Streaming Data Analysis Tools to Study Structural Dynamics of Materials

Sameera Abeykoon

BNL

Yugang Zhang, Eric Dill, Thomas Caswell, Daniel Allan, Arman Akilic, Lutz Wiegart, Stuart Wilkins, Annie Heroux, Andrei Flerasu and Kerstin K. van Dam.

McGill University

Mark Sutton

August 16 , 2016.
NYSDS (2016)

BROOKHAVEN NATIONAL LABORATORY

a passion for discovery

U.S. DEPARTMENT OF ENERGY | Office of Science

Outline

- NSLS-II (National Synchrotron Light Source –II)
- Scikit-beam
- Coherent Hard X-ray beamline (CHX)
- Speckle patterns
- Data analysis tools for X-ray Photon Correlation Spectroscopy (XPCS)
  One-time correction, Two-time correlation
- Streaming data analysis tools
- Summary

National Synchrotron Light Source-II (NSLS-II)

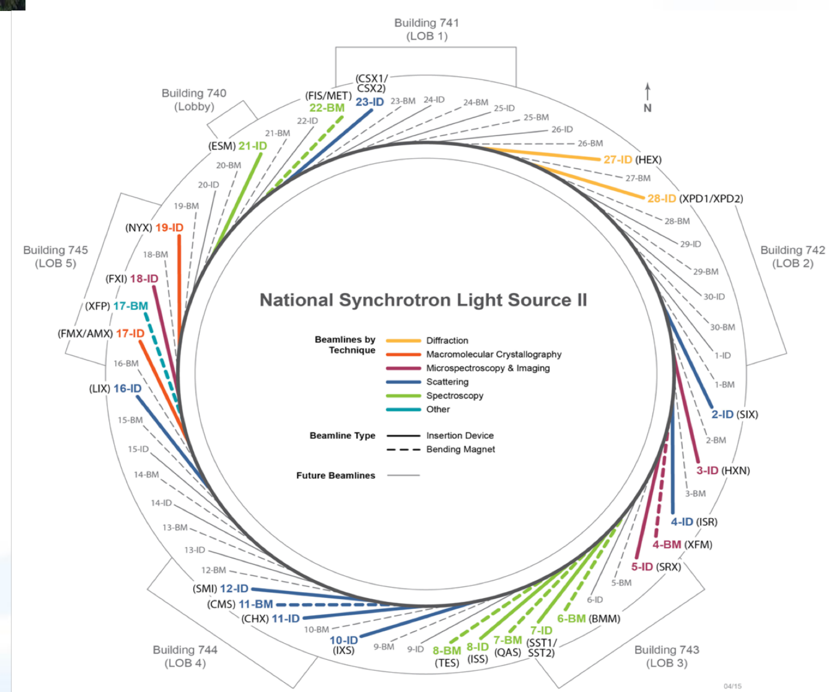
Ultra-high brightness synchrotron source

New studies, new X-ray techniques

Data rates GB/hour and several TB/day

50 – 60 beamlines

Need of "basic experimental toolkit" for data management and data analysis.



Scikit-beam

NSLS-II is developing a high performance data management structure capable of handling, processing visualization, analysis and storage of high-throughput data to optimize the scientific productivity of its beamlines.

CSI also plays a big role in this effort.

Scikit-beam is an analysis library developed by this effort.

Analysis tools for X-ray, electron, and neutron scattering techniques

Python package (Scientific data analysis)

Primary Objective - Minimize technique-specificity and maximize code re-usability.

Initiated by NSLS-II, DAMA group – Converted into multi-facilities effort with contributors from other DOE laboratories.

Scikit-beam and Scikit-beam examples

Scikit-beam supports all levels of user expertise, from novice to developer.

100% test coverage to all our analysis codes.

Full documentation for the users and developers.

We also provide use case examples for the tools included in the Scikit-beam. Two formats jupyter(a.k.a. ipython notebook) and python scripts.

https://github.com/scikit-beam/scikit-beam

The data analysis tools described here focus on X-ray photon correlation spectroscopy for coherent X-ray scattering.

Coherent Hard X-ray Beamline (CHX)

Third generation synchrotron sources enabled coherent X-ray scattering experiments.

X-ray source is 3m long in-vacuum undulator with 20mm magnetic period.

Double crystal Si(111) monochromator. Operational energy: (6 – 16)Kev.

Relative bandwidth $(\Delta\lambda / \lambda) = 10^{-4}$

Flux $\approx 10^{11}$ photons per second for the storage current, 250mA.

Dectris Eiger 4M area detector, detection rate of 750Hz resulting in an uncompressed data rate of 4.5 GB/s.

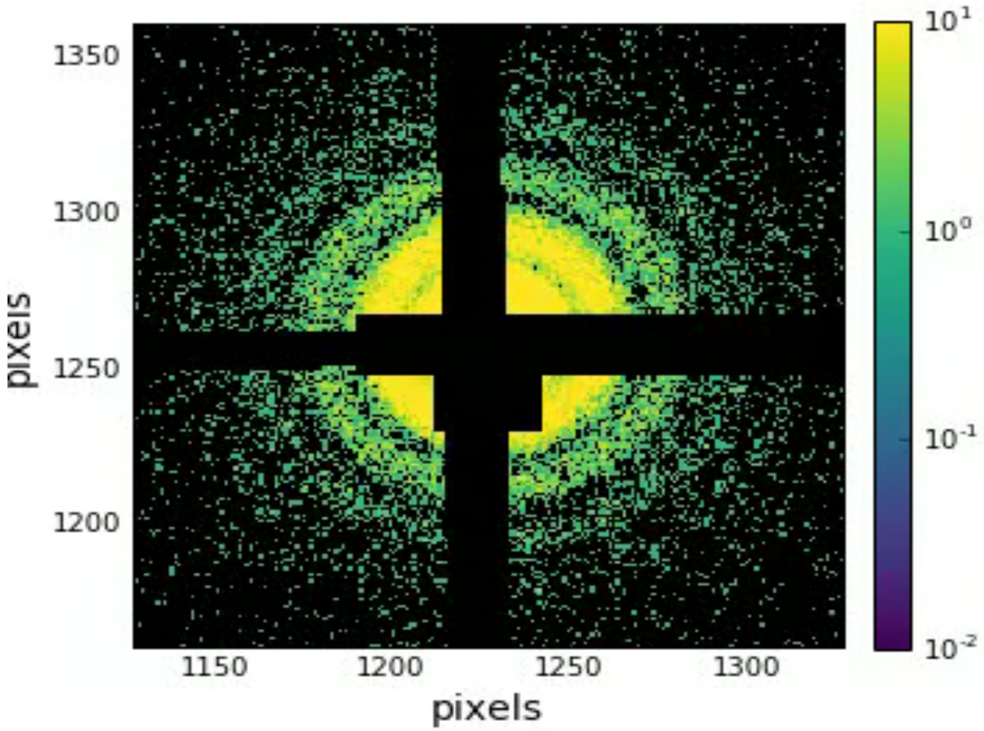
CHX beamline will allow studies of dynamics on time scales that can be 100 times faster and on shorter length scales than currently achievable elsewhere.

Material Science

- Structure-property relationship in disordered materials.
- Non equilibrium materials.
- Behavior of the materials, defects

Ex: Polymers, colloids, bio-memberance, bio polymers,

Speckle Patterns



Coherent light scattered by a heterogeneous sample produce a speckle pattern characteristic of the instant spatial arrangement of the scatterers.

The resultant speckle intensity changes as the scatters fluctuate over time.

The time auto correlation functions of the speckle fluctuations reveal the scattering wave-vector dependent timescales associated with the motion of the scatterers.

X-ray Photon Correlation Spectroscopy (XPCS)

Powerful experimental tool to

1. Study nano- and meso-scale dynamics of materials.
2. Study both equilibrium and non-equilibrium dynamics.

Can be used to measure structural dynamics of materials for long time (until retirement). It can be measure structural dynamics over wide range of time scales (e.g. from microseconds to minute)

Compared to other techniques that are currently being used to measure structural dynamics, XPCS access a unique combination of length and timescales which are difficult or impossible to measure otherwise.

One-time correlation

In XPCS, the sample dynamics are quantified by the normalized one-time correlation function.

$$g^{(2)}(q,\ \mathrm{t}) = \frac{\langle I(q,\ \mathrm{t}_0)\, I(q,\ \mathrm{t}_0 + \mathrm{t}) \rangle}{\langle I(q,\ \mathrm{t}_0) \rangle^2}$$

Where;
<…> - ensemble average over the “initial” time t_0 over nominally equivalent pixels.

$I(q,\ t_0)$ – scattered intensity at momentum transfer, q and time, t_0

One-time correlation

Intermediate scattering function (ISF) or dynamic structure factor $g^{(1)}(q,t)$ is calculated in theoretical simulations.

$$g^{(2)}(q,\ \mathrm{t}) = \beta_2\left[g^{(1)}(q,\ \mathrm{t})\right]^2 + g_\infty$$

β_2 is the correlation function contrast factor and g_∞ is the baseline which is equal to 1 for ergodic samples.

ISF describes all dynamics information about a system. Therefore, XPCS method provides a direct measurement of the nano- and micro- scale dynamics of the scatterers.

One-time correlation

When the scatterers are undergoing a simple free-diffusive (a.k.a. Brownian) motion, the ISF can be described by a simple exponential decay.

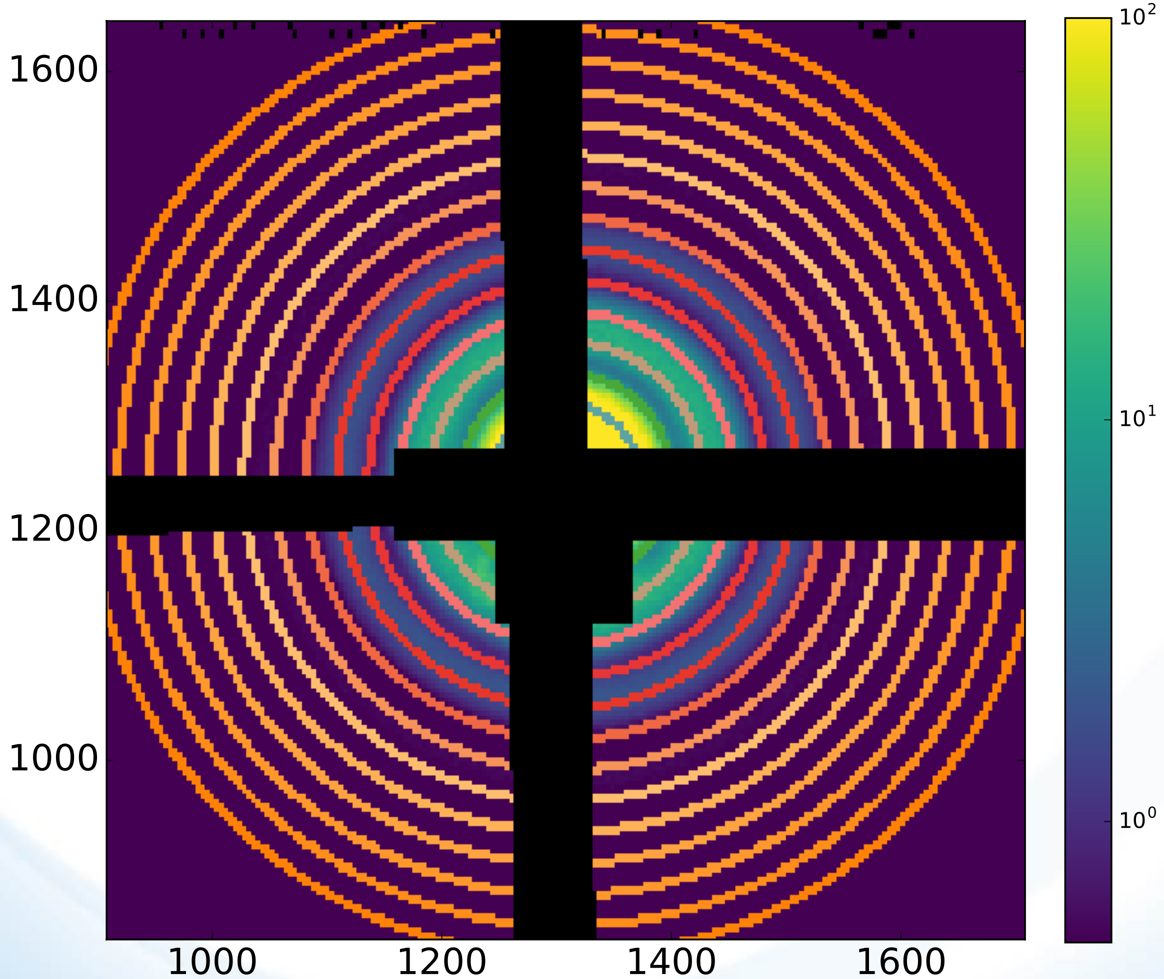
$$g^{(1)}(q,\ t) = \exp[-\gamma(\mathrm{q})\mathrm{t}]$$

$\gamma(q)$ is the relaxation rate related the diffusive coefficient D_0 and the momentum transfer q by

$$\gamma(q) = D_0 q^2$$

$$g^{(2)}(q,\ t) = \beta_2 \exp\left[-2D_0 q^2 t\right] + g_\infty$$

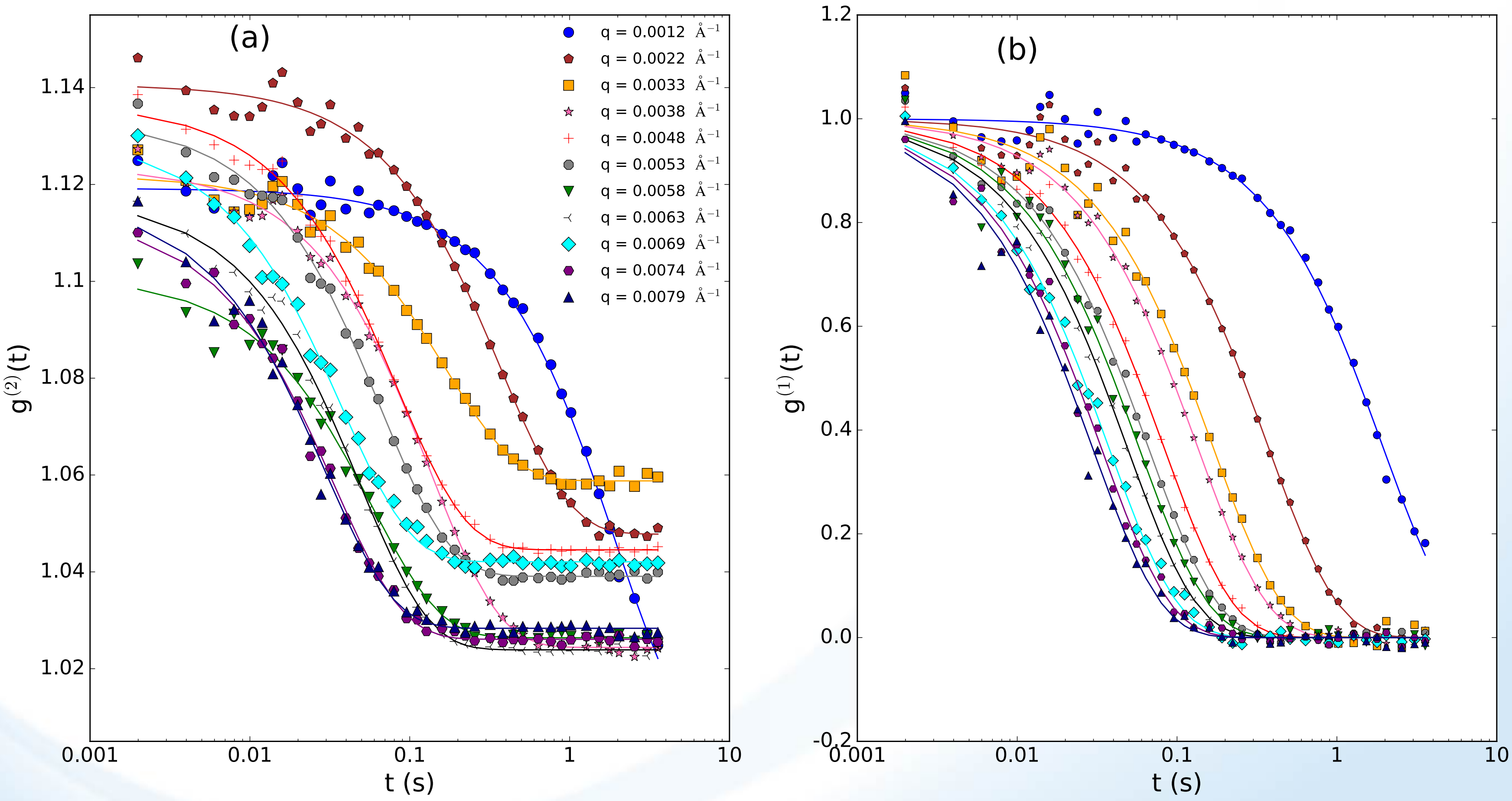
Ring "labels" (q rings)



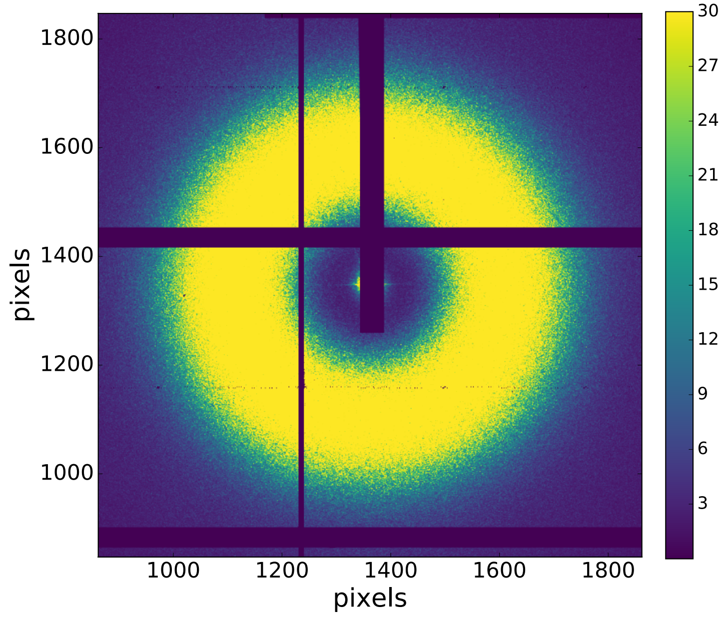
Q rings marked on the silica colloidal average image (Average of 10000 images with 2 ms integration time)

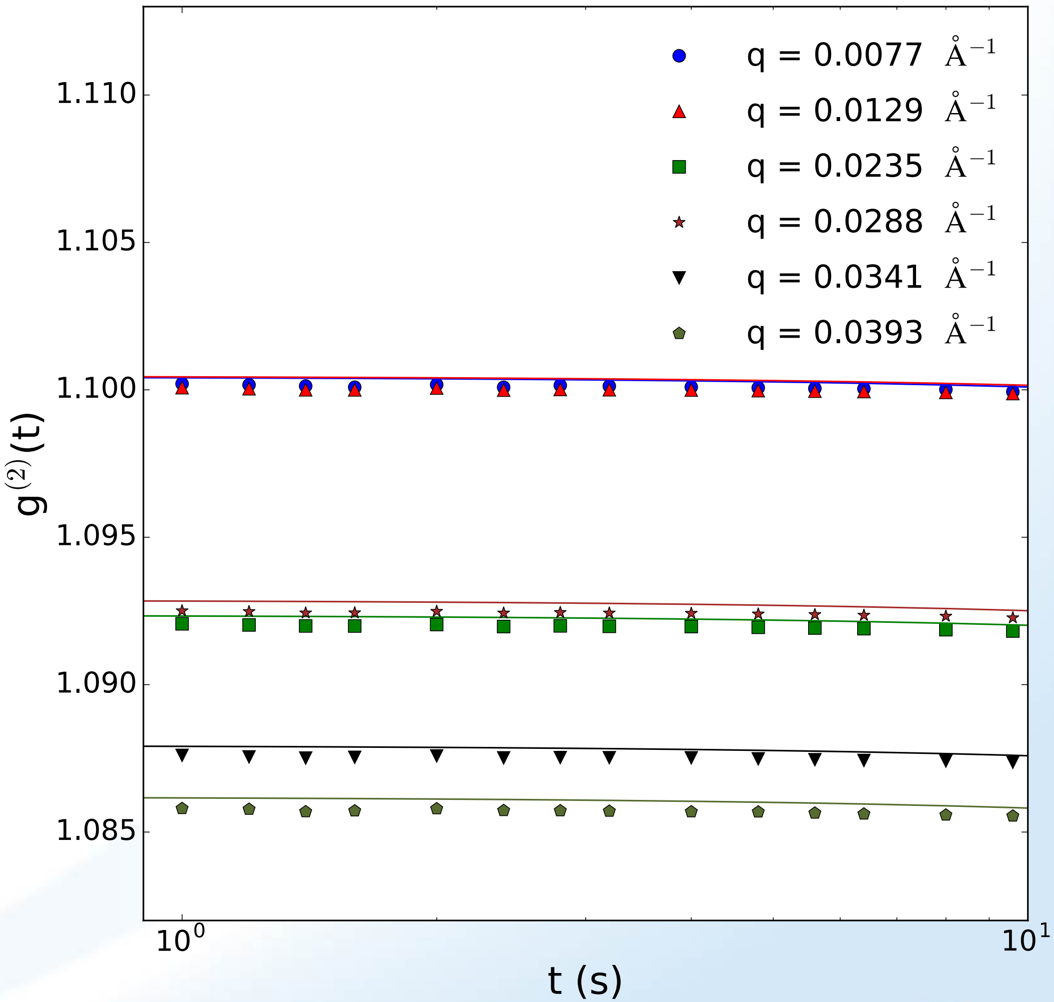
Scikit-beam tools allow the definition of customized ROI's of simple shapes, such as, rings, rectangle, circles etc..

One-time correlation plots from Sckit-beam for a dynamic sample



One time-correlation for a static sample





Any correlation time associated with possible experimental instabilities are large enough to make the correlation function constant on the timescales probed here.

Two-time correlation

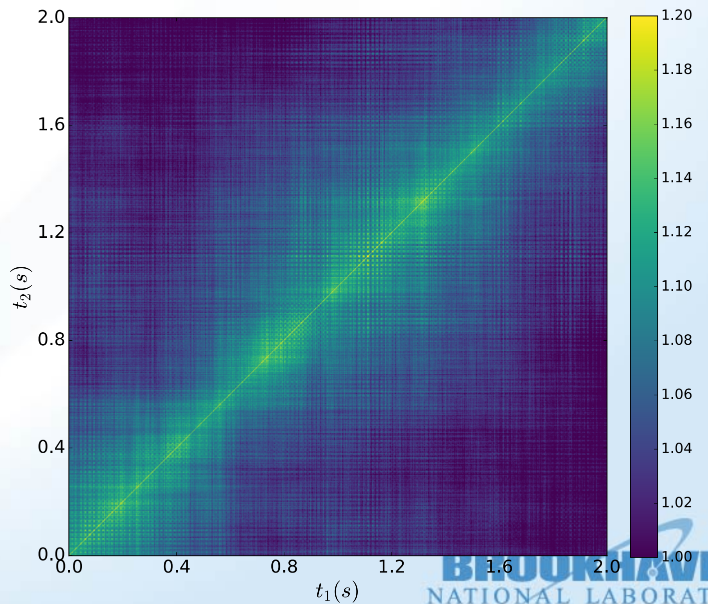
Non-equilibrium systems

$$C(q, t_1, t_2) = \frac{\langle I(q, t_1)\, I(q, t_2) \rangle}{\langle I(q, t_1)\rangle \langle I(q, t_2)\rangle}$$

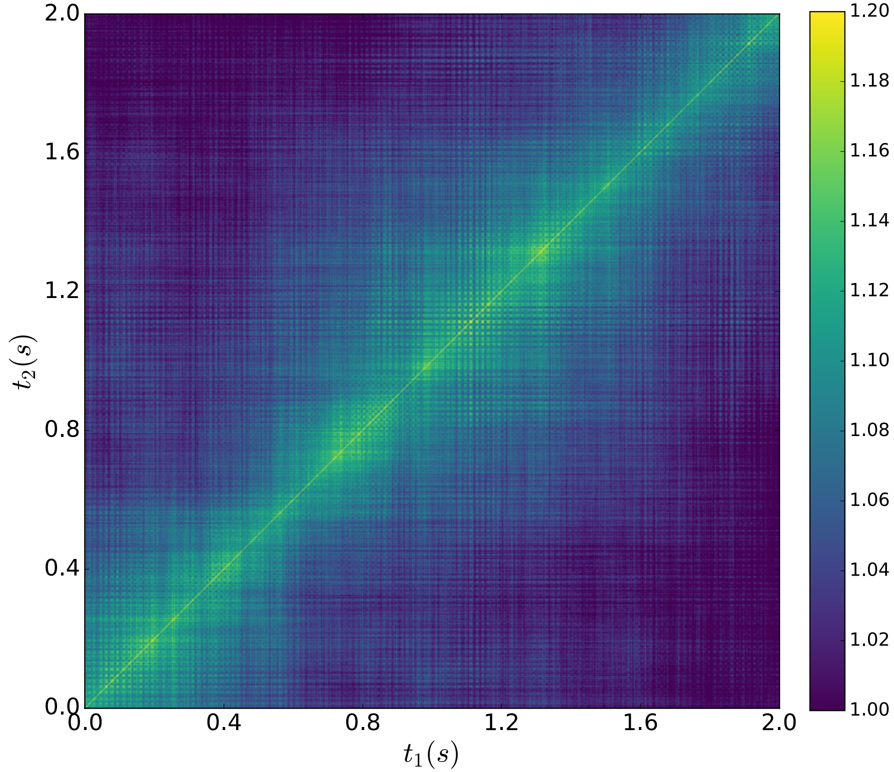
<…> - ensamble average over equivalent pixels of the detector.

Average time (age)
$t_a = (t_1 + t_2)/2$

Represents the time along the
$t_1 = t_2$ diagonal.



Two-time correlation

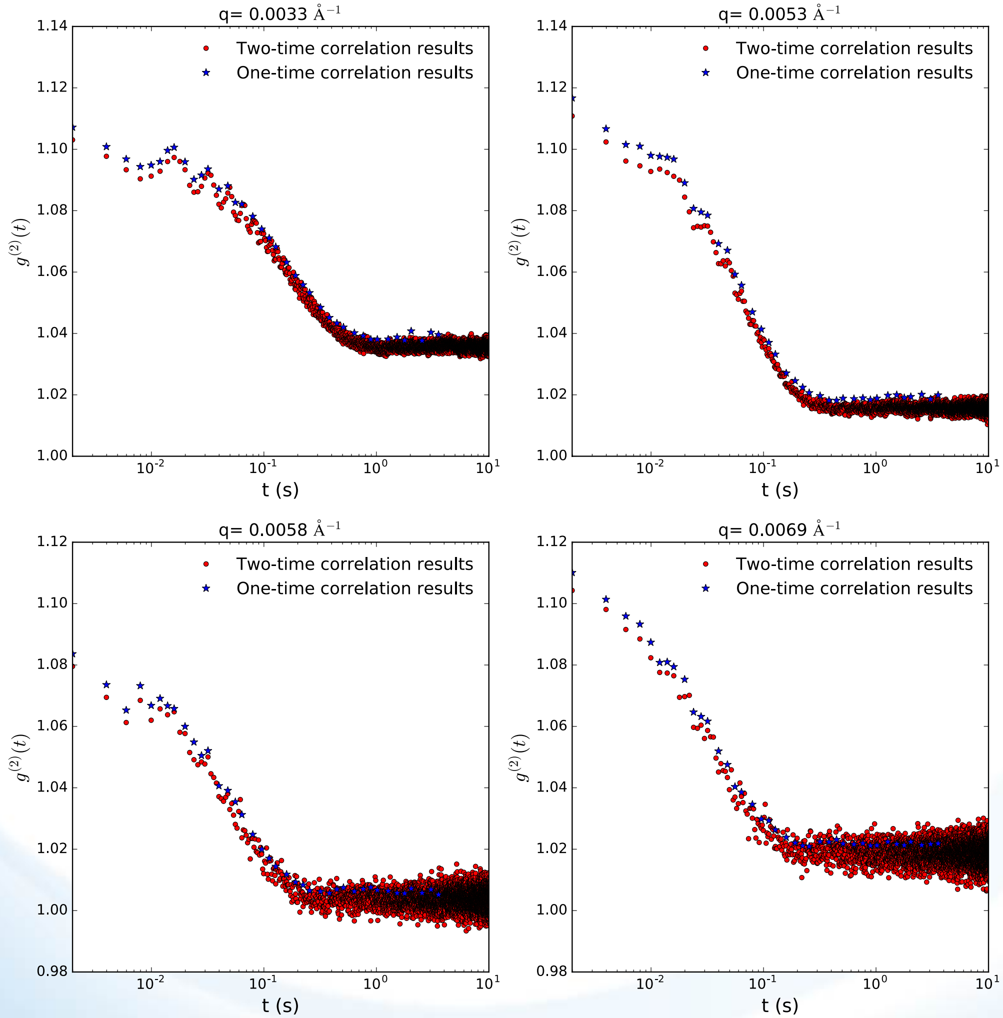


$t = | t_1 – t_2 |$ measures the "distance" from $t_1 = t_2$ diagonal

When the system is in equilibrium, the two-time correlation functions depend only on the time difference t, and not on the sample "age" t_a.

In this case, the one-time correlation function $g^{(2)}(q, t)$ is simply an average of $C(q, t_1, t_2)$ over all ages t_a.

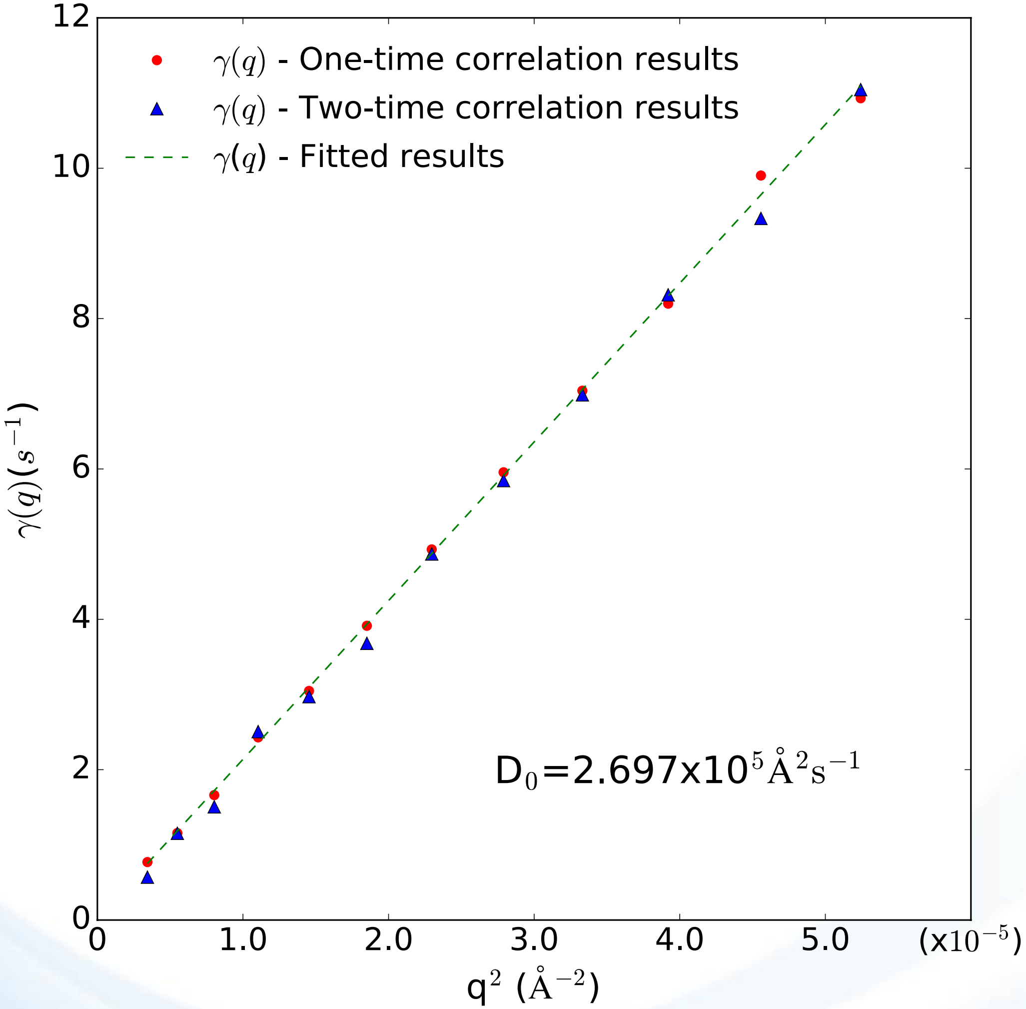
One-time correlations from two-time correlations



Two-time correlation of a sample from an equilibrium fluid. The correlation times are age-independent.

Therefore, two-time correlations can be used to calculate one-time correlations, by simply averaging all the "slices" corresponding to different sample ages.

Diffusion co-efficient



The relaxation rate *γ(q)* from a colloidal sample obtained by the one-time correlation and two-time correlation analysis plotted against q^2.

The experimental values are fitted by theoretical simulation.

Diffusion coefficient D_0 is calculated by characterizing the (random) motion of the Brownian colloidal particles in solvent.

BROOKHAVEN NATIONAL LABORATORY

Streaming data analysis pipelines

In a typical XPCS experiment, 20,000 speckle patterns are recorded at the maximum rate of the Eiger 4M detector with an acquisition time of 1.34ms/per pattern.

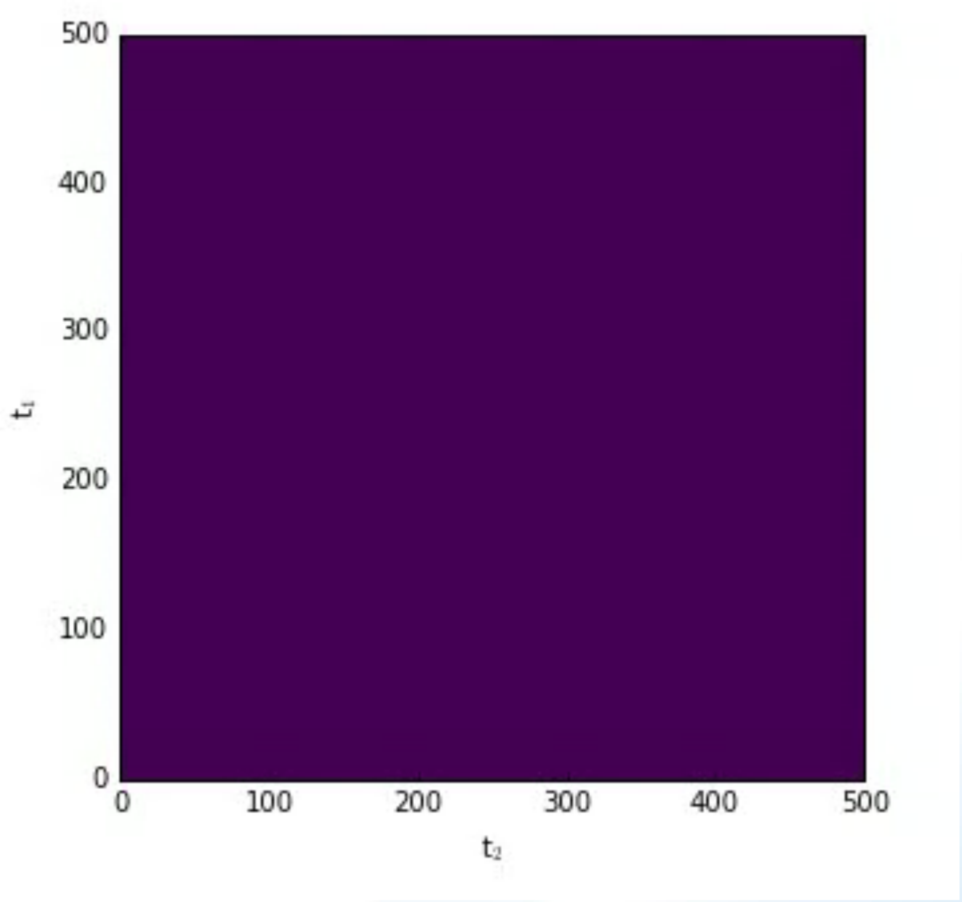
User beamtimes are short and awarded in competitive basis.

Therefore, processing, analysis and interpretation of a continues stream of data is required in an efficient manner.

Real time data analysis will help users to make decisions regarding the data collection strategies during experiments to optimize the scientific outcome.

Streaming tools for one-time correlation

Streaming tools for two-time correlation



Summary

New beamlines at NSLS-II generate large volumes of experimental data.

Therefore, it is required to have efficient data analysis tools to process, analyze and interpret continuous streams of structured and unstructured data.

A set of new streaming data analysis software tools have been developed for the studies of structural dynamics of materials using coherent scattering and photon correlation techniques.

These tools support wide range of user expertise, from novice to developer, and available in Scikit-beam python package, which is available at https://github.com/scikit-beam/scikit-beam.

Currently the Coherent Soft X-ray -1 (CSX-1) beamline uses these new data analysis tools.

Acknowledgment

Funding Sources: BNL, Laboratory Directed Research and Development (LDRD) program and U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-SC0012704.

Li Li, Kenneth Lauer, Nikolay Malitsky, Daron Chabot and Yong-Nian Tang (NSLS-II, DAMA group)

Claudio Mazzoli, Andi Barber (BNL, NSLS-II, CSX-1 Beamline)

Wei Xu, Michal McGuigan, Robert Harrison and Dantong Yu and all BNL, CSI staff

John Hill (BNL, NSLS-II Director)

Paul Zschack (BNL, Photon Science Division Director)

BNL, NSLS-II scientific staff

BNL, NSLS-II controls group