and positivity requirements. When the reconstruction starts, it is very likely that many pixels will belong to the S^c set. This means that E_j will initially be a large number.

As the phase retrieval proceeds, more and more pixels will move to the S set. In principle, when the reconstruction has converged we expect many of the pixels to be in the S set. Moreover, at that point, all the pixels that are outside of S should be the ones outside the support, which means that they will have close to zero density; their contribution to E_j is minimal. All in all, E_j should steadily decrease as the iterations go by, hopefully reaching zero by the end.

With that in mind, the smaller supports should intuitively show a slower decrease in E_j within an equal number of iterations. This can be seen in the test by Miao *et al.* in Figure 12, where the smaller support leads to a steep decrease in E_j within fewer than 2500 iterations.

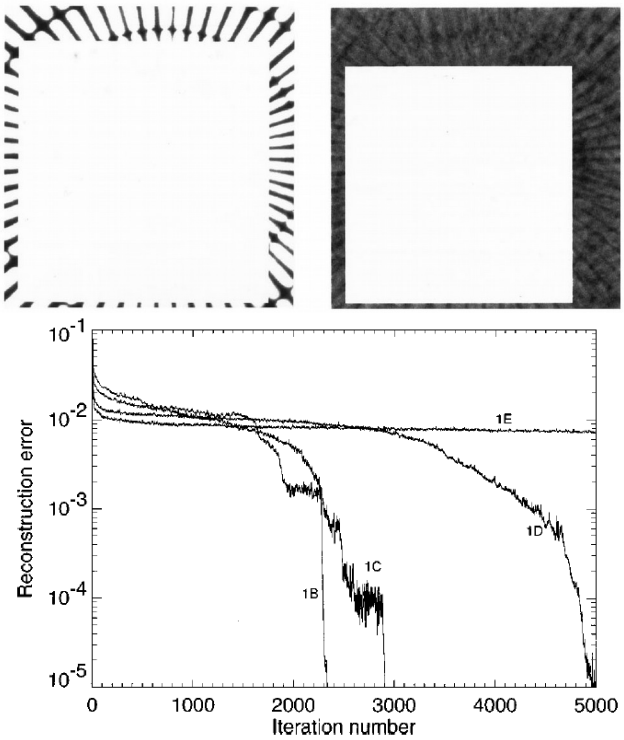

Figure 12: (From Reference 25) Top: recovered density (in modulus) of a complex-valued object. The blank squares contain zero-valued pixels, yielding σ=4 to the left and σ=2.5 to the right. Bottom: reconstruction error for several σ values; curves 1B and 1E correspond to the left and right reconstructions, respectively.

The larger support, however, leads to no remarkable decrease in E_j, signaling that the algorithm is having trouble to reconstruct this low σ (though still greater than 2) system. This is further made obvious by looking at the reconstructed images themselves, and it leads to the conclusion that, indeed, one should always attempt to use a support which is as small as possible.

Therefore, if the boundary of the object is well-known (from a previous image, as mentioned above), it is common to use a tight support in order to ensure a faster convergence. On the other hand, if the boundary is not very steep or only approximately known, one uses a loose support; this is the technical way of pointing out that one knowingly overestimates the size of the object in order to ensure that all of it fits inside the support.

It is now time to move on to imaging of physical objects. In that line, one of the most famous examples of the iterative phase retrieval procedure described so far is the determination of the structure of molecules and small cells[xxvi]; this is sometimes referred to as *single-particle* diffraction.

One such single-particle diffraction experiment was performed by Chapman and it involves what is the most well-known application of CDI so far. That is, the determination of protein structures from diffraction patterns[xxvii]; roughly speaking, this is achieved by having a suspension of the proteins sprayed in the way of the incoming XR pulses.

In that setup, a large number of patterns can be recorded from very similar samples over and over again; after some treatment, these patterns yield the $|F(\mathbf{Q})|$ values. The reconstructed objects can impressively reach sub-nm resolution. Consequently, although the sample being used is significantly different than a regular BHJ, this experiment gives an idea of the resolutions that can be achieved in CDI.

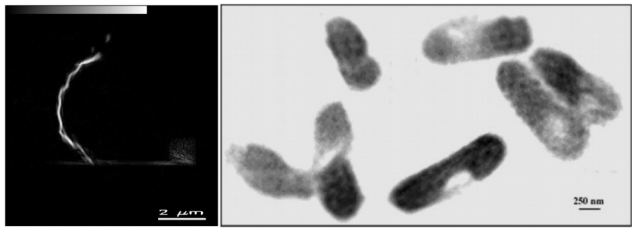

Figure 13. From Refs. 28 and 29, reconstructed images from Spiroplasma melliferum *(left) and* Escherichia coli *(right) cells. Reportedly, the dense regions in the right image are in agreement with confocal microscope measurements in similar cells.*

Less known experiments involved imaging bacteria cells with a resolution close to 70[xxviii] or 30nm[xxix], which is still remarkable in comparison with the resolution achieved in a similar sample using URAs, as shown previously. Moreover, in this case the cells contained manganese-labeled proteins, which reportedly provide excellent electron density contrast. It is undeniable that this knowledge is very valuable if one wishes to perform CDI in solar cells which likewise contain heavy elements, such as those based on hybrid layers or quantum dots. The contrast mechanisms for purely organic solar cells work similarly, as has been mentioned for STXM.

CDI works just as well with inorganic samples. These include, for instance, metal particles[xxx,xxxi] and nanofabricated patterns, such as particle arrays[xxxii] or the silicon nitride sample in Figure 4. The reason this experiment is mentioned again is that it used the so-called *Shrinkwrap* algorithm; this is similar to the

one by Fienup, with the added feature of adjusting the support size as the iterations take place. This is done, in broad terms, by setting a pixel density threshold and periodically shrinking the support so it only contains pixels that meet this threshold.

Undeniably, *some* region needs to be the starting support. A surefire way to find that region is to Fourier-transform the diffraction pattern[xxxiii]; as seen previously, this will yield the AC function of the sample, which by definition has a support *at least* as large as the sample itself. From this point, Shrinkwrap can go on, with the assurance that the starting support contains the whole object *and* that it will eventually become as small (and therefore efficient) as possible.

What is extremely interesting about this procedure is that it arguably requires no previous information on the sample. Although this involves some added computational cost, the fact is that it effectively negates the need for any data besides the diffraction pattern, thus making it one of the most powerful tools available when working with finite samples.

· Imaging of extended samples:

So far, many experimental cases of diffractive imaging have been shown. However, be it holography or CDI, they all had in common that they used finite objects. Even the FTH case of a magnetic multilayer was limited to a circular region surrounded by a gold mask. These finite objects are easily obtained via nanofabrication, or directly bought as test patterns, and they are the samples of choice in a good amount of the literature on XR imaging. However useful they are for explaining the imaging procedures, their possibilities and their limitations, at some point the discussion needs to transition to methods that actually allow the use of an extended sample such as a BHJ.

To that end, a rather intuitive technique would consist on successively illuminating several (overlapping) areas of an extended sample. This is equivalent to using a support which is limited to the illuminated area in each case. Therefore, after obtaining a number of diffraction patterns they can be reconstructed separately as finite objects. That would lead to a set of images which, when put back together, can span a larger area. This technique exists and it is called keyhole coherent diffractive imaging, or KCDI[xxxiv].

KCDI is the logical follow-up to regular CDI. However, in order to work properly, it relies on starting from a very sharp wavefield; in other words, the illuminated region cannot have blurred boundaries. This is achieved through a focussing optic, which invariably means that FZP need to be used. This is not a limitation *per se*, and the next extended imaging method also typically uses FZPs. However, it seems that there should be a better approach to extended samples than just repeating single CDI reconstructions over and over.

In that direction, a method named Ptychography[xxxv] stands out. Like KCDI, it involves recording diffraction patterns from partially overlapping areas of an object. Also like KCDI, it is necessary that the wavefield has sharp edges (a lens or a simple aperture are enough for that), although it has otherwise very few requirements on the illumination. Unlike KCDI, it uses a phase retrieval algorithm[xxxvi] which is quite different from the one discussed so far.

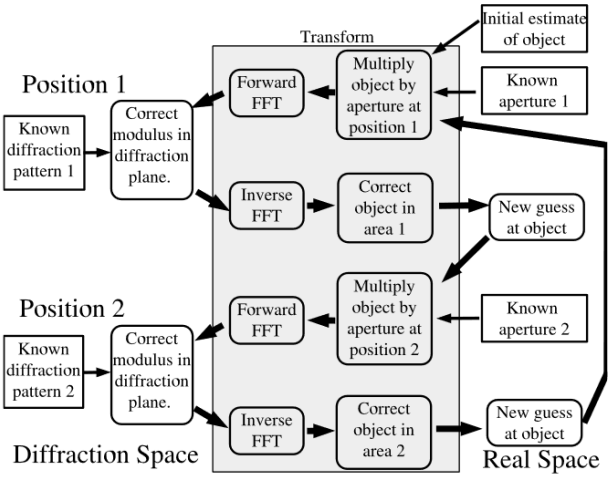

Figure 14: From Ref. 36, Ptychographical Iterative Engine (PIE). Note that, because region 1 and region 2 overlap, their reconstruction processes will be coupled. If they did not, this algorithm would be conceptually equivalent to KCDI.

The way in which it is different is that the reconstruction of contiguous regions is no longer carried out independently. Before calculations begin, one divides the object into several, partially overlapping regions (1 to n), each of which has an associated measured diffraction pattern. Then, the algorithm starts with a density function guess for the whole object. This guess is then multiplied by the illumination (aperture) function for the first diffraction experiment. In practice, this just means

that the object density is weighted so that the only nonzero pixels will be those in region 1. In other words, it conceptually means the same as adding an object support.

Then, the algorithm procedes as usual, Fourier-transforming this density, substituting the moduli for the measured values in diffraction pattern 1, and Fourier-transforming the result back into an object density for region 1. After that, the object density guess is corrected *only in region 1* with this first result (this correction is made using one's choice of several "Update functions"; these are analogous to the β factor in Formula 9). The new guess then undergoes the same treatment for region 2, using diffraction pattern 2, and so on until it has gone through all n regions in some order. Then, one has a new object density function, which is the starting point for the second iteration. In this iteration, regions 1 to n may be updated in a different order, and it would appear that this helps convergence. The iterations go on until convergence is reached.

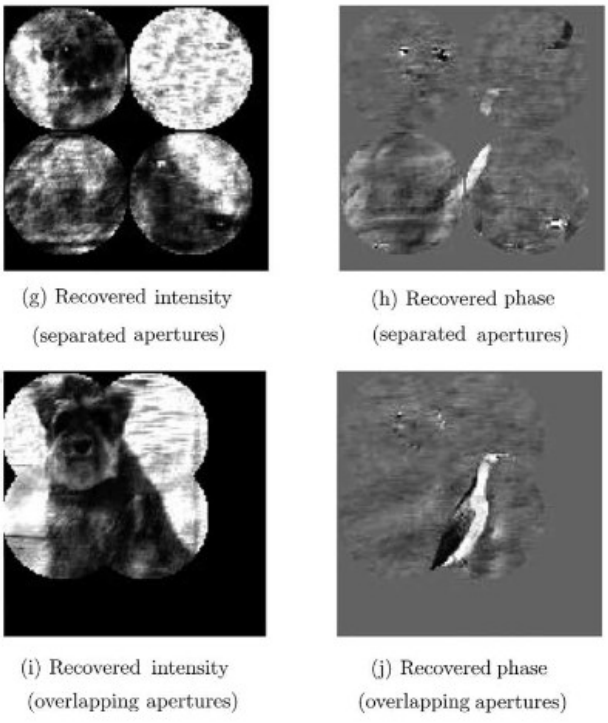

Figure 15: Simulated reconstruction with ptychography from Ref. 36. Overlapping the apertures has worse area coverage but the reconstructed image appears clearer.

It becomes obvious that now the n regions of the object are reconstructed all at once rather than separately. The scheme by Faulkner and Rodenburg in Figure 14 summarizes the procedure in the case of n=2.

According to the authors, the overlap between neighbouring regions, which is usually over 50% in area is what ensures that the global solution is unique in this algorithm. It also helps avoid stagnation, that is, a stationary state which is *not* the solution. Moreover, convergence seems to be faster, in number of iterations, than in Fienup type algorithms. Of course, this all comes at a reasonably larger computational cost, owing to the fact that all regions are corrected in series, whereas KCDI would solve them in parallel.

Much like in regular CDI, the potential resolution of this lensless technique is much higher than those in FTH or XR microscopes. Provided the right experimental setup, Rodenburg and colleagues foresee resolutions below the nm scale. Its main apparent drawback is that one needs to properly characterize the illumination function. Otherwise, it is logical that an error would be introduced at every multiplication step.

That is, unless the algorithm includes corrections for the illumination (probe) function as well. Indeed, the later extended PIE (ePIE) starts the process with an estimate for *both* the object (empty area) and the probe (approximate shape guess)[xxxvii]. EPIE only differs from the algorithm explained above in that this time, whenever the object is corrected, the probe is corrected as well, albeit using an analogous, but different update function.

It goes without saying that this addition is quite reminiscent of the Shrinkwrap algorithm. From computer simulations, it seems that ePIE converges faster and tolerates more noise than older ptychographical engines. That being said, the final notes of this paper will focus on reviewing a few recent experimental instances of ptychography being used outside of simulation environments.

The first one studied the effect of using a semi-transparent beamstop[xxxviii]. Focusing on the PIE details, synchrotron XR with a wavelength of about 1.57Å were diffracted off of a Siemens star (resolution test pattern). For reference, it took 9 minutes to collect diffraction patterns for 441 regions spanning 9μm^2. The resolution exceeded the star's smallest feature (50nm), and it was calculated to be, in the best case, an impressive 12nm.

The second experiment used 2.38Å XR on green alga samples, which were kept at 110K to diminish

radiation damage[xxxix]. XR fluorescence spectra (resolution limited by spot size) and diffraction patterns were measured at the same time. Interestingly, the fluorescence data yielded low-resolution (90nm) elemental composition maps for some light elements, whereas the ePIE reconstructions had calculated resolutions down to 26nm. Although this is not directly related to the OPV case, it nicely shows how, with the proper setup, a ptychography experiment can simultaneously serve as a scanning probe for other measurements.

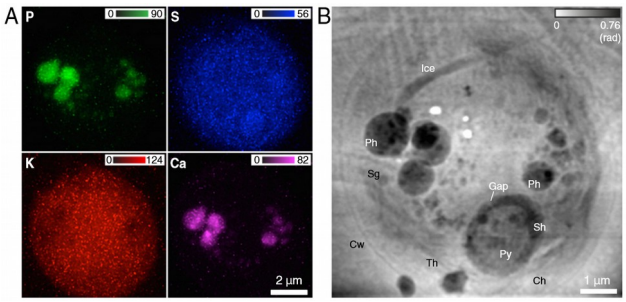

Figure 16: From Ref. 39, ePIE reconstruction of a C. Reinhardtii *alga (right) with its organelles labeled. The fluorescence maps (left) show the distribution of some elements within the cell.*

The last practical case of ptychography that will be mentioned is the work by Shapiro and colleagues[xl], who performed ptychographical reconstructions on inorganic samples using soft XR from a bending magnet. These experiments reach one of the finest resolutions achieved with ptychography to this date. By precise positioning of the probe on the sample and computationally dealing with background noise, a resolution test pattern is resolved down to an astonishing 5nm resolution.

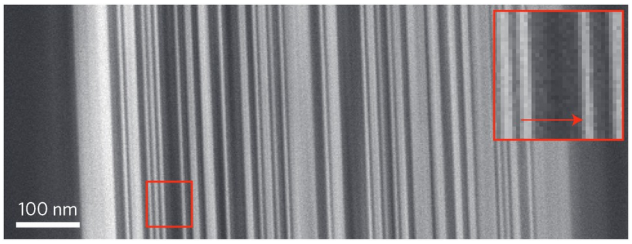

Figure 17: From Ref. 40, reconstruction of a test pattern. The white line to the right is 5nm wide.

Moreover, using the same setup and reconstructing patterns from XR energies close to certain absorption edges seems to allow the introduction of chemical contrast in the images. This is, however, only shown for $LiFePO_4$ particles, a relevant material in some electrodes.

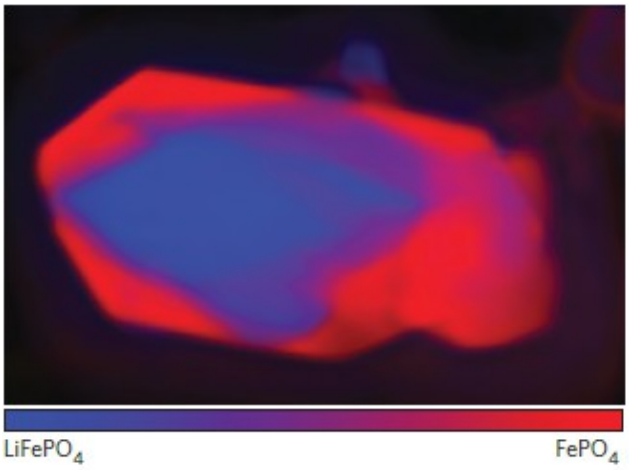

Figure 18: (From Ref. 40) Composition map in the $LiFePO_4$ sample. This map is calculated from several reconstructions at different XR energies around the Fe L absorption edge.

Similar contrast mechanisms for soft XR with organic samples have already been mentioned at the beginning of this paper (see Figure 3). As a consequence, it is not unreasonable to think that such a contrast mechanism could be combined with the high resolution reported by Shapiro. Achieving this would be a good step towards our initial goal of imaging organic BHJ samples.

Conclusions

In this paper, the possibility of using coherent XR diffraction in the imaging of OPV active layer BHJ has been explored. Beyond the goal of BHJ imaging, this paper also intended to give a simple, unified view of (part of) the field of diffractive imaging. This is not always easy; many of the mathematical reasonings have been outright excluded, or just simplified so that the main points could still be made.

A preliminary discussion on electron and XR microscopies has shown some current imaging techniques, their strengths and their limitations. This both serves as a way to review widely known methods and as an exposition vehicle for XR basic concepts.

Afterwards, actual diffraction has come into play. The first of many coherent diffractive methods, XR holographhy, has been considered for our case. Reviewing several cases has shown the possibilities FTH offers; the coupling between resolution and SNR has been addressed, as well as several strategies to solve it. All things considered, it is safe to conclude that FTH boasts incredibly straightforward computational procedures in the case of point references, and that they do not seem to pose any substantial problem either in the case of URAs or FZPs. What is more, generally FTH reconstructions are nonambiguous, which generally does not hold true for iterative methods.

That being said, using complicated, high quality apertures is still required for high-resolution imaging. Combining this with single-shot destructive imaging results in a relatively wasteful process. Even if imaging is non-destructive, a large amount of work is still needed in sample preparations; the low computational cost of the reconstructions (that is, a single FT) does not make up for it. Therefore, FTH in our case can, at most, be used as a stepping stone towards iterative methods.

Moving on to iterative phase retrieval a rundown has been given on the F-system, positivity constraints, oversampling and object supports. Solution non-uniqueness has not been addressed in depth, but it does not appear to be a huge problem in practice, and it is dealt with in some of the methods described. It has been shown that tighter supports perform better overall, and then some single-particle experiments have been reviewed as a means to show both the resolution and contrast capabilities that these technique has. Finally, the Shrinkwrap algorithm has been shown to alleviate the need for previous knowledge about the sample by using a dynamic support.

In spite of occasional slow convergence and moderate computational cost, it has been established that this imaging method has a strong potential in the case of BHJ. The next logical step, that is, taking CDI to extended samples, has been approached from two sides. The first one, KCDI, is the most straightforward, but does not always give the best results. Therefore, the case of ptychography has been presented. In that approach, a more convoluted algorithm seems to result in solution uniqueness and, if the simulations are to be believed, better results overall. Expansion of the algorithm to include probe updates further improves convergence.

At this stage, the ePIE method was already the favourite over both FTH and CDI (including KCDI) methods. This is because, in our mindset, computational cost is generally preferred over actual, physical work on the sample. This is not to say that ptychography measurements require *no* sample preparation, but that they require very little.

Finally, three practical cases of PIE use have been detailed to assess the power of the technique. The first two were not out of the ordinary; in the first one, quite a respectable resolution was achieved, and the second one showed the versatility of ptychography setups. However, the last case is the one which holds the most promise regarding the OPV imaging goal. Not only does it report a resolution of 5nm, well within the realm of BHJ domain sizes, but it also shows application of the absorption contrast capabilities of soft XR. Therefore, it is reasonable to conclude that, although it may not happen in the short run, high-resolution imaging of BHJ in organic solar cells should be possible using XR ptychography.

All in all, the quest for BHJ diffractive imaging using XR has taken the discussion through several approaches. Their working principles and some illustrative results have been accompanied by an exposition of the typical flaws in each method. These limitations have lead to slight changes in the approach, or new methods altogether, until eventually the case of ptychography has emerged as the favourite.

Reference List

i J. R. Fienup. *Reconstruction of a complex-valued object from the modulus of its Fourier transform using a support constraint*. Journal of the Optical Society of America A, 1987. Volume 4, pages 118-123.

ii G.W. Stroke. *Lensless Fourier-Transform method for optical holography.* Applied Physics Letters, 1965. Volume 6, pages 201-202

iii C.J. Brabec, S. Gowrisanker, J.J.M. Halls, D. Laird, S. Jia, S.P. William. *Polymer–Fullerene Bulk-Heterojunction Solar Cells.* Advanced Materials, 2010. Volume 22, pages 3839-3856.

iv M.C. Scharber, N.S. Sariciftci. *Efficiency of bulk-heterojunction organic solar cells*. Progress in Polymer Science, 2013. Volume 38, pages 1929-1940.

v W. Chao, J. Kim, S. Rekawa, P. Fischer, E.H. Anderson. *Demonstration of 12 nm Resolution Fresnel Zone Plate Lens based Soft X-ray Microscopy*. Optics Express, 2009. Volume 17, pages 17669-17677.

vi B.G. Mendis, K. Durose. *Prospects for electron microscopy characterisation of solar cells: Opportunities and challenges*. Ultramicroscopy, 2012. Volume 119, pages 82-96.

vii S.S. van Bavel, J. Loos. V*olume Organization of Polymer and Hybrid Solar Cells as Revealed by Electron Tomography*. Advanced Functional Materials, 2010. Volume 20, pages 3217-3234.

viii C. Jacobsen. *Soft x-ray microscopy.* Trends in Cell Biology, 1999. Volume 9, pages 44-47.

ix H. Ade, X. Zhang, S. Cameron, C. Costello, J. Kirz, S. Williams. *Chemical contrast in X-ray microscopy and spatially resolved XANES spectroscopy of organic specimens. Science, 1992*. Volume 258, pages 972-975.

x F. Livet. *Diffraction with a coherent X-ray beam: dynamics and imaging*. Acta Crystallographica A, 2007. Volume 63, pages 87-107.

xi J.C.H. Spence, U. Weierstall, M. Howells. *Coherence and sampling requirements for diffractive imaging*. Ultramicroscopy, 2004. Volume 101, pages 149-152.

xii M.D. Seaberg *et al. Ultrahigh 22 nm resolution coherent diffractive imaging using a desktop 13 nm high harmonic source.* Optics Express, 2011. Volume 19, pages 22470-22479.

xiii C. Jacobsen, S. Williams, E. Anderson, M.T. Browne, C.J. Buckley, D. Kern, J. Kirz, M. Rivers, X. Zhang. *Diffraction-limited imaging in a scanning transmission X-ray microscope*. Optics Communications, 1991. Volume 86, pages 351-364.

xiv Kahraman Keskinbora *et al. Ion beam lithography for Fresnel zone plates in X-ray microscopy.* Optics Express, 2013. Volume 21, pages 11747-11756.

xv J. Als-Nielsen, D. McMorrow. *Elements of Modern X-ray Physics. 2nd ed.* Chichester: John Wiley & Sons Ltd., 2011.

xvi W. I. F. David, K. Shankland. *Structure determination from powder diffraction data.* International Union of Crystallography, 2008. Volume 64, pages 52-64.

xvii H. N. Chapman *et al. Femtosecond diffractive imaging with a soft-X-ray free-electron laser.* Nature Physics, 2006. Volume 2, pages 839-843.

xviii A.V. Baez. *A Study in Diffraction Microscopy with Special Reference to X-Rays.* Journal of the Optical Society of North America, 1952. Volume 42, pages 756-762.

xix W. F. Schlotter, R. Rick, K. Chen, A. Scherz, J. Stöhr, J. Lüning, S. Eisebitt, Ch. Günther, W. Eberhardt, O. Hellwig, I. McNulty. *Multiple reference Fourier transform holography with soft x rays*. Applied Physics Letters, 2006. Volume 89, article 163112.

xx L. M. Stadler *et al. Hard X-Ray Holographic Diffraction Imaging.* Physical Review Letters, 2008. Volume 100, article 245503.

xxi J. Geilhufe *et al. Monolithic focused reference beam X-ray holography.* Nature Communications, 2014. Volume 5,

article 3008.

xxii S. Marchesini *et al*. *Massively parallel X-ray holography.* Nature Photonics, 2008. Volume 2, pages 560-563.

xxiii A. V. Martin *et al. X-ray holography with a customizable reference.* Nature Communications, 2014. Volume 5, article 4661.

xxiv S. Eisebitt, J. Lüning, W. F. Schlotter, M. Lörgen, O. Hellwig, W. Eberhardt, J. Stöhr. *Lensless imaging of magnetic nanostructures by X-ray spectro-holography*. Nature, 2004. Volume 432, pages 885-888.

xxv J. Miao, D. Sayre, H. N. Chapman. *Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects*. Journal of the Optical Society of America A, 1998. Volume 15, pages 1662-1669.

xxvi J. Miao, H.N. Chapman, J. Kirz, D Sayre, K.O. Hodgson. *Taking X-ray diffraction to the limit: macromolecular structures from femtosecond X-ray pulses and diffraction microscopy of cells with synchrotron radiation*. Annual Review of Biophysics and Biomolecular Structure, 2004. Volume 33, pages 157-176.

xxvii H.N. Chapman *et al*. *Femtosecond X-ray protein nanocrystallography*. Nature, 2011. Volume 470, pages 73-77.

xxviii H.N. Chapman *et al. Coherent imaging at FLASH.* Journal of Physics: Conference Series, 2009. Volume 186, article 012051.

xxix J. Miao, K.O. Hodgson, T. Ishikawa, C.A. Larabell, M.A. LeGros and Y. Nishino. *Imaging whole* Escherichia coli *bacteria by using single-particle x-ray diffraction*. Proceedings of the National Academy of Sciences, 2003. Volume 100, pages 110-112.

xxx I. K. Robinson, I. A. Vartanyants, G. J. Williams, M. A. Pfeifer, J.A. Pitney. *Reconstruction of the Shapes of Gold Nanocrystals Using Coherent X-Ray Diffraction*. Physical Review Letters, 2001. Volume 87, article 195505.

xxxi M.A. Pfeifer, G.J. Williams, I.A. Vartanyants, R. Harder and I.K. Robinson. *Three-dimensional mapping of a deformation field inside a nanocrystal*. Nature, 2006. Volume 442, pages 63-66.

xxxii J. Miao, P. Charalambous, J. Kirz, D. Sayre. *Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens*. Nature, 1999. Volume 400, pages 342-344.

xxxiii S. Marchesini *et al*. *Coherent X-ray diffractive imaging: applications and limitations.* Optics Express, 2003. Volume 11, pages 2344-2353.

xxxiv B. Abbey *et al. Keyhole coherent diffractive imaging*. Nature Physics, 2008. Volume 4, pages 394-398.

xxxv J.M. Rodenburg, A.C. Hurst, A.G. Cullis. *Transmission microscopy without lenses for objects of unlimited size*. Ultramicroscopy, 2007. Volume 107, pages 227-231.

xxxvi H.M.L. Faulkner and J.M. Rodenburg. *Movable Aperture Lensless Transmission Microscopy: A Novel Phase Retrieval Algorithm.* Physical Review Letters, 2004. Volume 93, article 023903.

xxxvii A.M. Maiden and J.M. Rodenburg. *An improved ptychographical phase retrieval algorithm for diffractive imaging*. Ultramicroscopy, 2009. Volume 109. Pages 1256-1262.

xxxviii R.N. Wilke, M. Vassholz and T. Salditt. *Semi-transparent central stop in high-resolution X-ray ptychography using Kirkpatrick-Baez focusing*. Acta Crystallographica Section A, 2013. Volume 69, pages 490-497.

xxxix J. Deng *et al. Simultaneous cryo X-ray ptychographic and fluorescence microscopy of green algae.* Proceedings of the National Academy of Sciences, 2014. Volume 112, pages 2314-2319.

xl D.A. Shapiro *et al. Chemical composition mapping with nanometre resolution by soft X-ray microscopy*. Nature Photonics, 2014. Volume 8, pages 765-769.