LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-TR-210230

Estimates of signals in LCLS diffraction imaging experiments

H. N. Chapman

March 4, 2005

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

Estimates of signals in LCLS diffraction imaging experiments

Henry Chapman, LLNL
December, 2004

In the coherent X-ray diffraction imaging experiments, samples will be injected or placed in the beam and a two-dimensional diffraction pattern will be collected for a single pulse. This is repeated for a large number of pulses, with the data being read out of the detector each pulse, and stored if the data meets a requirement of enough total recorded counts. There must be sufficient pixels in the detector to over-sample the diffraction pattern, which depends on the sample size and desired resolution, as described below. The scattering from the sample covers a large dynamic range: it is strong very close to the central core and at high angles there will be much less than one photon per pixel. Since the technique relies upon classifying and averaging a large number of patterns, the read noise must be considerably less than the photon count per pixel averaged over these patterns. Estimates of the noise level and dynamic range are given below, after first listing the requirements of pixel count and sampling.

Pixel requirements

The pixel requirements simply depend on the number of resolution elements to sample the object of a given size at a given resolution, as described in Huldt (2003), for example. To record to a resolution $f_{max} = 1/d$ requires a maximum scattering angle 2θ given by $\sin\theta = \lambda f_{max}/2$. For an object of finite extent of width D, its molecular transform (in reciprocal space) is band limited. The Nyquist sampling rate of the transform is $1/D$ in each dimension. To measure this transform to a resolution $1/d$, in one dimension, requires samples from $-1/d$ to $+1/d$ or $2D/d$ samples. In real space this corresponds to samples at intervals $\Delta x = d/2$, which is the largest sufficient interval to measure periods larger than d. The detector measures the diffraction intensities, which are the modulus squared of the molecular transform, or equivalently, the Fourier transform of the object's autocorrelation function. For an object of extent D the extent of its autocorrelation is $2D$, which means that the diffraction intensities are band limited with a Nyquist rate of $1/(2D)$. The phase retrieval algorithms do not necessarily require sampling at this rate but experimental experience shows better results with higher sampling. Note that sampling at a higher rate than $1/(2D)$ does not add any information to the measurement, but may increase the signal to noise of the measurement. However, pixellated detectors do not sample at points but integrate over the active area of the pixels. This corresponds to a Modulation Transfer Function (MTF) that may decrease to zero at spatial frequencies (at the detector) of period $2p$, where p is the pixel width. The effect of the MTF is to apply an envelop to the reconstructed real-space image, which should be no less than 0.7 at the largest radial extent of the object. As such, the detector's MTF influences the required pixel count. The number of pixels along the width of the detector is given by $N = 2Ds/d$, where s is a sampling ratio per dimension (relative to the molecular transform Nyquist rate), with $s = 2$ in the case of maximum required sampling (for which the 0.7 MTF level should occur for pixel frequencies no lower than $1/(4p)$).

We estimate the maximum requirement for number of pixels is $N = 2000$, which corresponds to a particle size of 100 nm at a resolution of 1/(0.3 nm) and a sampling ratio of $s = 3$, or a particle size of 200 nm at the same resolution and a sampling ratio of $s = 1.5$. The larger sampling would be required if the detector MTF at $1/(2p)$ is about 50%. These are likely parameters for imaging of nanoparticles, and for the imaging of arrays of biological particles. For the the cow-pea mosaic virus (CPMV) test object described below, which has $D = 32$ nm, we require $N = 450$ pixels for $s = 2$ and a resolution of 1/(0.3 nm). This reduced pixel count will be sufficient for most biological samples, and a larger pixel count detector will only add to the readout time in these cases. In general we would prefer modules of 512 × 512 or 1024 × 1024 pixels.

Image reconstruction can be achieved with considerable missing data due to a beamstop. However, the larger this region, the larger the uncertainty of various components of the image, and the less quantitative the image. If the beamstop covers no more than the central speckle, then the only missing data is essentially F_{000}. The speckle size for an object of width D is $1/D$ or $2s$ pixels in width. The dynamic range values given below were based on patterns with the central $2s \times 2s$ pixels excluded.

Simulations

Detector signals were computed for a test sample of a cow-pea mosaic virus (CPMV), with labeling of nanoclusters of gold [Wang, 2002]. This may be an early test sample for LCLS experiments. The atomic coordinates of the virus capsid (1NY7) were obtained from the EMBL-EBI Macromolecular Structure Database (http://pqs.ebi.ac.uk/pqs-bin/macmol.pl?filename=1ny7). The capsid structure is hollow (since the DNA structure inside is unknown), and this was filled in with carbon atoms in random locations and average density of 1.3 g/cm^3 (less mass than DNA). When the gold labeling was applied, clusters were attached to the 65 symmetry sites (CYS 295) as described in Wang (2002). Each gold cluster was spherical with a diameter of 1.4 nm diameter, and contained 82 gold atoms (density of 18.8 g/cm^3). The gold increases the total scattered photons by less than 10%. The CPMV has a diameter of 32 nm, and a total molecular weight of 14.8 MDalton (13.7 MDalton without the gold labeling). Larger sized samples have been simulated by arraying the CPMV structure in ordered and disordered groups of particles. The incident beam was modeled as a single-mode coherent Gaussian beam, with a waist diameter of 0.1 or 0.2 micron and total flux of 10^{12} photons. The diffraction patterns were computed for a wavelength of 0.15 nm, with a painfully slow program that computes the scattered intensity from a collection of atoms illuminated by a focused single Gaussian mode, in the Born approximation (and no atom motion or ionization):

$$I(\mathbf{k}_{out}) = \Omega r_e^2 I_0 \left| \sum_j f_j(\mathbf{q})\, u(\mathbf{x}_j) \exp\{i\mathbf{k}_{out} \cdot \mathbf{x}_j\} \right|^2$$

where f_j is the structure factor for the j^{th} atom, r_e the electron radius, Ω the solid angle of a pixel, and an incident field described by [Seigman, 1986]

$$u(\mathbf{x}) = u(\mathbf{r};z) = \sqrt{\frac{2}{\pi}}\frac{1}{w(z)}\exp\{-ikz + i\psi(z)\}\exp\{-r^2/w^2(z)\}\exp\{-ikr^2/(2R(z))\},$$

with $w(z) = w_0\sqrt{1+(z/z_R)^2}$; $R(z) = z + z_R^2/z$; $\psi(z) = \tan^{-1}(z/z_R)$.

The parameter w_0 is the waist radius (radius at which intensity is $1/e^2$), and $z_R = kw_0^2/2$ is the Rayleigh range. The length $R(z)$ is the radius of curvature of the wavefront at a distance z, along the propagation axis, from the waist. The wave-vector magnitude is defined as $k = 2\pi/\lambda$, and the momentum transfer $\mathbf{q}$ is given by $\mathbf{q} = \mathbf{k}_{out} - \mathbf{k}_{in}$ with

$$\mathbf{k}_{in} = [\mathbf{x}_j, -R(z_j)]\, k/\sqrt{\mathbf{x}_j^2 + R^2(z_j)}\,.$$

Note that $\int_0^\infty 2\pi r\, u^2(r;z)\, dr = 1$, and thus the parameter I_0 is the total number of photons in the beam. In the simulations the intensity I is quantized and photon noise added (normally distributed with standard deviation $\sqrt{I}$).

Signal

Results of simulations are shown in Fig.1. For these calculations, the center of the particle was always positioned in the center of the beam, at the z location of minimum waist ($z = 0$). For the calculations shown here we used $I_0 = 10^{12}$ and $w_0 = 0.1$ μm (0.2 μm diameter waist) and $w_0 = 0.05$ μm (0.1 μm diameter waist). The calculations were carried out for an X-ray wavelength of 0.15 nm. The simulated array size was 422x422 pixels, with $s = 2$, corresponding to a resolution of 1/(0.3 nm). Most of the pixel values are zero or one photon, and the maximum and total photon counts (excluding the central 4x4 pixels) was 1530 and 7.3 x 10^4, respectively, for the 0.2 μm diameter waist and 5540 and 3.0 x 10^5, respectively, for the 0.1 μm diameter waist.

Note that larger samples do not necessarily give larger signals; there are a fixed number of photons per pulse and samples larger than the beam will require a proportionally larger beam and hence lower fluence. Larger signals will be achieved with thicker objects and objects of higher-Z, and as such the experiments of nanoscale inorganic samples will give stronger signals. Larger signals will also be achieved with arrays (2D or 3D) of particles, due to coherent addition in the Bragg directions. Similarly, the imaging of large objects at low resolution (e.g. single-shot imaging of micrometer-sized cells beyond the radiation damage limit) will produce larger photons per pixel due to the coherent addition, in the forward direction, of scattering from atoms within a single resolution voxel [Sayre, 1995]. For the case of crystals and arrays of identical unit cells, the photon count will increase in the Bragg peaks by a factor n^2, where n is equal to the total number of unit cells illuminated. The signal between the Bragg peaks will only increase in proportion to n, but this signal can be built up by averaging, once classification has been achieved on the Bragg peaks. For a 2D crystal with n unit cells, if the beam size matches the object size then the incident fluence will be proportional to $1/n$ and the Bragg peaks will increase in photon count by n, not n^2. For the simulations we have run on 2D and 3D

crystals of 5×5 and 5×5×5 unit cells, we typically see a factor <10 increase in photon count. In all these cases, the pattern is strongest near the zero frequency and locally (within a 10x10 pixel patch) the intensity changes at most by a factor of 1000. Therefore, it seems feasible to record patterns including the weakest to the strongest intensities by placing an absorber of a specific thickness profile in close proximity to the detector surface, in order to attenuate the strongest regions of the pattern.

Noise

Datasets will be assembled by classifying patterns into classes of like-orientation and averaging patterns within each class. Huldt et al. (2003) showed that accurate classification could be performed with as little as 0.1 photon counts on average per pixel, at the highest resolution of the pattern. For the CPMV particle simulation, this means we can classify out to the full resolution of the simulation, of 1/(0.3 nm), for the case of a 0.1 μm diameter waist (Fig. 2). This requires that when we sum together ten diffraction patterns the accumulated noise must still be less than one photon. Choosing a noise level of 0.2 photon (SNR = 5, Rose criterion) in the ten-frame sum, the noise per pixel for each pattern should be no larger than $0.2/\sqrt{10} \approx 0.06$ photons per pixel.

Summary of requirements:

Number of pixels: 1024 × 1024, and up to 2k × 2k
Noise per pixel: <0.06 photons
Maximum signal: $<10^4$ photons in a pixel (locally $<10^3$ intensity range)
Total data rate: $<10^6$ photons
MTF: >70% for modulations of period 4*p*, where *p* is the pixel pitch
Hole in detector: 4 × 4 pixels

Acknowledgements

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-ENG-48.

References

[Huldt, 2003] G. Huldt, A. Szoke, and J. Hajdu, "Diffraction imaging for single particles and biomolecules, "J. Struct. Biol. **144** 219-227 (2003).
[Miao, 1998] J. Miao, D. Sayre, and H. N. Chapman, "Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects," J. Opt. Soc. Am. **15** 1662-1669 (1998).
[Sayre, 1995] D. Sayre and H. N. Chapman, "X-ray microscopy," Acta Cryst. A. **51**, 237–252 (1995).
[Siegman, 1986] A. E. Siegman, "Lasers," University Science Books, 1986.
[Wang, 2002] Q. Wang *et al*., "Icosahedral virus particles as addressable nanoscale building blocks," Angew. Chem. Int. Ed. **41**, 459-462 (2002).

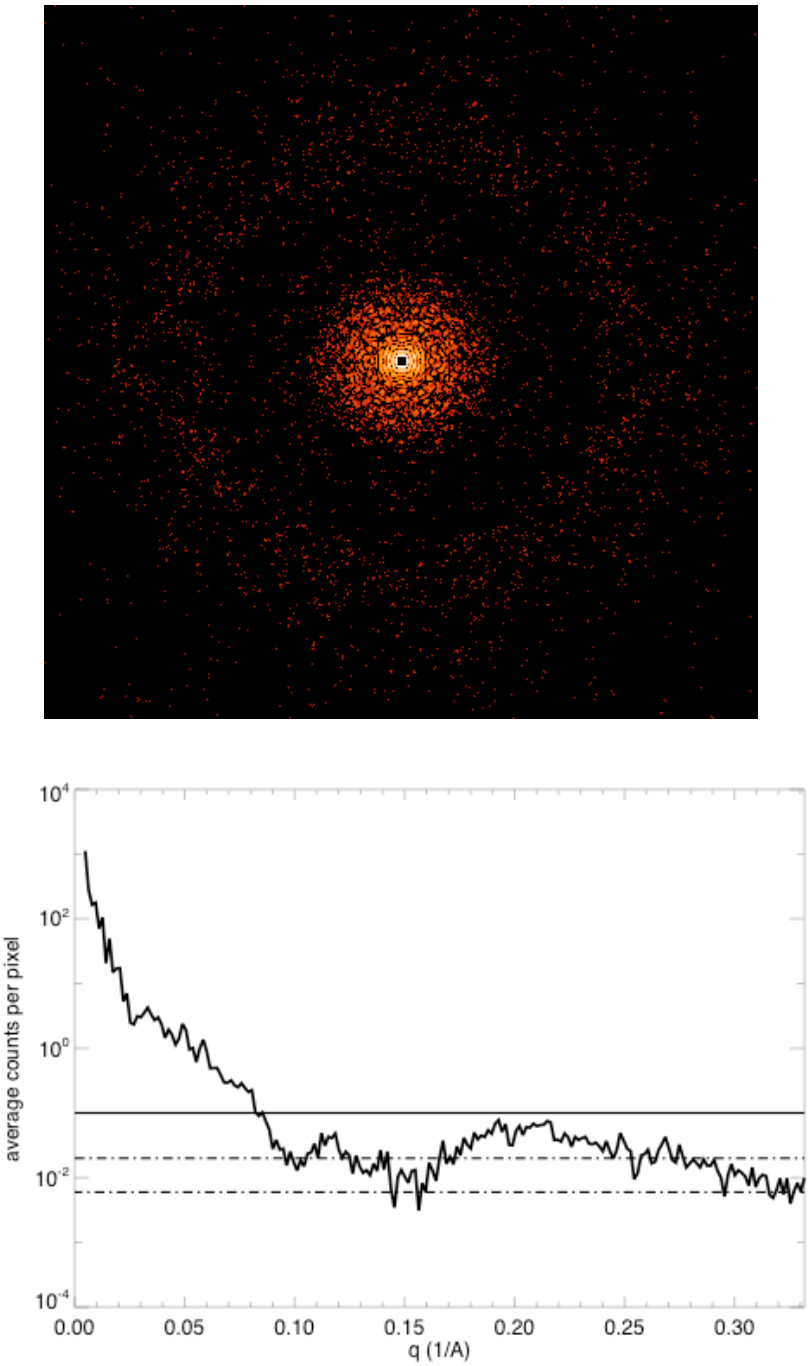
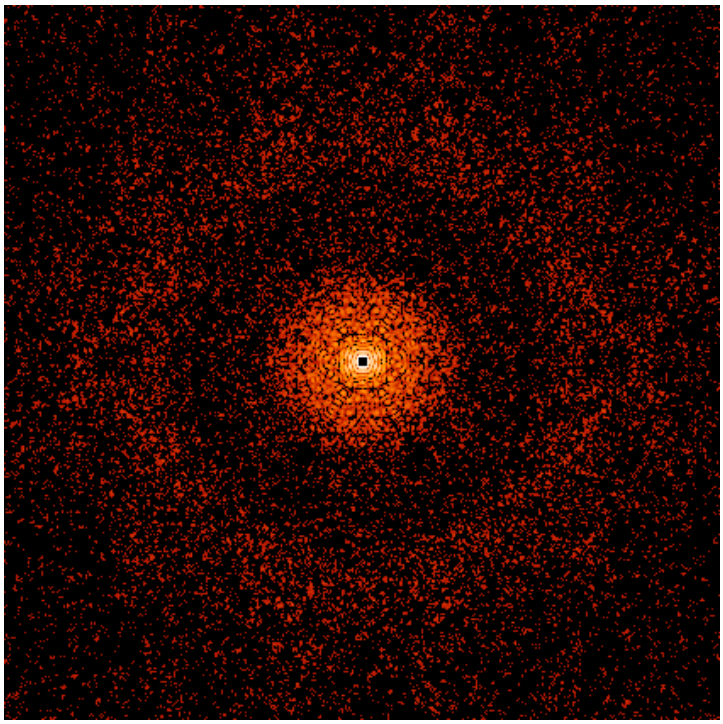


Fig. 1: Simulated diffraction data from the CPMV object, for a beam waist diameter of 0.2 μm. The total photons in the incident beam was 10^{12}. The array is 422x422 pixels, corresponding to s=2, and a resolution of 1/(0.3 nm). The central 4x4 pixels were blocked. The total integrated photons is 7.3×10^4. The intensities are displayed on a logarithmic greyscale. Away from the center most pixels have one or no photons. The maximum photon count is 1530. The plot shows the radial average of the photon counts. The solid line is at 0.1 counts, the minimum counts needed to classify, and the dotted lines denote the noise level for 10 and 100 averages, for a detector noise of 0.06



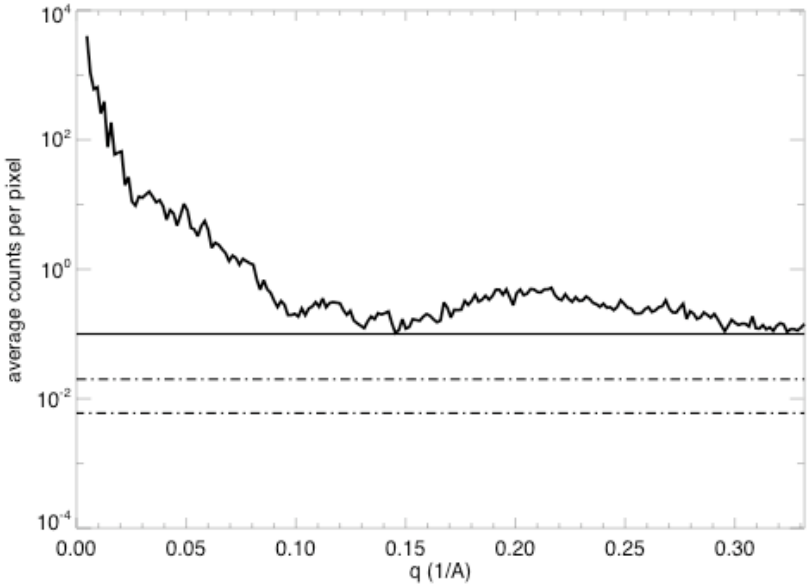


Fig. 2: Simulated diffraction data from the CPMV object, for a beam waist diameter of 0.1 μm, and all other parameters the same as for Fig. 1. The maximum photon count is 5540, excluding the beamstop. The total integrated photons is 3.0×10^5. The plot shows the radial average of the photon counts. The solid line is at 0.1 counts, the minimum counts needed to classify, and the dotted lines denote the noise level for 10 and 100 averages, for a detector noise of 0.06 photons per pixel.