

**Title:** Observation of correlated x-ray scattering at atomic resolution

**Authors:** Derek Mendez, Thomas J. Lane, Jongmin Sung, Clément Levard, Herschel Watkins, Aina Cohen, Michael Soltis, Shirley Chisholm, James Spudich, Vijay Pande, Daniel Ratner, Sebastian Doniach

## Abstract

Tools to study structure of disordered systems, such as proteins in solution, remain limited. Such understanding is essential for *e.g.* rational drug design. Correlated x-ray scattering (CXS) has recently attracted new interest as a way to leverage next-generation light sources to study such disordered matter. The CXS experiment measures the intensity correlation caused by the scattering of x-rays from an ensemble of identical particles, with disordered orientation and position. Averaging over 15,496 snapshot images obtained by exposing a sample of silver nanoparticles in solution to a micro-focused synchrotron beam, we report an experimental CXS signal obtained from an ensemble in three dimensions. The correlation function was measured at wide angles corresponding to atomic resolution and matches theoretical predictions. These results suggest that other CXS experiments on disordered ensembles – such as proteins in solution – may be possible in the future.

## 1 Introduction

In a pioneering paper, Kam [4] showed that correlated x-ray scattering (CXS) from an ensemble of randomly oriented particles could in principle reveal information about the internal structure of the particles beyond usual small and wide angle solution scattering measurements. The extraction of such information in the absence of an ordered system (*e.g.* a crystal) can be beneficial in biological studies, as many biological systems are inherently disordered (*e.g.* proteins in solution).

In order to gauge the feasibility of Kam’s method at atomic length scales and to assess the associated difficulties, we conducted experiments measuring CXS from silver nanoparticle (NP) solutions at wide angles. Crucially, each measurement was conducted on an ensemble of NPs oriented randomly in three dimensions, extending previous experimental work done in two dimensions [13], at small angles in 3-dimensions [6, 18], and on single particles in 3-dimensions [5, 17].

From these experiments, we obtained empirical correlation functions measuring the correlation between all pixel pairs in the first two silver Bragg rings. The three correlation functions (two rings with themselves, one between rings) show sharp peaks consistent with analytical and simulated predictions based on the crystal structure of silver.

By successfully measuring CXS signal from an ensemble of silver NPs, we have demonstrated the effectiveness of Kam’s method given the current advances in x-ray technology. This experiment will serve as a benchmark for future experiments involving much weaker scatterers, such as proteins. It is our hope that the refinement of our analysis techniques on a well known sample such as silver NPs will help facilitate the extension of CXS to studies of biomolecules in solution.

## 2 Theory

We briefly review the portions of [4] relevant to this manuscript. Let  $S(\vec{q}, \omega)$  represent the structure factor of an isolated particle in solution, i.e.

$$S(\vec{q}, \omega) = \left| \sum_j^{N_j} f_j(|\vec{q}|) e^{i\vec{q} \cdot (\hat{R}_\omega \cdot \vec{r}_j)} \right|^2 \quad (1)$$

where  $\vec{q}$  is the scattering momentum transfer vector,  $\vec{r}_j$  and  $f_j(|\vec{q}|)$  are the coordinates and form factor of the  $j^{th}$  atom in the particle respectively ( $1 \leq j \leq N_j$ ),  $\omega$  is a triple of Euler angles, and  $\hat{R}_\omega$  is a 3-dimensional rotation operator.

Kam showed that the correlation function

$$C(\vec{q}_1, \vec{q}_2) = \frac{N}{8\pi^2} \int S(\vec{q}_1, \omega) S(\vec{q}_2, \omega) d\omega \equiv C(|\vec{q}_1|, |\vec{q}_2|, \psi) \quad (2)$$

( with  $\psi$  the angle between vectors  $\vec{q}_1$  and  $\vec{q}_2$  ), may be extracted from a CXS measurement where one repeatedly records snapshots of a solution of  $N$  identical particles, with each snapshot representing a unique ensemble of particles frozen in solution. Kam argued (assuming negligible inter-particle scattering interference) that the empirical correlation function averaged over shots would converge to (2), *i.e.*,

$$\langle n_s(\vec{q}_1) n_s(\vec{q}_2) \rangle_s - \langle n_s(\vec{q}_1) \rangle_s \langle n_s(\vec{q}_2) \rangle_s \Rightarrow C(\vec{q}_1, \vec{q}_2) \quad (3)$$

where  $n_s(\vec{q})$  is the total photons scattered from all particles in snapshot  $s$  ( $1 \leq s \leq N_s$ ) into a pixel along scattering vector  $\vec{q}$ . Neglecting inter-particle scattering interference,  $n_s(\vec{q})$  can be thought of as a linear combination of  $S(\vec{q}, \omega)$  for a set of orientations  $\{\omega\}_s$ .

While one can show that inverting a complete set of correlations to an image of the electron density is an underdetermined problem [2], we would like to emphasize that (2) can be calculated using a model of  $\rho$ . This opens up the possibility for refining such a model against a CXS dataset using (3) [10, 1, 15], especially in cases where prior information (such as protein primary sequence) can be included. Other routes to analyzing structure given CXS data include iterative phasing [14], implicit alignment of particles [12], or direct analysis of local internal symmetries of the system [9, 8].

We have computed the expected correlation function (2) for the silver NPs studied as a benchmark for our experimental results (Figure 1-A). A silver NP may be represented by a simple model consisting of a face-centered-cubic lattice cutoff by a spherical boundary. The scattering can then be represented by reciprocal lattice vectors cutting the Ewald sphere and giving rise to Bragg peaks broadened owing to the finite size of the NPs. Hence, each snapshot records a series of Bragg rings. The sub-population of all NP orientations that simultaneously subtend two Bragg peaks on the Ewald sphere give rise to the correlations in (3). In what follows, we will only consider scattering vectors which meet the Bragg

condition for silver. Let  $hkl$  denote a set of miller indices [cite] representing a silver Bragg condition. Let  $\phi$  be the azimuthal coordinate of  $\vec{q}_{hkl}$  projected into the plane perpendicular to an incident beam (corresponding to the azimuthal coordinate of a pixel measuring  $\vec{q}_{hkl}$  on a farfield planar detector). Let  $\Delta = |\phi_{hkl} - \phi_{h'k'l'}|$  be the angle between two such vector projections. Then define the average angular correlation of Bragg ring intensity fluctuations as

$$D(|\vec{q}_{hkl}|, |\vec{q}_{h'k'l'}|, \Delta) = \left\langle \int_0^{2\pi} \delta n_s(|\vec{q}_{hkl}|, \phi) \delta n_s(|\vec{q}_{h'k'l'}|, \phi + \Delta) d\phi \right\rangle_s \quad (4)$$

where

$$\delta n_s(|\vec{q}_{hkl}|, \phi) = n_s(|\vec{q}_{hkl}|, \phi) - \mu_s(|\vec{q}_{hkl}|)$$

and  $\mu_s(|\vec{q}_{hkl}|)$  is the average photon counts per Bragg ring per snapshot. Assuming proper sampling of orientation space per snapshot, and sufficiently large numbers of snapshots, it can be argued that  $D(|\vec{q}_{hkl}|, |\vec{q}_{h'k'l'}|, \Delta)$  and the Kam correlation  $C(|\vec{q}_{hkl}|, |\vec{q}_{h'k'l'}|, \psi)$  contain the same information content. Such discussions, however, are beyond the scope of this report. One should expect CXS signal at values of  $\Delta$  corresponding to the geometry of the reciprocal lattice and the angles between reciprocal lattice vectors. For instance,  $D(|\vec{q}_{111}|, |\vec{q}_{111}|, \Delta)$  should display strong signal at  $\Delta_1 = \arccos[\frac{-2}{3\cos^2\theta} + 1]$  and  $\Delta_2 = \arccos[\frac{-4}{3\cos^2\theta} + 1]$  where  $2\theta$  is the standard scattering angle.

Theoretical results suggest the CXS signal to noise scales as  $\sqrt{N_s}$ , but is independent of the number of illuminated particles  $N$  for sufficiently large  $N$  [4, 6, 7]. These facts were employed to optimize our experimental design, which emphasized collecting a large number of independent images.

### 3 Methods

Data from 15,496 x-ray diffraction images of silver NP solution were collected and analyzed at the micro-crystallography beamline (12-2) at SSRL. We prepared a sample containing an estimated  $10^9$  20 nm NPs per snapshot in the illuminated volume (350 mg/mL conc.) [CITE: NP PREP], but we observed significant numbers of NPs that were larger. Samples were loaded and oriented in the X-ray beam using the Stanford Automated Mounting System (SAM), controllable from the experimental hutch. Using a liquid nitrogen-cooled double crystal monochromator we tuned the beam energy to 17 keV. The beam was focused down to about  $20 \times 50 \mu\text{m}^2$  using Rh coated Kirkpatrick-Baez mirrors. Snapshots were recorded on a Dectris Pilatus 6M photon counting detector.

To successfully measure a correlation function via the scheme (3), the sample must be frozen in time or space. Any random motion due to diffusion of particles will reduce the scattering correlation, which is a function of the particle structure and orientation (eq. 1). To prevent diffusion during the long exposure times (order 1 second) necessary to scatter a sufficient number of photons to measure a correlation signal at a synchrotron, we cooled

the sample 100 Kelvin using a nitrogen cryo jet, ensuring that the particles remained immobilized during each exposure. The NPs were held in glycerol antifreeze used to prevent the formation of solvent crystals at the low temperature, forming a colloidal suspension. To house the solutions we used kapton capillaries with a 500 and 600  $\mu\text{m}$  inner and outer diameter respectively. Kapton scatters into relatively lower angles as does glycerol, and we did not observe corruption of our silver NP signal with scattering from the kapton or glycerol. Samples were prepared a day early and stored in a liquid nitrogen bath. Further, no changes in the diffraction images were observed during the course of a single snapshot exposure, indicating the sample does not undergo significant diffusion or radiation damage.

Our goal was to record as many snapshots as possible, each one representing a different ensemble of particle orientations frozen in time. The sample holder was equipped to automatically rotate the capillary 150 degrees about its longitudinal axis, perpendicular to the beam. Photon counts were read out and reset every 0.7 second as the capillary rotated 0.3 degrees under continuous beam irradiation, yielding 500 shots per 150 degree rotational scan. This was deemed an optimal timing to simultaneously maximize signal and minimize damage and heating. Every 500 shots, between scans, the capillary was moved longitudinally so as to always probe different regions of the sample.

A bicubic interpolation algorithm was used to convert the cartesian pixel lattice to polar coordinates for calculation of equations like (4). Using the Scherrer equation [8] relating the width of a Bragg ring to the average NP size, we concluded that the majority of silver NPs in each snapshot were roughly 20 nm in diameter. Histograms of photon counts into the  $|\vec{q}_{111}|$  Bragg ring indicate a rather large distribution of particle sizes which we discuss in the following section.

## 4 Results

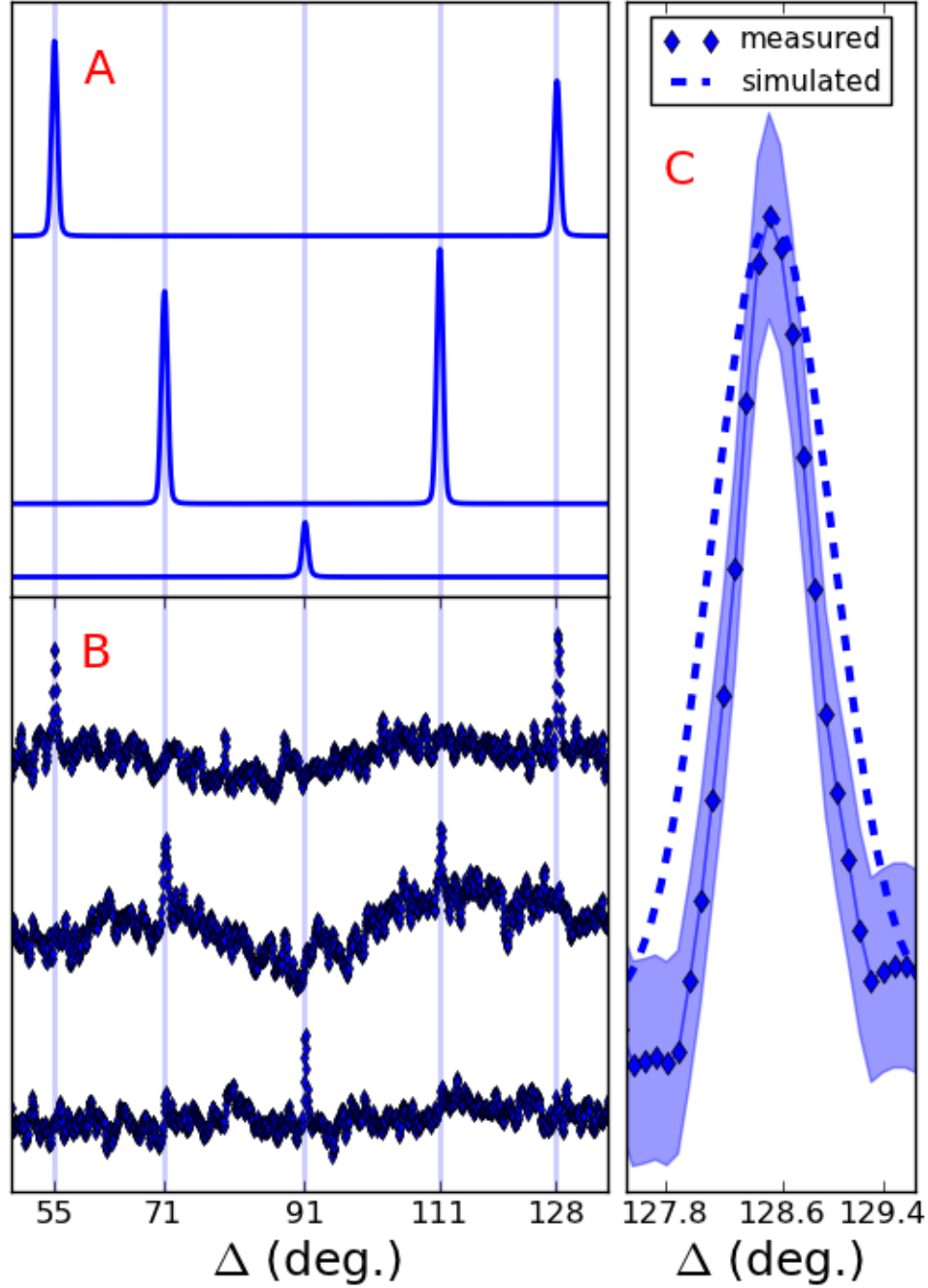
Here we report on our attempts to resolve (4) at the two brightest silver Bragg rings. Specifically, we calculated  $D(|\vec{q}_{111}|, |\vec{q}_{111}|, \Delta)$ ,  $D(|\vec{q}_{200}|, |\vec{q}_{200}|, \Delta)$ , and  $D(|\vec{q}_{111}|, |\vec{q}_{200}|, \Delta)$ .

Initial attempts to see convergence of the CXS signal were unsuccessful. In addition to the inherent statistical noises described in [1,7], CXS measurement is very sensitive to systematic noises associated with the experimental conditions [6], *e.g.* detector artifacts. The Pilatus 6M detector is made up of 60 pixel modules separated by 1-2 mm gaps, and each Bragg ring subtends multiple modules. The overall electronic response of each module is slightly different, causing false intensity correlations that dominate any underlying CXS.

In order to minimize these parasitic correlations, we applied a binary filter to the data. We defined the filter as follows: Let  $\{\vec{q}_{hkl}\}_m$  represent the set of Bragg pixels on the  $m^{\text{th}}$  detector module ( $1 \leq m \leq 60$ ). Let  $\mu_m$  and  $\sigma_m$  represent the average and standard deviation of  $n_s(\vec{q})$  for all  $\vec{q} \in \{\vec{q}_{hkl}\}_m$ , respectively. Then for each  $\vec{q} \in \{\vec{q}_{hkl}\}_m$ , we apply the photon count filter:

$$n_s(\vec{q}) = \begin{cases} 1 & \text{if } n_s(\vec{q}) > \mu_m + \sigma_m \\ 0 & \text{otherwise} \end{cases}$$

*i.e.* all pixels less than one standard deviation above the mean are set to zero and the rest



**Figure 1** **A** From top to bottom, 20 nm silver NP simulations of  $D(|\vec{q}_{111}|, |\vec{q}_{200}|, \Delta)$ ,  $D(|\vec{q}_{111}|, |\vec{q}_{111}|, \Delta)$ , and  $D(|\vec{q}_{200}|, |\vec{q}_{200}|, \Delta)$  **B** Corresponding measurements of the correlations plotted in **A**. Vertical lines mark analytical predictions. We truncated the signal to isolate the correlation peaks. The entire angular plot is symmetric about  $\Delta = \pi$ . Regions not shown contain artifacts similar to those on the figure, yet nothing greater in magnitude than the CXS peaks. **C** Width of a correlation peak. The simulation was for 20 nm particles. Peak width scales inversely with particle size, hence we expect the measured CXS resulted from particles larger than 20 nm. Shading represents 95% confidence intervals on the measurement.

are set to unity. Hence, we only looked for correlations from the strongest Bragg peaks in each shot. Furthermore, the filter placed each detector module on the same scale, removing correlations between modules. [ will explain why 1 sigma was chosen here, after running some tests – TJJ sounds good. you don’t need to be toooo precise, the fact it worked is good enough] After applying the filter and calculating (4), the resulting correlations averaged to display peaks corresponding to the double Bragg scattering discussed above (Figure 2). Despite the filtering, some artificial correlations persisted. We have discovered that the response of pixels within a module can also vary spatially, biasing the binary filter selection within the module and leading to artificial correlations. The low-order oscillations seen in Figure 2-B are most likely generated from such artifacts, as intra-module pixel response variation is relatively weak. Refinement of the data analysis is an ongoing process.

As mentioned above, we infer from the width of the Bragg rings that the majority of our sample consisted of  $10^9$  20 nm silver NPs. From tabulated coherent atomic cross sections [3] we estimated that each 20nm silver NP scattered roughly 1.6 photons per 0.7 sec exposure. We can further estimate that roughly 8.3% of these NPs scatter into the Bragg ring at  $q_{111}$  on the detector. (This is done by tracing out volumes of rotation of the reciprocal lattice points whose diameter is given by the Scherrer equation [8], and then determining which fraction of that volume intersects the Ewald sphere). These estimates are in line with the average photon counts per pixel per snapshot that we observed in the  $q_{111}$  Bragg ring,  $\approx 4 \cdot 10^4$ . Many of the snapshots had Bragg peaks with counts  $\approx 8 \cdot 10^8$ , much greater than one would expect given Poisson photon statistics. Thus we conclude that our sample consisted of larger particles which scatter more photons. To put a lower bound on the size of these particles we again turn to the Scherrer equation, this time relating the width of the correlation peak to the particle size (Figure 2-C). We simulated (4) for 20 nm NPs and compared the width of a simulated correlation peak to the width of the corresponding measured one. The width of the correlation peak is in line with estimates for 50 nm particles. Thus, we concluded that the correlations we are measuring are most likely generated by scattering off of thousands of  $\geq 50$  nm particles. We would like to emphasize that this is an estimated lower bound, for example thermal vibrations might cause a Bragg peak of a very large particle to broaden, whereas so far we have assumed these kinds of effects are negligible. We estimated that roughly 0.5% [ rough estimate, I’ll put a number on it tomorrow] of 50 nm silver NP orientations will simultaneously subtend 2 Bragg peaks on the detector at either  $|\vec{q}_{111}|$  or  $|\vec{q}_{200}|$ .

With the result of our measurement, we show that the theoretical Poisson noise is not limiting the convergence of (3) in our simple case. We conclude that the main impediment to accurate measurements of CXS arises from systemic errors resulting from anisotropy artifacts induced by the detector system, which we have been able to partially overcome by the nonlinear filtering to a binary signal.

## 5 Discussion

Kam’s original results [4] show that (2), which may be directly calculated from the structure factors of the particle (1), may be derived from measurement of an ensemble of particles as a whole. Hence accurate measurement of CXS in the 3-dimensional  $\{|\vec{q}_1|, |\vec{q}_2|, \Delta\}$  space can lead to constraints which can be placed on the electron density in a model, thus giving a route to iterative refinement.

Our results show that it is possible to obtain atomic scale information on the internal structure of a particle, from a bulk sample containing many identical but randomly oriented particles, for the simple case of a silver NP. These results suggest that it should be feasible to obtain more detailed atomic scale constraints on models of more complex biomolecules by measuring scattering using pulses from x-ray free electron lasers [11, 16], provided one can effectively correct for experimentally induced sources of noise, particularly detector-induced correlations.

## References

- [1] Gang Chen, Peter H Zwart, and Dongsheng Li. Component Particle Structure in Heterogeneous Disordered Ensembles Extracted from High-Throughput Fluctuation X-Ray Scattering. *Phys. Rev. Lett.*, 110(19):195501, May 2013.
- [2] Veit Elser. Strategies for processing diffraction data from randomly oriented particles. *Ultramicroscopy*, 111(7):788–792, June 2011.
- [3] B L Henke, E M Gullikson, and J C Davis. X-Ray Interactions: Photoabsorption, Scattering, Transmission, and Reflection at  $E= 50\text{-}30,000$  eV,  $Z= 1\text{-}92$ . *Atomic data and nuclear data tables*, 1993.
- [4] Z Kam. Determination of macromolecular structure in solution by spatial correlation of scattering fluctuations. *Macromolecules*, 10(5):927–934, 1977.
- [5] Z Kam and I Gafni. Three-dimensional reconstruction of the shape of human wart virus using spatial correlations. *Ultramicroscopy*, 17(3):251–262, 1985.
- [6] Z Kam, M H Koch, and J Bordas. Fluctuation x-ray scattering from biological particles in frozen solution by using synchrotron radiation. *Proceedings of the National Academy of Sciences*, 78(6):3559–3562, June 1981.
- [7] Richard A Kirian, Kevin E Schmidt, Xiaoyu Wang, R Bruce Doak, and John C H Spence. Signal, noise, and resolution in correlated fluctuations from snapshot small-angle x-ray scattering. *Phys. Rev. E*, 84(1):011921, July 2011.
- [8] R P Kurta, M Altarelli, and I A Vartanyants. X-ray cross-correlation analysis of disordered systems: potentials and limitations. *arXiv*, February 2013.
- [9] R P Kurta, M Altarelli, E Weckert, and I A Vartanyants. X-ray cross-correlation analysis applied to disordered two-dimensional systems. *arXiv*, February 2012.
- [10] Haiguang Liu, Billy K Poon, Dilano K Saldin, John C H Spence, and Peter H Zwart. Three-dimensional single-particle imaging using angular correlations from X-ray laser data. *Acta Crystallogr A Found Crystallogr*, 69(4):365–373, May 2013.
- [11] R Neutze, R Wouts, D van der Spoel, E Weckert, and J Hajdu. Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature*, 406(6797):752–757, August 2000.
- [12] H C Poon, P Schwander, M Uddin, and D K Saldin. Fiber Diffraction without Fibers. *Phys. Rev. Lett.*, 110(26):265505, June 2013.

- [13] D Saldin, H Poon, M Bogan, S Marchesini, D Shapiro, R Kirian, U Weierstall, and J Spence. New Light on Disordered Ensembles: Ab Initio Structure Determination of One Particle from Scattering Fluctuations of Many Copies. *Phys. Rev. Lett.*, 106(11):115501, March 2011.
- [14] D K Saldin, H C Poon, V L Shneerson, M Howells, H N Chapman, R A Kirian, K E Schmidt, and J C H Spence. Beyond small-angle x-ray scattering: Exploiting angular correlations. *Phys. Rev. B*, 81(17):174105, May 2010.
- [15] D K Saldin, V L Shneerson, R Fung, and A Ourmazd. Structure of isolated biomolecules obtained from ultrashort x-ray pulses: exploiting the symmetry of random orientations. *J. Phys.: Condens. Matter*, 21(13):134014, March 2009.
- [16] J C H Spence, U Weierstall, and H N Chapman. X-ray lasers for structural and dynamic biology. *Rep. Prog. Phys.*, 75(10):102601, September 2012.
- [17] D Starodub, A Aquila, S Bajt, M Barthelmess, A Barty, C Bostedt, J D Bozek, N Coppola, R B Doak, S W Epp, B Erk, L Foucar, L Gumprecht, C Y Hampton, A Hartmann, R Hartmann, P Holl, S Kassemeyer, N Kimmel, H Laksmono, M Liang, N D Loh, L Lomb, A V Martin, K Nass, C Reich, D Rolles, B Rudek, A Rudenko, J Schulz, R L Shoeman, R G Sierra, H Soltau, J Steinbrener, F Stellato, S Stern, G Weidenspointner, M Frank, J Ullrich, L Str uuml der, I Schlichting, H N Chapman, J C H Spence, and M J Bogan. Single-particle structure determination by correlations of snapshot X-ray diffraction patterns. *Nature Communications*, 3:1276–7, 1.
- [18] Peter Wochner, Christian Gutt, Tina Autenrieth, Thomas Demmer, Volodymyr Bugaev, Alejandro Díaz Ortiz, Agnès Duri, Federico Zontone, Gerhard Grübel, and Helmut Dosch. X-ray cross correlation analysis uncovers hidden local symmetries in disordered matter. *P Natl Acad Sci Usa*, 106(28):11511–11514, July 2009.