MODERN X-RAY SOURCES: FROM BASICS TO DATA ANALYSIS

O.M. Yefanov
Center of Free Electron Lasers, Notkestr. 85, 22607 Hamburg, Germany,
oleksandr.yefanov@desy.de

The range of applications covered by modern x-ray facilities stretches from atomic physics to biological processes. The mostly used techniques are: diffraction/scattering, spectroscopy and imaging/microscopy. We proposed to use additional scattered intensity between Bragg peaks, to improve the resolution of reconstructed structure and, in future, to develop a new method of ab initio structure determination that can solve the phase problem of classical structure determination.

FELs and 3rd generation synchrotrons

Modern x-ray sources, such as Free Electron Lasers (FELs) and synchrotrons, are widely used for material and life sciences applications.

3rd generation synchrotrons (in fact storage rings) are being built around the Globe during last 20 years and provide big variety of end station for different techniques that require x-ray beams of high brilliance and in the range of energies from 100s ev to 100s keV. New generation of synchrotrons, like MAX IV and NSLS 2, sometimes referred as 3.5th generation, are being currently commissioned.

X-ray Free Electron Lasers (XFELs) are relatively new facilities that utilize similar principles for generating x-rays as 3rd generation storage rings, but with small modification – without a ring. Such device consists of linear accelerator of electrons (Figure 1) and undulator system to generate photons. In this way it is possible to avoid spreading of electron bunches (mostly horizontally) that is typical for synchrotrons. This leads to much higher transverse coherence of XFELs. Also undulator system works in saturation mode and energy of electrons bunches is higher, therefore peak brilliance of XFELs is several orders of magnitude higher than at synchrotrons. One more unique feature of XFELs is the photon pulse length which is of the order of 1-100fs. This allows performing time-dependent measurements at a scale not available at any other x-ray (or electron and neutron) generation facility.

Fig. 1. 2.1km of accelerator modules in the tunnel of European XFEL (Hamburg)

Even though peak beam parameters are much better at XFELs there is one huge disadvantage of these facilities originating from its linear design – maximum number of experiments that can be done simultaneously is 1 or 2. With new multiplexing technique this parameter can be increased to about 5-6. At the same time this parameter for a synchrotron is usually of the order of 50.

Beam manipulations

For most of experiments a monochromatic x-ray beam is being used. Design of monochromators didn't change much during last 100 years – one or usually several perfect crystals (mostly silicon) are used. For soft x-rays gratings are utilized.

While experiments performed with monochromatic beam are easier to interpret, more and more development is carried out in the field of application of pink beam – the beam generated by undulator that has typically about 5% bandwidth. This allows to increase the flux by 3 orders of magnitude and therefore makes it possible to use native picosecond structure of electron bunches that are typically used in modern storage rings. One such bunch can produce up to 1e8 pink photons, that is enough for many experiments and allows time resolved measurements at picosecond scale. Of course such experiments require more sophisticated data analysis procedures that are being developed.

While monochromators didn't change much, focusing optics was greatly improved during last 20 years. KB mirrors, 1D and 2D CRL lenses and zone plates are widely used at all modern facilities to make a beam of down to 50nm. New lenses based on Laue diffraction are being developed and can produce 5nm spot.

Detectors

One of the most important parts of any experiments with x-rays is detection of scattered photons. For this purpose 2D detectors are usually used. In former times CCD-based detectors were widely spread, while now more and more facilities are being equipped with so called pixel detectors. In this type of detectors signal from each pixel is read individually and this allows to greatly improve the speed of the detection.

High speed of detection as well as wide dynamical range and single photon sensitivity are desired for most x-ray experiments. With very intense beam, available at modern x-ray sources, many measurements can be done in fraction of a second. Even protein crystals usually produce enough diffracted photons in milliseconds at 3rd generation synchrotrons. And those crystals survive just a couple of minutes, therefore high speed of the detectors is highly demanded. At the same time measuring of weak signals requires noise-free detectors with single photon sensitivity. And wide dynamical range allows to measure signals with very different intensity, like Bragg peaks together with diffuse scattering. Also it is needed to compensate the fact that diffracted intensity in most cases rapidly decreasing with increasing of scattering angle. Modern counting detectors, that simply count photons incident at each pixel with very high (several MHz) rate, such as Pilatus, Eiger, MediPix-based satisfy all these parameters for measurements at synchrotrons.

Some scientific applications also require energy sensitivity or small pixel size. Special detectors are usually designed to meet these requirements.

FELs cannot use modern counting detectors developed for synchrotrons because pulse length is just 10s of femtoseconds and it is currently impossible to count photons in such small time period. Instead separate brunch of integrating detectors is developed and used (different CCDs like mpCCD, pnCCD, pixel detectors like CS-PAD, AGIPD, JoungFrau, ePix, etc.). These detectors just integrate the charge produced by the scattered photons, thus allowing to measure very short pulses.

The cost of modern detectors is high (from $100k to $10M), but many experiments are simply not possible without such detectors.

Data collection and processing

High rate of data collection requires new strategies for data processing and storing. Most facilities already faced the difficulties arising from usage of new fast detectors. Quite often the whole data transfer and storage system has to be re-designed. Just some numbers to understand the scale of the problem. Pilatus 6M detector at full speed (25Hz) will produce at least 150MB/sec (0.5TB/hour) of 4 times compressed (!) data per hour. Therefore 1Gbit network connection is already not enough and in 5 days user can collect 60Tb of data. At LCLS with CS-PAD detector current experiments often produce 100TB of data during a standard 2.5 full days beamtime. And this will only get worse – for example Eiger detector (1Mp, 3kHz) can produce 20 times more data then Pilatus 6M (3GB/sec) and new FEL detectors like AGIPD and ePix can fill hard drives even faster.

With such high speed and huge data volume manual data processing is not feasible anymore. Therefore both online and offline data processing has to be automated. First of all some “on-the-fly” (online) data analysis has to be done during the measurements to understand if the experiment is going well. It is difficult to overestimate the usefulness of such analysis – a lot of experiments failed completely without proper online feedback. And one has to remember – beamtime at an FEL is precious and if the experiment fails, there is only a small probability to get a new beamtime and even if you’re lucky the new beamtime can be allocated not earlier then in one year. For FELs several packages like OnDa, Cass, Hummingbird, etc. were developed by different groups. When some technique is well established, some features of community development software became a part of facility provided pipeline.

Collected data has to be either processed onsite via remote access or transferred for further processing. For onsite analysis the facility has to take responsibility to provide users enough computational power and space and fast remote access. From the other hand huge data volumes makes the task of data transfer rather complicated – system like Globus Online are used to speed up the process. And also a user has to have enough computing power and storage.

Next step is offline data processing. For different well established tasks, like classical protein crystallography, there are already quite a lot of tools that do processing (or some parts of processing) rather efficiently and quite often almost automatic. Commercial and open source packages exist for such tasks. But for all non-standard experiment all data processing has to be done by the user. And once again: manual processing of such amount of data is often rather bad idea. So highly parallelized software, sometimes even special hardware (for example supporting CUDA or PhysX), have to be used.

Scientific applications

Our group is mostly interested in development of new methods for structure determination [1-4]. Recently we have a lot of progress in developing new strategies to get more information from analysis of diffraction data. Classical structure analysis uses crystals and analyzes diffracted Bragg peaks intensities to solve the structure [5]. We proposed to use additional scattered intensity, like shown at Figure 2, to improve the resolution of reconstructed structure [6] and, in future, to develop a new method of ab initio structure determination that can solve the phase problem of classical structure determination.

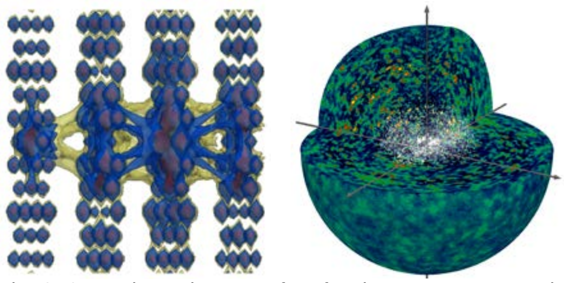

Fig. 2. 3D reciprocal space of: Left: Photosystem I crystals showing both Bragg peaks (red) and truncation rods between peaks (blue and yellow). Right: Photosystem II crystals, Bragg peaks (white) and continuous scattering

References

1. *Chapman H.N.* // Nature. 2007. 448 (7154), P. 666-671.
2. *Chapman H.N.* // Science. 2008. 321 (5887). P. 352-353
3. *Chapman H.N.* // Nature. 2010. 467 (7314), P. 409-410.
4. *Chapman H.N., et al.* // Nature. 2011. V. 470. P. 73-77.
5. *Yefanov O., et al.* // Phil. Trans. B. 2014. V. 369. P. 20130333.
6. *Ayyer K., Yefanov O., et al.* // Nature. 2016. V. 530. P. 202-206.