Hindawi Publishing Corporation
Advances in Condensed Matter Physics
Volume 2013, Article ID 959835, 15 pages
http://dx.doi.org/10.1155/2013/959835

Research Article

X-Ray Cross-Correlation Analysis of Disordered Ensembles of Particles: Potentials and Limitations

R. P. Kurta,[1] M. Altarelli,[2] and I. A. Vartanyants[1,3]

[1] *Deutsches Elektronen-Synchrotron DESY, Notkestraße 85, 22607 Hamburg, Germany*
[2] *European X-Ray Free-Electron Laser Facility, Albert-Einstein-Ring 19, 22761 Hamburg, Germany*
[3] *National Research Nuclear University, "MEPhI", Moscow 115409, Russia*

Correspondence should be addressed to I. A. Vartanyants; ivan.vartaniants@desy.de

Received 8 February 2013; Accepted 18 April 2013

Academic Editor: Dilano K. Saldin

Angular X-ray cross-correlation analysis (XCCA) is an approach to study the structure of disordered systems using the results of X-ray scattering experiments. In this paper we summarize recent theoretical developments related to the Fourier analysis of the cross-correlation functions. Results of our simulations demonstrate the application of XCCA to two- and three-dimensional (2D and 3D) disordered ensembles of particles. We show that the structure of a single particle can be recovered using X-ray data collected from a 2D disordered system of identical particles. We also demonstrate that valuable structural information about the local structure of 3D systems, inaccessible from a standard small-angle X-ray scattering experiment, can be resolved using XCCA.

1. Introduction

Correlation methods are widely used in the physics of disordered materials such as amorphous and glassy systems [1, 2]. Recently it was suggested [3–10] that angular cross-correlations of scattered intensities in coherent diffraction experiment could provide essential structural information about the system that is not accessible by other scattering techniques. The growing interest in the application of these methods is driven by two main factors. The first is its application to the physics of disordered or partially ordered systems, and the second is an attempt to solve the structure of biological molecules in their native environment. Although ideas of applying cross-correlation techniques to study disordered systems go back to the work of Kam over 30 years ago [11, 12], new perspectives are opened by the emerging high power free-electron lasers (FELs) [13–16], which provide intense coherent femtosecond pulses of X-rays.

In the study of disordered systems the application of cross-correlation methods is driven by the hope to unveil the hidden symmetries and dynamics in a disordered collection of elements composing the system. This leads to the problem of understanding systems with complicated correlation behavior between dynamical heterogeneity and medium-range order [17–19] in a large class of glass-forming liquids [20–22]. One of the mysterious liquid systems whose structure is not fully understood is water. It exists in different forms known as low-density and high-density amorphous states [23]. The phase diagram of different forms is still under debate [24, 25]. In such systems the relevant task could be, for example, the detection of an n-fold symmetry axis of an individual molecular species in a liquid composed of such molecules. An identification of bond angles or local structure in an amorphous system, from the study of angular correlations of the diffracted intensity, is another attractive topic of research. There are partly affirmative answers in the study of partially ordered quasi two-dimensional systems like liquid crystals [26] in which hexatic bond order can be detected from the angular intensity correlation. Angular correlations with pronounced periodic character that can be directly related to the local symmetry of the system were also observed in colloidal suspensions [3, 27, 28].

For biological systems it was proposed to use high power FEL pulses in so-called single particle coherent diffraction imaging experiments [29, 30]. In these experiments individual particles, for example, macromolecules, are injected in the FEL beam and their diffraction patterns are measured

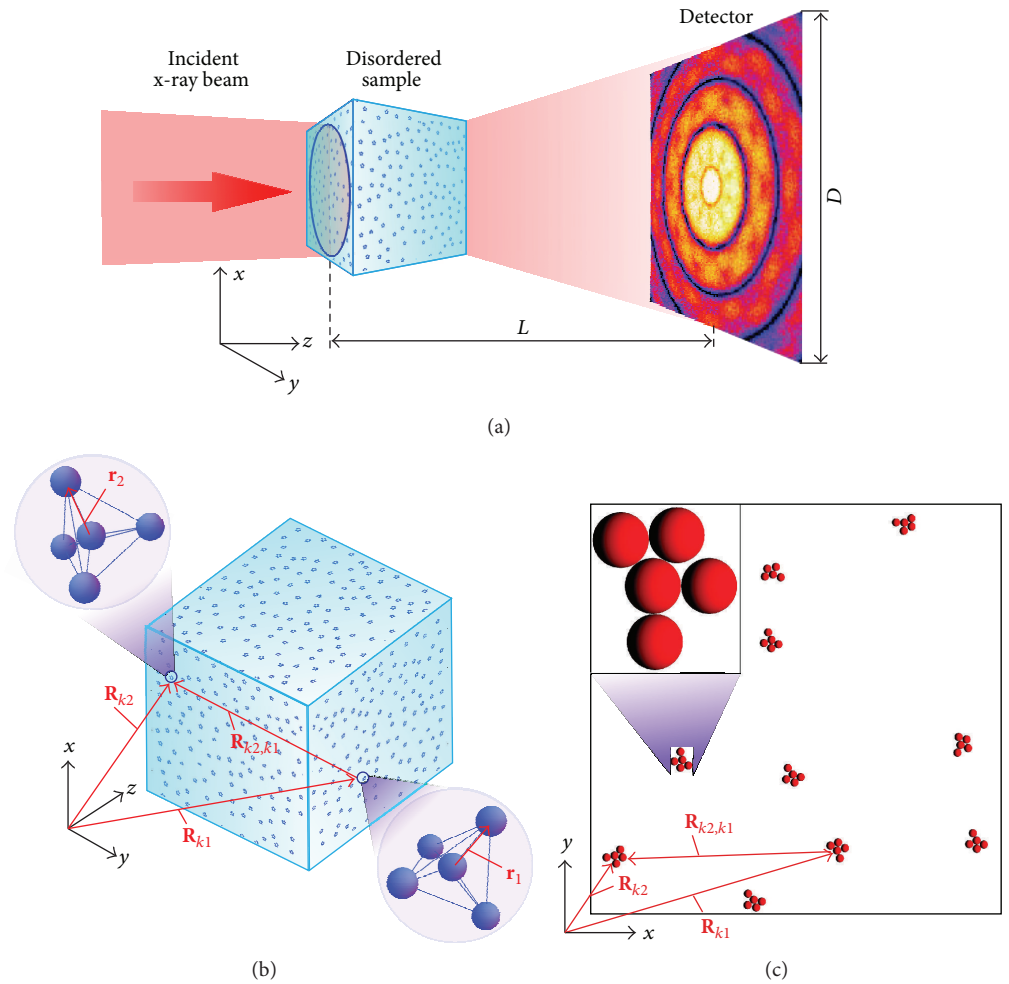

FIGURE 1: (a) Geometry of the diffraction experiment. A coherent X-ray beam illuminates a disordered sample and produces a diffraction pattern on a detector. The direction of the incident beam is defined along the z-axis of the coordinate system. (b) A disordered 3D sample composed of tetrahedral pentamers. All clusters have random position and orientation in the 3D space. (c) A disordered 2D sample composed of asymmetric clusters. All clusters have random position and orientation in the 2D plane.

in the far-field. After the first experiments performed at LCLS in Stanford [31], it was realized that even high power FELs provide single pulse resolution only up to tens of nanometers for gigantic viruses. Extension of the technique to sub-nanometer resolution for biological systems with the size below 100 nm is still under development. Clearly, it will be even more challenging to resolve molecules in their native liquid environment. One of the possible ways to enhance the scattered signal is the illumination of a large number of particles by a single FEL pulse. An approach based on cross-correlation analysis could provide a tool to determine three-dimensional (3D) diffraction pattern to high resolution with the available FEL sources [32–34].

Here, we discuss potentials of application of X-ray cross-correlation analysis (XCCA) to the study of a certain class of disordered systems. Our model sample consists of N identical non-interacting particles with orientational and positional disorder (see Figure 1). A coherent X-ray beam with a pulse duration shorter than the translational and rotational relaxation times of the system is scattered from the sample, and the diffraction pattern is measured in the far-field. The main concept of our approach based on application of XCCA to the study of disordered systems is depicted in Figure 2. A large number of diffraction patterns is measured by successive pulses of the FEL beam with different realizations of the system. We assume that the time between the pulses is sufficient that two sequential realizations of the system are completely uncorrelated. XCCA is performed for each diffraction pattern in a sequence, and angular cross-correlation functions (CCFs) are averaged over the whole data set. In this way structural information about an individual particle in the system can be obtained.

In this paper we summarize our recent theoretical results [4–7, 35] based on the analysis of the high-order cross-correlation functions. It is organized as follows: in Section 2 the definition of the two-point and three-point CCFs, as well as their angular Fourier decomposition, is given. In Section 3 intensities scattered from our model system are defined. In Section 4 a detailed description of XCCA for 2D dilute and dense systems is provided; this is followed in the next section

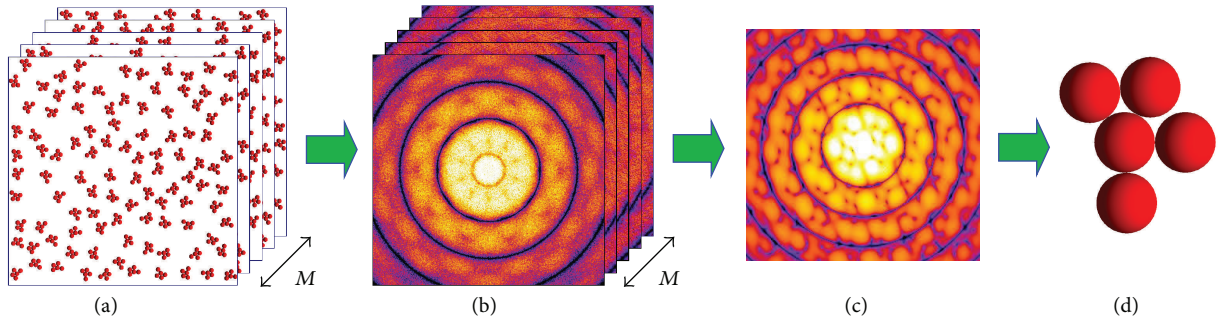

FIGURE 2: A concept of recovery of the structure of a single particle using X-ray scattering data from many particles. A large number M of realizations of a disordered system (a) composed of many identical particles is used to collect M diffraction patterns (b). X-ray cross-correlation analysis is applied to this X-ray data set to recover a diffraction pattern (c) corresponding to a single particle. The structure of a single particle (d) is determined applying phase retrieval algorithms to the recovered diffraction pattern (c).

by the description of a more general case of scattering from 3D systems. In Section 6 major results of the recovery of the structure of an individual 2D cluster and analysis of the scattering from a 3D system of pentameric structures are discussed. This is followed by the conclusions and an overview section.

2. Definition of the Two- and Three-Point CCFs and Their Angular Fourier Decomposition

The two-point CCF defined for a single realization of a disordered system at two resolution rings q_1 and q_2 is given by [4, 5, 11, 36] (see Figure 3(c))

$$C(q_1, q_2, \Delta) = \left\langle \tilde{I}(q_1, \varphi)\tilde{I}(q_2, \varphi + \Delta)\right\rangle_{\varphi}, \tag{1}$$

where $0 \le \Delta \le 2\pi$ is the angular coordinate, $\tilde{I}(q, \varphi) = I(q, \varphi) - \langle I(q, \varphi)\rangle_{\varphi}$, is the intensity fluctuation function, $\langle f(\varphi)\rangle_{\varphi} = (1/2\pi)\int_0^{2\pi} f(\varphi)d\varphi$ denotes the average over the angle φ.

In a similar way, the three-point CCF for a single realization of a system is defined at three resolution rings q_1, q_2, and q_3 as [7] (see Figure 3(d)),

$$C(q_1, q_2, q_3, \Delta_1, \Delta_2) = \left\langle \tilde{I}(q_1, \varphi)\, \tilde{I}(q_2, \varphi + \Delta_1)\, \tilde{I}(q_3, \varphi + \Delta_2)\right\rangle_{\varphi}, \tag{2}$$

where $0 \le \Delta_1 \le 2\pi$ and $0 \le \Delta_2 \le 2\pi$ are the angular coordinates.

In practical applications one would need to consider CCFs $\langle C(q_1, q_2, \Delta)\rangle_M$ and $\langle C(q_1, q_2, q_3, \Delta_1, \Delta_2)\rangle_M$ averaged over a sufficiently large number M of diffraction patterns [6],

$$\langle C(q_1, q_2, \Delta)\rangle_M = \frac{1}{M}\sum_{m=1}^{M}\{C(q_1, q_2, \Delta)\}^m, \tag{3a}$$

$$\langle C(q_1, q_2, q_3, \Delta_1, \Delta_2)\rangle_M = \frac{1}{M}\sum_{m=1}^{M}\{C(q_1, q_2, q_3, \Delta_1, \Delta_2)\}^m, \tag{3b}$$

where $\{C(q_1, q_2, \Delta)\}^m$ and $\{C(q_1, q_2, q_3, \Delta_1, \Delta_2)\}^m$ are the CCFs defined by (1) and (2) for the mth realization of a disordered system. The importance of this statistical averaging will be discussed in the following sections of the paper.

It is convenient to analyze the two-point CCF $C(q_1, q_2, \Delta)$ using a Fourier series decomposition in the $(0, 2\pi)$ interval [4, 5]

$$C(q_1, q_2, \Delta) = \sum_{n=-\infty}^{\infty} C^n_{q_1,q_2} e^{in\Delta}, \tag{4a}$$

$$C^n_{q_1,q_2} = \frac{1}{2\pi}\int_0^{2\pi} C(q_1, q_2, \Delta)\, e^{-in\Delta} d\Delta. \tag{4b}$$

Here, $C^n_{q_1,q_2}$ is the nth component in the Fourier series expansion of $C(q_1, q_2, \Delta)$, and $C^0_{q_1,q_2} = 0$ at $n = 0$ by definition (see (1)). Substituting (1) into (4b) and applying the Fourier convolution theorem we get

$$C^n_{q_1,q_2} = I^{n*}_{q_1} \cdot I^n_{q_2}. \tag{5}$$

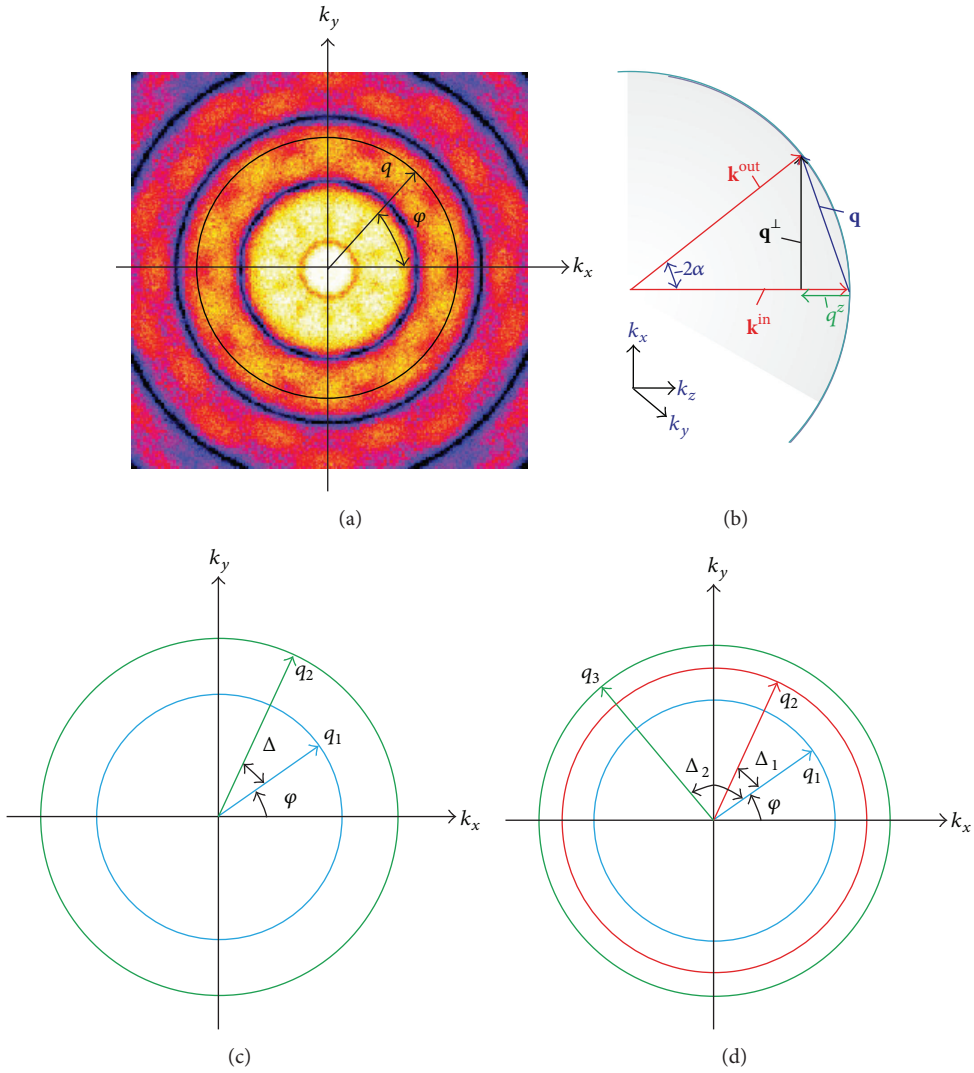

FIGURE 3: Scattering geometry in reciprocal space. (a) Scattered intensity $I(\mathbf{q})$ defined in the detector plane in the polar coordinate system, $\mathbf{q} = (q, \varphi)$. (b) Ewald sphere construction. Here $\mathbf{k}^{\text{in}}$ is the wavevector of the incident beam directed along the z axis, $\mathbf{k}^{\text{out}}$ is the wavevector of the scattered wave with the scattering angle 2α. The scattering vector $\mathbf{q} = (\mathbf{q}^{\perp}, q^z)$ is decomposed into two components that are perpendicular $\mathbf{q}^{\perp}$ and parallel q^z to the direction of the incident beam. Definition of the momentum transfer vectors in the derivation of the two-point $C(q_1, q_2, \Delta)$ (c), and three-point CCFs $C(q_1, q_2, q_3, \Delta_1, \Delta_2)$ (d).

Here I_q^n are the components of the Fourier expansion of the scattered intensity $I(q, \varphi)$ on the ring of radius q (see Figure 3(a)),

$$I(q, \varphi) = \sum_{n=-\infty}^{\infty} I_q^n e^{in\varphi}, \tag{6a}$$

$$I_q^n = \frac{1}{2\pi} \int_0^{2\pi} I(q, \varphi) e^{-in\varphi} d\varphi. \tag{6b}$$

Since scattered intensities are always real quantities, it is easy to show that $I_q^{-n} = I_q^{n^*}$ and $C_{q_1,q_2}^{-n} = C_{q_1,q_2}^{n^*}$. By its definition the 0th order Fourier component represents an angular averaged intensity, $I_q^0 = \langle I(q, \varphi) \rangle_{\varphi}$ [4–6].

In the specific case, when $q_1 = q_2 = q$, (4a) and (5) reduce to

$$C(q, \Delta) = 2 \sum_{n=1}^{\infty} C_q^n \cos(n\Delta), \tag{7a}$$

$$C_q^n = \left| I_q^n \right|^2, \quad C_q^n \geq 0. \tag{7b}$$

According to (7a) a strong single cosine dependence of $C(q, \Delta)$ can be observed for those values of q, at which one of the Fourier components C_q^n significantly dominates over all others [3]. Such components can be related to the structure and symmetry of the system [4–6].

The Fourier series expansion of the three-point CCF $C(q_1, q_2, q_3, \Delta_1, \Delta_2)$ can be written as

$$C(q_1, q_2, q_3, \Delta_1, \Delta_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} C_{q_1,q_2,q_3}^{n_1,n_2} e^{in_1\Delta_1} e^{in_2\Delta_2}, \tag{8a}$$

$$C_{q_1,q_2,q_3}^{n_1,n_2} = \left(\frac{1}{2\pi}\right)^2 \iint_0^{2\pi} C(q_1, q_2, q_3, \Delta_1, \Delta_2) \times e^{-in_1\Delta_1} e^{-in_2\Delta_2} d\Delta_1\, d\Delta_2, \tag{8b}$$

where $C_{q_1,q_2,q_3}^{n_1,n_2}$ are the Fourier components of the three-point CCF, and $C_{q_1,q_2,q_3}^{n_1,n_2} = 0$ for $n_1 = 0, n_2 = 0$ and $n_1 = -n_2$. Substituting (2) into (8b) one can get [7]

$$C_{q_1,q_2,q_3}^{n_1,n_2} = I_{q_1}^{(n_1+n_2)^*} I_{q_2}^{n_1} I_{q_3}^{n_2}. \tag{9}$$

In general, (9) determines a relation between three different Fourier components of intensity I_q^n of the order n_1, n_2 and $n_1 + n_2$, defined on three resolution rings, q_1, q_2, and q_3.

In practical applications one would need to consider the Fourier spectra $\langle C_{q_1,q_2}^n \rangle_M$ and $\langle C_{q_1,q_2,q_3}^{n_1,n_2} \rangle_M$ averaged over a large number M of diffraction patterns [6, 7],

$$\left\langle C_{q_1,q_2}^n \right\rangle_M = \frac{1}{M} \sum_{m=1}^{M} \left\{ C_{q_1,q_2}^n \right\}^m, \tag{10a}$$

$$\left\langle C_{q_1,q_2,q_3}^{n_1,n_2} \right\rangle_M = \frac{1}{M} \sum_{m=1}^{M} \left\{ C_{q_1,q_2,q_3}^{n_1,n_2} \right\}^m, \tag{10b}$$

where $\{C_{q_1,q_2}^n\}^m$ and $\{C_{q_1,q_2,q_3}^{n_1,n_2}\}^m$ are the Fourier components of the CCFs $\{C(q_1, q_2, \Delta)\}^m$ and $\{C(q_1, q_2, q_3, \Delta_1, \Delta_2)\}^m$, respectively, defined for the mth realization of the system (see (3a) and (3b)).

3. Scattering from a Disordered System of Identical Particles

We consider a scattering experiment in transmission geometry as shown in Figure 1(a). A coherent X-ray beam scatters from a disordered sample, and a speckle pattern is measured on the detector in the far-field regime. In our simulations we consider a kinematical scattering approximation. As a general model system we assume a 3D sample consisting of N identical particles with random positions and orientations. This model includes a variety of systems, for example, clusters or molecules in the gas phase, local structures formed in colloidal systems, viruses in solution, and so forth.

The amplitude $A_k(\mathbf{q})$ scattered from the kth particle at the momentum transfer vector $\mathbf{q}$ can be defined as [4, 5]

$$A_k(\mathbf{q}) = \int \rho_k(\mathbf{r})\, e^{i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r}, \tag{11}$$

where $\rho_k(\mathbf{r})$ is an electron density of the kth particle at the position $\mathbf{R}_k$ (see Figure 1(b)) and the integration is performed over the volume of the particle. Using (11) the intensity $I(\mathbf{q})$ coherently scattered from a disordered sample consisting of N particles is given by

$$I(\mathbf{q}) = \sum_{k_1,k_2=1}^{N} e^{i\mathbf{q}\cdot\mathbf{R}_{k_2,k_1}} A_{k_1}^*(\mathbf{q})\, A_{k_2}(\mathbf{q}) = \sum_{k_1,k_2=1}^{N} \int\int \rho_{k_1}^*(\mathbf{r}_1)\, \rho_{k_2}(\mathbf{r}_2)\, e^{i\mathbf{q}\cdot\mathbf{R}_{k_2,k_1}^{21}} d\mathbf{r}_1\, d\mathbf{r}_2, \tag{12}$$

where the double summation is performed over all N particles, and the integration is performed over the volume of the k_ith particle ($i = 1, 2$). Here, the following notation for the radius vectors is used, $\mathbf{R}_{k_2,k_1}^{21} = \mathbf{R}_{k_2,k_1} + \mathbf{r}_{21}$, where $\mathbf{R}_{k_2,k_1} = \mathbf{R}_{k_2} - \mathbf{R}_{k_1}$ is the vector connecting two different particles k_1 and k_2, and $\mathbf{r}_{21} = \mathbf{r}_2 - \mathbf{r}_1$, where the vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ define the positions of scatterers (e.g., colloidal spheres or atoms) inside the particles k_1 and k_2, respectively (see Figures 1(b) and 1(c)).

In the case of a partially coherent illumination and a dilute disordered system when the mean distance between the particles is larger than the coherence length of the incoming beam, the inter-particle correlations due to coherent interference of scattered amplitudes from the individual particles in (12) can be neglected. In these conditions, the total scattered intensity $I(\mathbf{q})$ can be represented as a sum of intensities $I_k(\mathbf{q}) = |A_k(\mathbf{q})|^2$ corresponding to individual particles in the system

$$I(\mathbf{q}) = \sum_{k=1}^{N} I_k(\mathbf{q}). \tag{13}$$

4. 2D Disordered Systems

In this section we consider a particular case of a 2D system in a small angle scattering geometry when all vectors are defined in a 2D plane (see Figure 1(c)), and the electron density of the kth particle transforms to a projected electron density, $\tilde{\rho}_k(\mathbf{r}) = \int \rho_k(\mathbf{r}, z) dz$. It was shown [4, 5], that in the case of plane wavefront illumination of a 2D system only even ($n = 2l$, $l = 1, 2, 3, \ldots$) Fourier components of the intensity I_q^n can have non-zero values.

4.1. Dilute Systems. The intensity $I_{\psi_0}(\mathbf{q})$ scattered from a single particle in some reference orientation ψ_0 (without loss of generality we fix the reference orientation to $\psi_0 = 0$) is related to the projected electron density of the particle $\tilde{\rho}_{\psi_0}(\mathbf{r})$ through its scattered amplitude (see (11)) as $I_{\psi_0}(\mathbf{q}) = |A_{\psi_0}(\mathbf{q})|^2$. Similar to $I(\mathbf{q})$ (see (6a) and (6b)), the intensity $I_{\psi_0}(\mathbf{q}) \equiv I_{\psi_0}(q, \varphi)$ can be represented as an angular Fourier series expansion,

$$I_{\psi_0}(q, \varphi) = \sum_{n=-\infty}^{\infty} I_{q,\psi_0}^n e^{in\varphi}, \tag{14}$$

where I_{q,ψ_0}^n are the Fourier components of $I_{\psi_0}(q, \varphi)$.

For a dilute 2D system of identical particles, the intensity $I_{\psi_k}(q, \varphi)$ scattered from a particle in an arbitrary orientation

ψ_k is related to the intensity $I_{\psi_0}(q,\varphi)$ scattered from a particle in the reference orientation ψ_0 as $I_{\psi_k}(q,\varphi) = I_{\psi_0}(q,\varphi-\psi_k)$. Applying the shift theorem for the Fourier transforms [37] we obtain for the corresponding Fourier components of the intensities, $I^n_{q,\psi_k} = I^n_{q,\psi_0}\exp(-in\psi_k)$. Using these relations we can write for the Fourier components I^n_q of the intensity $I(q,\varphi)$ scattered from N particles (see (13))

$$I^n_q = I^n_{q,\psi_0}\sum_{k=1}^{N} e^{-in\psi_k} = I^n_{q,\psi_0}\mathbf{A}_n, \tag{15}$$

where $\mathbf{A}_n = \sum_{k=1}^{N}\exp(-in\psi_k)$ is a random phasor sum [38]. Equation (15) leads to the following expression for the small-angle X-ray scattering (SAXS) intensity, $\langle I(q,\varphi)\rangle_\varphi = NI^0_{q,\psi_0}$.

The Fourier component I^n_{q,ψ_0} is related to the projected electron density $\tilde{\rho}_{\psi_0}(\mathbf{r})$ of the particle as [4–6]

$$I^n_{q,\psi_0} = \int\int \tilde{\rho}^*_{\psi_0}(\mathbf{r}_1)\,\tilde{\rho}_{\psi_0}(\mathbf{r}_2) \times J_n(q|\mathbf{r}_{21}|)\,e^{-in\phi_{\mathbf{r}_{21}}}d\mathbf{r}_1\,d\mathbf{r}_2, \tag{16}$$

where $\phi_{\mathbf{r}_{21}}$ is the angle of the vector $\mathbf{r}_{21}$ in the detector plane, $J_n(\rho)$ is the Bessel function of the first kind of integer order n, and the integration is performed over the area of a particle. According to the structure of I^n_{q,ψ_0} its value strongly depends on the symmetry of a particle and determines selection rules [4–6] for the values n of non-zero Fourier components $C^n_{q_1,q_2}$. These selection rules can be used for the identification of the symmetry of particles in dilute systems. For example, for a cluster with 5-fold symmetry only $n = 10l, (l = 1, 2, \ldots)$ will give a non-zero contribution to the Fourier components of CCFs.

Substituting (15) in (5), in the limit of dilute systems we have for the Fourier components $C^n_{q_1,q_2}$ of the CCF the following expression:

$$C^n_{q_1,q_2} = I^{n*}_{q_1,\psi_0}I^n_{q_2,\psi_0}|\mathbf{A}_n|^2. \tag{17}$$

The statistical behavior of $\mathbf{A}_n$ has been analyzed for different angular distributions of orientations of particles in the system (see [4–6]). It is clear, that in the case of a completely oriented system of particles (all $\psi_k = 0$) the square amplitude of the random phasor sum $|\mathbf{A}_n|^2$ is equal to N and $C^n_{q_1,q_2} = NI^{n^*}_{q_1,\psi_0}I^n_{q_2,\psi_0}$. In the case of a uniform distribution of orientations of particles $|\mathbf{A}_n|^2$ fluctuates around its mean value $\langle|\mathbf{A}_n|^2\rangle = N$ with the standard deviation $\sigma_{|\mathbf{A}_n|^2} = N$. Averaging the Fourier components $C^n_{q_1,q_2}$ over a large number M of diffraction patterns decreases these fluctuations and leads to the following asymptotic result [4–6]:

$$\left\langle C^n_{q_1,q_2}\right\rangle_M = I^{n^*}_{q_1,\psi_0}I^n_{q_2,\psi_0} \times \left\langle|\mathbf{A}_n|^2\right\rangle_M \xrightarrow[M\to\infty]{} NI^{n^*}_{q_1,\psi_0}I^n_{q_2,\psi_0}, \tag{18}$$

where $\langle\cdots\rangle_M$ denotes statistical averaging over M diffraction patterns. Importantly, the ensemble-averaged Fourier components $\langle C^n_{q_1,q_2}\rangle_M$ converge to a scaled product of the two Fourier components of intensity $I^{n^*}_{q_1,\psi_0}$ and $I^n_{q_2,\psi_0}$ associated with a single particle.

The amplitudes $|I^n_{q,\psi_0}|$ and phases ϕ^n_{q,ψ_0} (for $n\neq 0$) of the Fourier components $I^n_{q,\psi_0} = |I^n_{q,\psi_0}|\exp(i\phi^n_{q,\psi_0})$ associated with a single particle can be determined using (18) [7]. This equation determines the phase difference between two Fourier components $I^n_{q_1,\psi_0}$ and $I^n_{q_2,\psi_0}$ of the same order n, defined at two different resolution rings q_1 and q_2,

$$\arg\left[\left\langle C^n_{q_1,q_2}\right\rangle_M\right] = \phi^n_{q_2,\psi_0} - \phi^n_{q_1,\psi_0}. \tag{19}$$

Similar to the Fourier components of the two-point CCF, the Fourier components of the three-point CCF (9) can be expressed in the limit of a dilute system as [7]

$$C^{n_1,n_2}_{q_1,q_2,q_3} = I^{(n_1+n_2)^*}_{q_1,\psi_0}I^{n_1}_{q_2,\psi_0}I^{n_2}_{q_3,\psi_0}\cdot\mathbf{A}_{n_1,n_2}, \tag{20}$$

where $\mathbf{A}_{n_1,n_2} = \sum_{i,j,k=1}^{N}\exp\{i[(n_1+n_2)\psi_i - n_1\psi_j - n_2\psi_k]\}$ is another random phasor sum. Our analysis shows [7] that in the case of a uniform distribution of orientations of N particles the statistical average $\langle\mathbf{A}_{n_1,n_2}\rangle_M$ converges to N for a sufficiently large number M of diffraction patterns

$$\left\langle C^{n_1,n_2}_{q_1,q_2,q_3}\right\rangle_M \xrightarrow[M\to\infty]{} NI^{(n_1+n_2)^*}_{q_1,\psi_0}I^{n_1}_{q_2,\psi_0}I^{n_2}_{q_3,\psi_0}. \tag{21}$$

An important result of (21) is that the ensemble-averaged Fourier components $\langle C^{n_1,n_2}_{q_1,q_2,q_3}\rangle_M$ converge to a scaled product of three Fourier components of intensity $I^{(n_1+n_2)^*}_{q_1,\psi_0}$, $I^{n_1}_{q_2,\psi_0}$, and $I^{n_2}_{q_3,\psi_0}$ associated with a single particle. Equation (21) also provides the following phase relation:

$$\arg\left[\left\langle C^{n_1,n_2}_{q_1,q_2,q_3}\right\rangle_M\right] = \phi^{n_1}_{q_2,\psi_0} + \phi^{n_2}_{q_3,\psi_0} - \phi^{(n_1+n_2)}_{q_1,\psi_0}. \tag{22}$$

This equation determines the phase difference between three Fourier components $I^{(n_1+n_2)}_{q_1,\psi_0}$, $I^{n_1}_{q_2,\psi_0}$, and $I^{n_2}_{q_3,\psi_0}$ of different order n defined on three resolution rings. If $n_1 = n_2 = n$ and $n_3 = 2n$, (22) reduces to a particular form, giving the phase relation between Fourier components of only two different orders n and $2n$. Phase relations (19) and (22) can be used to determine the phases of the complex Fourier components I^n_{q,ψ_0} using measured CCFs from a disordered system of particles [7].

4.2. Dense Systems. In the case of a dense system, when the average distance between particles is of the order of the size of a single cluster, the Fourier components I^n_q of the intensity $I(\mathbf{q})$ (see (12)) can contain a substantial inter-particle contribution. In this case I^n_q can be presented as a sum of two terms,

$$I^n_q = I^n_{\text{part}}(q) + I^n_{\text{int-part}}(q), \tag{23}$$

where $I^n_{\text{part}}(q)$ is attributed to a single particle structure discussed previously, and $I^n_{\text{int-part}}(q)$ is defined by the inter-particle correlations [4–6]. In a 2D system these two terms are [4–6]

$$I^n_{\text{part}}(q) = I^n_{q,\psi_0}\mathbf{A}_n, \tag{24a}$$

$$I^n_{\text{int-part}}(q) = 2\sum_{k_2>k_1}\int\int \widetilde{\rho}^*_{k_1}(\mathbf{r}_1)\,\widetilde{\rho}_{k_2}(\mathbf{r}_2) \times J_n\left(q\left|\mathbf{R}^{21}_{k_2,k_1}\right|\right)e^{-in\phi_{\mathbf{R}^{21}_{k_2,k_1}}}d\mathbf{r}_1\,d\mathbf{r}_2, \tag{24b}$$

where $\phi_{\mathbf{R}^{21}_{k_2,k_1}}$ is the angle of the vector $\mathbf{R}^{21}_{k_2,k_1}$ in the sample plane.

Taking into account both terms of (23), the Fourier components of the two-point CCF for a single realization of a system (5) can be written as the following sum of four terms:

$$C^n_{q_1,q_2} = I^{n^*}_{q_1}\cdot I^n_{q_2} = S^n_1 + \left(S^n_2 + S^n_3\right) + S^n_4, \tag{25}$$

where

$$S^n_1 = I^{n^*}_{\text{part}}(q_1)\,I^n_{\text{part}}(q_2), \tag{26a}$$

$$S^n_2 + S^n_3 = I^{n^*}_{\text{part}}(q_1)\,I^n_{\text{int-part}}(q_2) + I^{n^*}_{\text{int-part}}(q_1)\,I^n_{\text{part}}(q_2), \tag{26b}$$

$$S^n_4 = I^{n^*}_{\text{int-part}}(q_1)\,I^n_{\text{int-part}}(q_2). \tag{26c}$$

In (25) the term S^n_1 is defined only by the structure of a particle, whereas the terms S^n_2, S^n_3, and S^n_4 contain contributions from the inter-particle correlations due to the second term in (23). A similar expansion can be performed for the Fourier components of the three-point CCF. The influences of inter-particle correlations on CCFs measured on the same resolution ring were discussed in detail in [6].

5. 3D Disordered Systems

In our previous discussion of X-ray scattering on 2D systems, we have seen that only even Fourier components of the CCFs have non-zero values. Here we consider a general case of 3D systems in which one or more particles are distributed with random positions and orientations in 3D space. In the case of 3D systems non-zero odd Fourier components can be also present when scattering to high angles is considered, due to Ewald sphere curvature effects.

In general, the scattering vector $\mathbf{q} = (\mathbf{q}^\perp, q^z)$ can be decomposed into two components: $\mathbf{q}^\perp$ that is, perpendicular, and q^z that is, parallel to the direction of the incident beam (see Figure 3(b)). We also define the perpendicular $\mathbf{R}^{\perp 21}_{k_2,k_1} = \mathbf{R}^\perp_{k_2,k_1} + \mathbf{r}^\perp_{21}$, and the z-components $Z^{21}_{k_2,k_1} = Z_{k_2,k_1} + z_{21}$ of the radius vectors (see Figures 1(a) and 3). Using these notations we can write (12) in the following form:

$$I(\mathbf{q}) = \sum_{k_1,k_2=1}^{N} e^{-iq^z\cdot Z_{k_2,k_1}}\int\int \widetilde{\rho}^*_{k_1}(\mathbf{r}^\perp_1, q^z)\,\widetilde{\rho}_{k_2}(\mathbf{r}^\perp_2, q^z) \times e^{i\mathbf{q}^\perp\cdot\mathbf{R}^{\perp 21}_{k_2,k_1}}d\mathbf{r}^\perp_1 d\mathbf{r}^\perp_2. \tag{27}$$

Here we introduced a modified complex valued electron density function, defined as

$$\widetilde{\rho}_{k_i}(\mathbf{r}^\perp_i, q^z) = \int \rho_{k_i}(\mathbf{r}^\perp_i, z)\,e^{-iq^z z}dz. \tag{28}$$

In the case of wide angle scattering, the effect of the Ewald sphere curvature (see Figure 3(b)), which manifests itself by the presence of the exponential factors $e^{-iq^z\cdot Z_{k_2,k_1}}$ and $e^{-iq^z z}$ in (27) and (28), may become important. This effect can break the scattering symmetry of a diffraction pattern, characteristic for the scattering on a positive valued electron density (Friedel's law) and may help to reveal symmetries that can not be directly observed in the small angle scattering case. A wide angle scattering geometry may become important for scattering on atomic systems with local interatomic distances of the order of few Ångstroms.

For simplicity we will consider here a 3D system consisting of particles composed of identical scatterers. The modified electron density of a particle according to (28) can be defined in the following form:

$$\widetilde{\rho}_k(\mathbf{r}^\perp, q^z) = f(q)\sum_{i=1}^{N_s}\delta(\mathbf{r}^\perp - \mathbf{r}^\perp_i)\,e^{-iq^z z_i}, \tag{29}$$

where $f(q)$ is a form-factor of a scatterer, and N_s is a number of scatterers in the cluster. The coordinates $(\mathbf{r}^\perp_i, z_i)$ define the position of the ith scatterer inside the kth cluster. Using this definition and performing a Fourier transformation of (27) we obtain [4, 5]

$$I^n(q^\perp, q^z) = (i)^n\left|f(q)\right|^2\sum_{k_1,k_2=1}^{N}\sum_{l,m=1}^{N_s} e^{-iq^z Z^{ml}_{k_2,k_1}} \times J_n\left(\left|\mathbf{q}^\perp\right|\cdot\left|\mathbf{R}^{\perp ml}_{k_2,k_1}\right|\right)e^{-in\phi_{\mathbf{R}^{\perp ml}_{k_2,k_1}}}, \tag{30}$$

where the summation over index l is performed over the positions of scatterers in the cluster k_1, and the summation over index m is performed over the positions of scatterers in the cluster k_2. We note here that due to the property of the Bessel functions [$J_n(0) = 0$ for $n \neq 0$] the terms with $k_1 = k_2$ and $l = m$ are equal to zero. Taking into account that the terms with interchanged indices, that is, k_1, k_2 and k_2, k_1, as well as l, m and m, l, differ from each other by a change of the sign of $Z^{ml}_{k_2,k_1}$ and by an additional factor $(-1)^n$, which arises due to

the change of the phase $\phi_{\mathbf{R}^{\perp ml}_{k_2,k_1}} = \phi_{\mathbf{R}^{\perp lm}_{k_1,k_2}} + \pi$, we have for *even* values of n in (30) [4, 5]

$$
\begin{aligned}
&I^n\left(q^{\perp}, q^z\right) \\
&\quad = 2(i)^n \left|f(q)\right|^2 \\
&\qquad \times \Bigg[\sum_{\substack{1\le k_1\le N \\ k_1<k_2\le N}} \sum_{\substack{1\le l\le N_s \\ 1\le m\le N_s}} \cos\left(q^z Z^{ml}_{k_2,k_1}\right) \\
&\qquad\qquad \times J_n\left(\left|\mathbf{q}^{\perp}\right| \cdot \left|\mathbf{R}^{\perp ml}_{k_2,k_1}\right|\right) e^{-in\phi_{\mathbf{R}^{\perp ml}_{k_2,k_1}}} \\
&\qquad + \sum_{1\le k\le N} \sum_{\substack{1\le l\le N_s \\ l\le m\le N_s}} \cos\left(q^z Z^{ml}_{k}\right) \\
&\qquad\qquad \times J_n\left(\left|\mathbf{q}^{\perp}\right| \cdot \left|\mathbf{R}^{\perp ml}_{k}\right|\right) e^{-in\phi_{\mathbf{R}^{\perp ml}_{k}}} \Bigg]
\end{aligned}
\tag{31}
$$

and for *odd* values of n:

$$
\begin{aligned}
&I^n\left(q^{\perp}, q^z\right) \\
&\quad = -2(i)^{n+1} \left|f\left(q\right)\right|^2 \\
&\qquad \times \Bigg[\sum_{\substack{1\le k_1\le N \\ k_1<k_2\le N}} \sum_{\substack{1\le l\le N_s \\ 1\le m\le N_s}} \sin\left(q^z Z^{ml}_{k_2,k_1}\right) \\
&\qquad\qquad \times J_n\left(\left|\mathbf{q}^{\perp}\right| \cdot \left|\mathbf{R}^{\perp ml}_{k_2,k_1}\right|\right) e^{-in\phi_{\mathbf{R}^{\perp ml}_{k_2,k_1}}} \\
&\qquad + \sum_{1\le k\le N} \sum_{\substack{1\le l\le N_s \\ l\le m\le N_s}} \sin\left(q^z Z^{ml}_{k}\right) \\
&\qquad\qquad \times J_n\left(\left|\mathbf{q}^{\perp}\right| \cdot \left|\mathbf{R}^{\perp ml}_{k}\right|\right) e^{-in\phi_{\mathbf{R}^{\perp ml}_{k}}} \Bigg],
\end{aligned}
\tag{32}
$$

where $\mathbf{R}^{\perp ml}_{k} = \mathbf{R}^{\perp ml}_{k,k}$ and $Z^{ml}_{k} = Z^{ml}_{k,k}$. From the performed analysis we can see that, due to the curvature of the Ewald sphere (non-zero q^z component), we obtain non-zero odd Fourier components of the CCF when scattering from a 3D system. These components become negligibly small for experimental conditions corresponding to a flat Ewald sphere, considered in the previous section.

In the case of a dilute sample, the Fourier components of intensity defined in (29) reduce to

$$
\begin{aligned}
&I^n\left(q^{\perp}, q^z\right) \\
&\quad = (i)^n \left|f\left(q\right)\right|^2 \sum_{k=1}^{N} \sum_{l,m=1}^{N_s} e^{-iq^z z^{ml}_k} \\
&\qquad \times J_n\left(\left|\mathbf{q}^{\perp}\right| \cdot \left|\mathbf{r}^{\perp ml}_{k}\right|\right) e^{-in\phi_{\mathbf{r}^{\perp ml}_{k}}}.
\end{aligned}
\tag{33}
$$

6. Results and Discussion

In this section we provide two examples of application of XCCA for investigation of the structural properties of disordered systems.

6.1. Recovery of the Structure of an Individual 2D Particle. In the pioneering work of Kam [11, 12] it was proposed to determine the structure of a single particle using scattered intensity from many identical particles in solution. However, this approach, based on a spherical harmonics expansion of the scattered amplitudes, was not fully explored until now. Recently, it has been revised theoretically [8, 33, 36, 39] and experimentally [9]. The possibility to recover the structure of individual particle was demonstrated in systems in two dimensions [8, 9, 36] and three dimensions [33] using additional *a priori* knowledge on the symmetry of the particles. Unfortunately, these approaches, based on optimization routines [33, 36] and iterative techniques [8, 9], do not guarantee the uniqueness of the recovered structure unless strong constraints are applied. Here we provide a brief overview of our recently developed approach [7] that enables a direct reconstruction of the single-particle intensity distribution using an algebraic formalism of two- and three-point CCFs without additional constraints. This approach is developed for a 2D system of particles, but can be also used to study 3D systems provided that particles can be aligned with respect to a certain axis.

Our goal here is to determine the scattering pattern of a single particle $I_{\psi_0}(\mathbf{q})$ using a large number of diffraction patterns $I(\mathbf{q})$ (see (12)) corresponding to different realizations of the system. Once the single-particle intensity is obtained, conventional phase retrieval algorithms [40, 41] can recover the projected electron density of a single particle. According to (14) the intensity scattered from a single particle can be uniquely determined by the set of complex coefficients $\{I^n_{q,\psi_0}\} = \{|I^n_{q,\psi_0}|, \phi^n_{q,\psi_0} = \arg(I^n_{q,\psi_0})\}$. We will determine the Fourier components $\{I^n_{q,\psi_0}\}$ applying two- and three-point CCFs to the measured intensities $I(q,\varphi)$ scattered from N particles.

Equations (18), (19), and (22) constitute the core of our approach and allow us to directly determine the complex Fourier components I^n_{q,ψ_0}. According to our simulations, these expressions, originally derived in the dilute limit approximation (see the incoherent sum (13)), can be also used in the case of coherent scattering from a system of particles. Also note, that in the general case of coherent scattering, the 0th order Fourier component can be determined as $I^0_{q,\psi_0} = 1/N\langle\langle I(q,\varphi)\rangle_{\varphi}\rangle_M$, which gives the same result as for the case of incoherent scattering [7]. The amplitudes $|I^n_{q,\psi_0}|$ for $n \neq 0$ can be determined using (18), and the phases ϕ^n_{q,ψ_0} using (19) and (22), in particular when $n_1 = n_2$. It might be tempting to use (7b) to determine the amplitudes $|I^n_{q,\psi_0}|$ by just taking the square root of $\langle C^n_q\rangle_M$. However, in the case of coherent scattering from N particles the spectrum $\langle C^n_q\rangle_M$ may contain a substantial inter-particle contribution or noise [4–6]. In contrast to this, even in the case of coherent scattering from

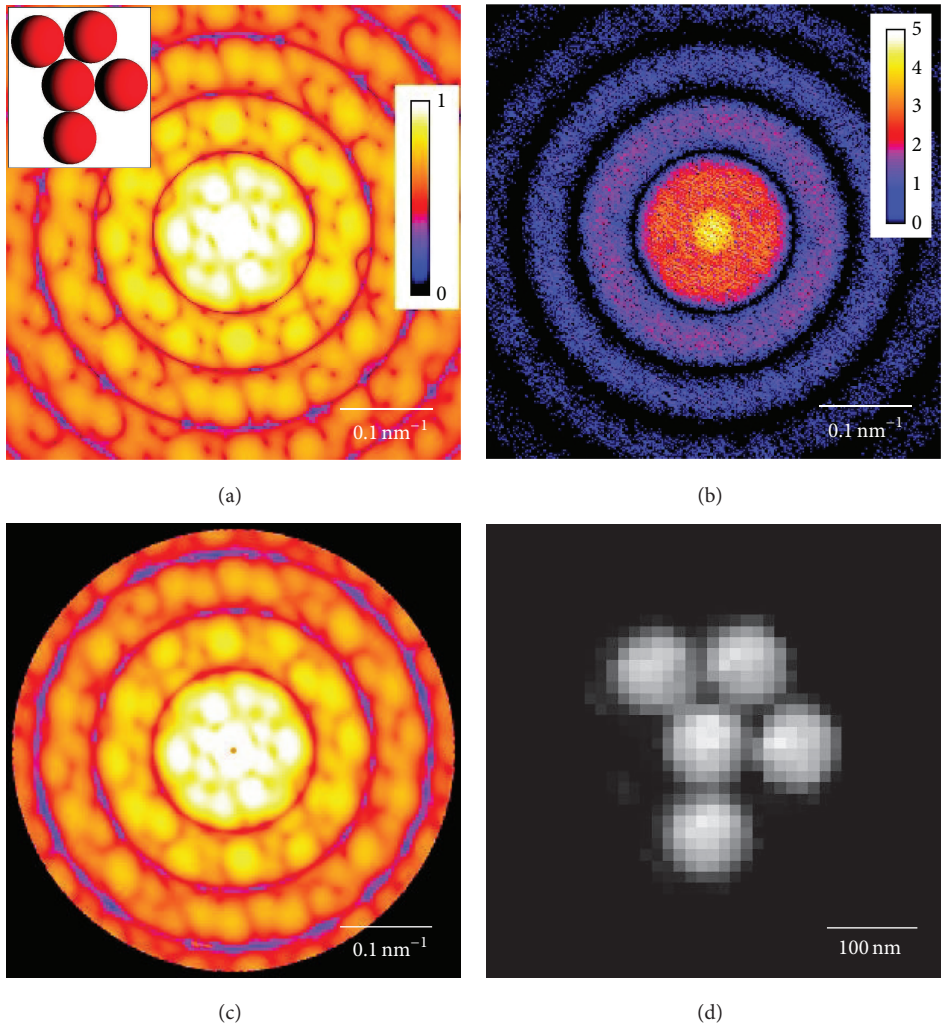

FIGURE 4: (a) Scattered intensity (logarithmic scale) calculated for a single asymmetric cluster (shown in the inset). (b) Coherently scattered intensity from a disordered system consisting of $N = 10$ clusters in random position and orientation. Scattered signal corresponds to the incident fluence of $4 \cdot 10^{11}$ photons/μm^2 and contains Poisson noise. (c) Scattered intensity corresponding to an asymmetric cluster recovered from $M = 10^5$ diffraction patterns of the form (b). (d) Structure of a single cluster reconstructed by an iterative phase retrieval algorithm using the diffraction pattern shown in (c). The intensities in (a), (c) are given in arbitrary units, and in (b) in photon counts.

a dilute system of particles, the inter-particle contribution to the spectrum $\langle C^n_{q_1,q_2} \rangle_M$ determined at *different* resolution rings $q_1 \neq q_2$ is negligibly small. For the same reason we used the three-point CCFs (see (2)) defined at different resolution rings q_1, q_2, and q_3. Once all amplitudes and phases of the Fourier components I^n_{q,ψ_0} are determined, the diffraction pattern corresponding to a single particle can be recovered using (14).

We demonstrate our approach for the case of a coherent illumination of a disordered system of particles (see (12)) with an incident fluence of $4 \cdot 10^{11}$ photons/μm^2 (that corresponds to 10^{13} photons focused on a sample area of $5 \times 5\,\mu$m^2), in the presence of Poisson noise in the scattered signal. As an example, we recover a diffraction pattern and projected electron density of an asymmetric cluster (Figure 4(a)) composed of PMMA spheres of 50 nm radius, with a size of a cluster $d = 300$ nm. We consider coherent scattering of X-rays with wavelength $\lambda = 1$ Å from a system of $N = 10$ clusters in random positions and orientations, distributed within the sample area (see Figure 1(c)). Diffraction patterns are simulated for a 2D detector of size $D = 24$ mm (with pixel size $p = 80\,\mu$m), positioned in transmission geometry at $L = 3$ m distance from the sample (see Figure 1(a)). This experimental geometry corresponds to a maximum resolution of 25 nm. For given experimental conditions the speckle size corresponding to the illuminated area is below the pixel size of the detector. At the same time the speckle size corresponding to the size of a single particle is about 12 pixels, which provides sufficient sampling for the phase retrieval algorithm.

The coherently scattered intensity simulated for a single realization of the system (all simulations of diffraction patterns were performed using the computer code MOLTRANS) is shown in Figure 4(b). The Fourier components of two-point and three-point CCFs, $\langle C^n_{q_1,q_2} \rangle_M$ and $\langle C^{n,n}_{q_1,q_2,q_3} \rangle_M$, were averaged over $M = 10^5$ diffraction patterns. It is important

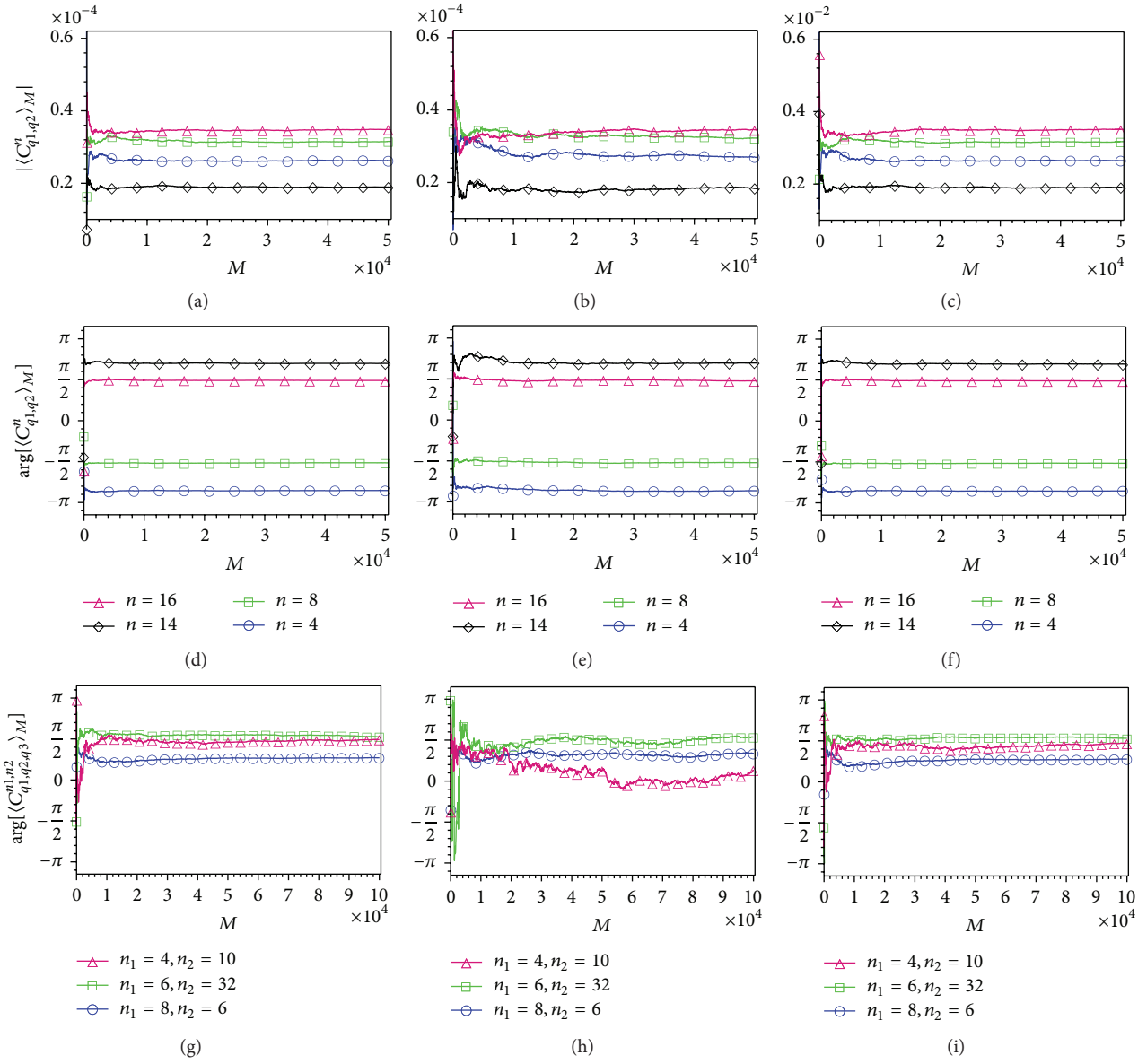

FIGURE 5: (a)–(f) Statistical convergence of the amplitudes $|\langle C^n_{q_1,q_2}\rangle_M|$ and phases $\arg[\langle C^n_{q_1,q_2}\rangle_M]$ at $q_1 = 0.25$ nm^{-1}, $q_2 = 0.09$ nm^{-1}, $n = 4, 8, 14$, and 16, calculated as a function of the number of diffraction patterns M used in the averaging. (g)–(i) The phases $\arg[\langle C^{n_1,n_2}_{q_1,q_2,q_3}\rangle_M]$ at $q_1 = 0.23$ nm^{-1}, $q_2 = 0.24$ nm^{-1}, and $q_3 = 0.25$ nm^{-1}, calculated for three different combinations of n_1 and n_2 as a function of M. Simulations were performed (a), (d), (g) without noise, (b), (e), (h) for the incident photon fluence $4 \cdot 10^{10}$ photons/μm^2, and (c), (f), (i) for the photon fluence $4 \cdot 10^{11}$ photons/μm^2, and the scattered intensity is expressed in photon counts. In (b), (e), (h), and (c), (f), (i) Poisson noise was included.

to perform the averaging of the two- and three-point CCFs over a sufficient number M of diffraction patterns to achieve their convergence to the mean value. In practical applications the convergence of the CCFs can be directly controlled as a function of the number of the diffraction patterns M considered in the averaging [6]. The results presented in Figure 5 demonstrate the evolution of the Fourier components $\langle C^n_{q_1,q_2}\rangle_M$ as a function of M. As one can see, after averaging the CCFs over a sufficiently large number M of diffraction patterns, the amplitudes and phases of $\langle C^n_{q_1,q_2}\rangle_M$ asymptotically converge to their mean values. Comparing Figures 5(a) and 5(d) with Figures 5(b) and 5(e) one can see that the CCFs converge slower in the presence of Poisson noise, and one needs to average over more patterns. This effect is more prominent for the three-point CCFs. In Figures 5(g)–5(i) the phases of the Fourier components $\langle C^{n_1,n_2}_{q_1,q_2,q_3}\rangle_M$ are presented for three different combinations of n_1 and n_2. It is well seen from these figures, that at the incident fluence of $4{\cdot}10^{11}$ photons/μm^2 the phases and amplitudes of the Fourier components are reaching their mean values at $M = 10^5$.

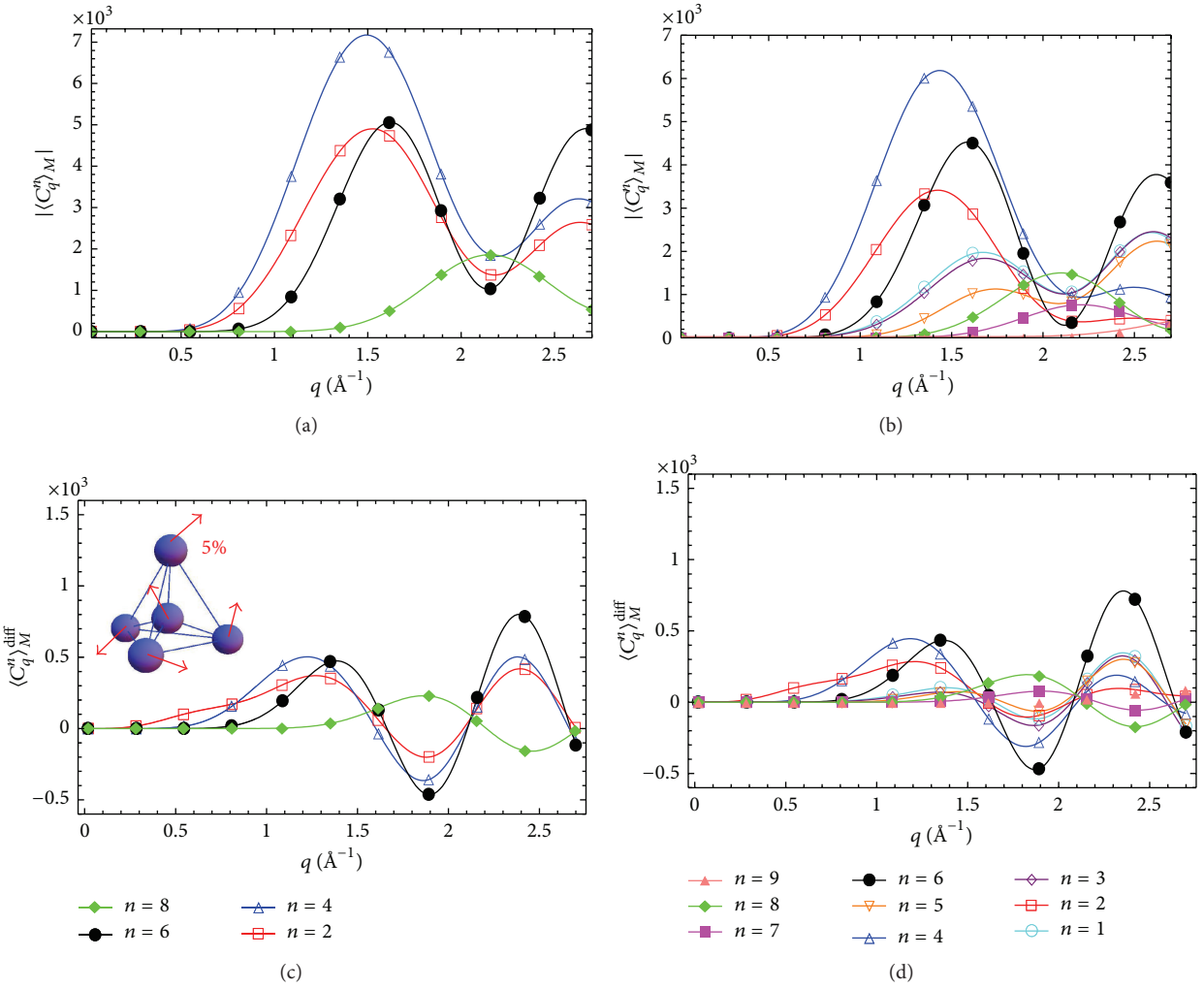

FIGURE 6: Fourier spectra $\langle C_q^n \rangle_M$ calculated as a function of q for $1 \leq n \leq 9$, for a flat (a), (c) and curved (b), (d) Ewald sphere. (a), (b) Amplitudes $|\langle C_q^n \rangle_M|$ calculated in dilute limit approximation of scattering (see (13)) from $N = 10$ clusters, averaged over $M = 10^4$ diffraction patterns. (c), (d) Difference spectra $\langle C_q^n \rangle_M^{\text{diff}}$ calculated in the dilute limit approximation for a distorted oxygen pentamer for $N = 10$, and $M = 10^4$ (see text for details). All atoms in each pentamer cluster in the sample were randomly displaced by 5% of the O-O distance from the positions of a regular pentamer. The scattered intensities used for calculation of the spectra shown in Figures 6 and 7 are expressed in the units of classical electron radius squared, and the spectra were normalized by the number of particles N in the system.

The diffraction pattern corresponding to a single particle recovered by our approach is presented in Figure 4(c). As one can see, it reproduces the diffraction pattern of an individual cluster shown in Figure 4(a) very well. The structure of the cluster reconstructed from this recovered diffraction pattern was obtained by a standard phase retrieval approach [40] and is presented in Figure 4(d). The comparison of the single cluster structure obtained by our approach (Figure 4(d)) with the initial model shown in the inset of Figure 4(a) confirms the correctness of our reconstruction. The particle was reconstructed with a resolution of 25 nm. These results clearly demonstrate the ability of our approach to recover the single-particle structure from noisy data obtained in coherent X-ray scattering experiments. A detailed discussion of different aspects that can affect the quality of reconstruction, such as particle density, fluctuations of the number of particles N in the system, and signal-to-noise ratio of the CCFs, can be found in [7, 42].

6.2. Wide Angle Scattering from a 3D System of Pentameric Structures. In the previous section we demonstrated the application of XCCA to the scattering data from a 2D disordered system of particles which led to a successful reconstruction of the structure of an individual particle. At the same time, for a 3D disordered system the same problem has been solved only if additional symmetry conditions were imposed [33, 34]. Hence, development of direct, self-consistent approaches for the recovery of the local structure of a 3D disordered system remains a challenge. As a step

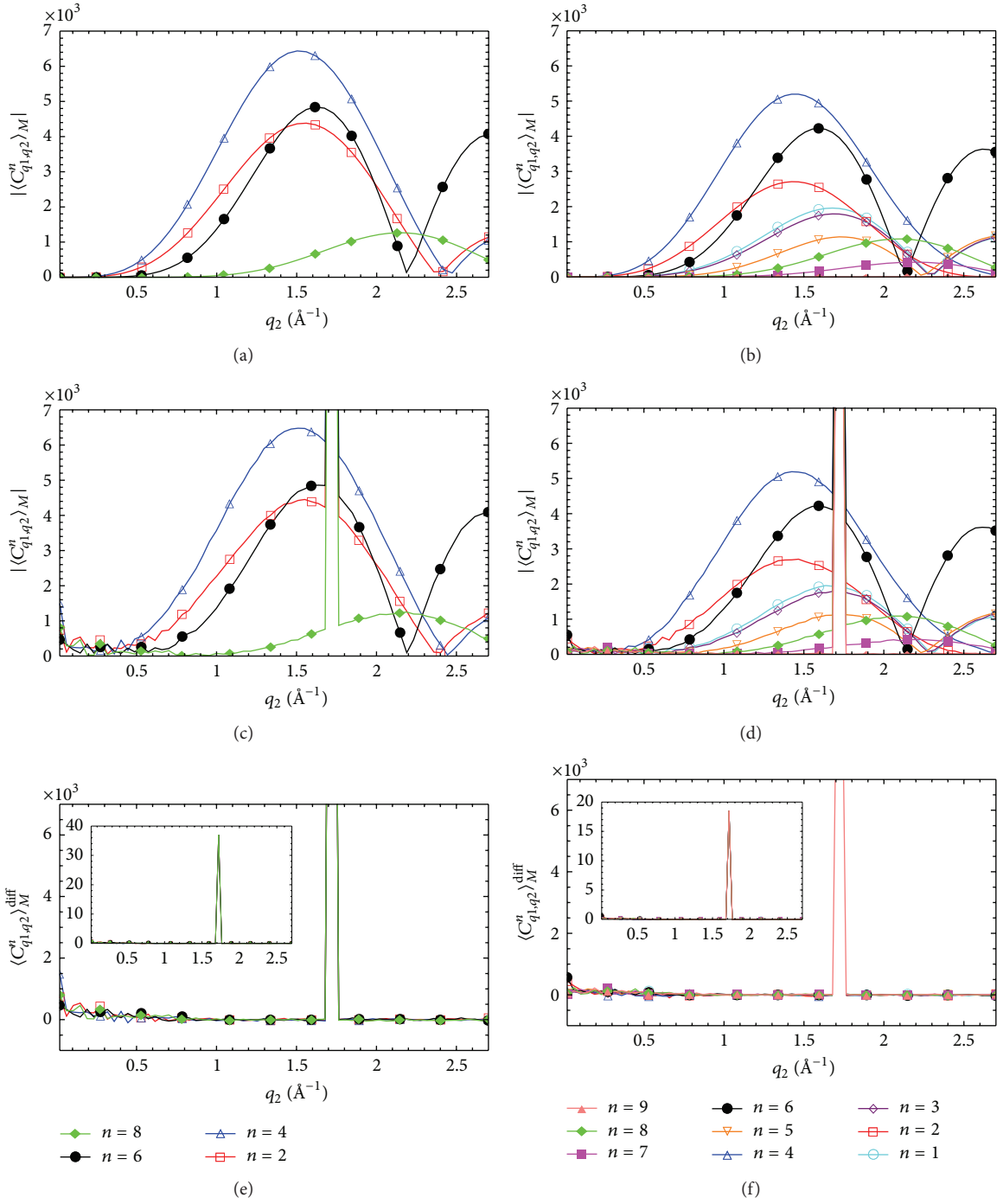

FIGURE 7: Fourier spectra $\langle C^n_{q_1,q_2}\rangle_M$ calculated as a function of q_2 ($q_1 = 1.71$ Å^{-1}) for $1 \le n \le 9$, for a flat (a), (c), (e) and curved (b), (d), (f) Ewald sphere. (a), (b) Amplitudes $|\langle C^n_{q_1,q_2}\rangle_M|$ calculated in dilute limit approximation of scattering (see (13)) from $N = 10$ clusters, averaged over $M = 10^4$ diffraction patterns. (c), (d) Amplitudes $|\langle C^n_{q_1,q_2}\rangle_M|$ calculated for the case of coherent scattering (see (12)) from a cubic sample of 800 nm in size consisting of $N = 400$ clusters, averaged over $M = 10^6$ diffraction patterns. (e), (f) Difference spectra $\langle C^n_{q_1,q_2}\rangle^{\text{diff}}_M$ show the difference (e) between the spectra in (c) and (a), and (f) between (d) and (b). The insets in (e) and (f) show full size of the peaks attributed to the inter-particle contribution.

towards the solution of this problem we present the results of simulations of X-ray scattering experiments from 3D disordered systems composed of oxygen tetrahedral pentamers (Figure 1(b)). The structure of an oxygen pentamer corresponds to the tetrahedral arrangement of water molecules in a so-called Walrafen pentamer [43], one of the favorable models of water that reproduces the results of X-ray and neutron scattering experiments [43–47]. We perform the XCCA of the diffraction patterns originating from such structures and show that the Fourier spectra of the CCFs contain information on the local structure of a disordered system that is not accessible in the standard SAXS experiment [35].

We consider here a scattering experiment on a 3D disordered system composed of N oxygen pentamers (see Figure 1(a)) with the following parameters: X-ray wavelength $\lambda = 1.0$ Å, detector size $D = 24$ mm (with a pixel size of $p = 80$ μm), sample-detector distance $L = 25$ mm. The samples were composed of a different number N of oxygen pentamers. In our model we assume a uniform distribution of orientations of clusters. Three Euler angles were used to generate this distribution. The regular pentamer (see Figure 1(b)) was defined by the O-O distance of 2.82 Å between the central oxygen atom and each atom in a vertex, and a tetrahedral angle 109.47° typical for water [44]. We demonstrate results of the calculations of the Fourier spectra $\langle C_q^n \rangle_M$ and $\langle C_{q_1,q_2}^n \rangle_M$ averaged over a large number M of realizations of a system.

First, we consider scattering from a dilute system of clusters, where the intensity $I(\mathbf{q})$ scattered from N clusters is approximated by a sum of intensities scattered from each cluster (see (13) and (33)). The Fourier spectra $\langle C_q^n \rangle_M$ defined on a single resolution ring q and averaged over $M = 10^4$ diffraction patterns for $N = 10$ clusters in the system are presented in Figures 6(a)-6(b). The spectrum C_q^n for each realization of the system was directly determined using (7b) and (33). Calculations were performed for a regular pentamer for the case of a flat (a) and curved (b) Ewald sphere. Calculations for the case of a flat Ewald sphere (such conditions can be achieved, for example, using high photon energy) were performed by setting the z-component of the momentum transfer vector to $q^z = 0$ in (33). In the case of a flat Ewald sphere only Fourier components with even n-values have nonzero values. The amplitudes $|\langle C_q^n \rangle_M|$ calculated for a single cluster $N = 1$ in the system (not presented here), and for a system containing $N = 10$ clusters (Figures 6(a) and 6(b)), converge to the same functional dependence of q. Therefore, in dilute limit approximation the results do not depend on the number N of particles in the system, but are just scaled by N. As one can see from Figures 6(a)-6(b), in the considered q-range only the Fourier components with $n \leq 9$ contribute to the spectrum, and in the case of the curved Ewald sphere the Fourier components with odd n have magnitudes comparable with the ones with even n. Therefore, experimental conditions that correspond to curved Ewald sphere can provide complementary structural information, as compared to the case with a flat Ewald sphere.

The effect of a distortion of the regular pentamer on the Fourier spectra of the two-point CCF is demonstrated in Figures 6(c) and 6(d). These results were also obtained in the dilute limit using (33). The difference spectra $\langle C_q^n \rangle_M^{\text{diff}} = |\langle C_q^n \rangle_M^{\text{dist}}| - |\langle C_q^n \rangle_M^{\text{reg}}|$ were calculated for the case when all atoms in each cluster were randomly displaced by 5% of the O-O distance. Here, $\langle C_q^n \rangle_M^{\text{reg}}$ corresponds to the spectra calculated for the regular pentamer [Figures 6(a) and 6(b)], and $\langle C_q^n \rangle_M^{\text{dist}}$ to the distorted one (not presented). These results show that CCFs preserve their functional dependence on momentum transfer vector q with small distortions of the cluster, and, importantly, the ratio between the values of different Fourier components is not significantly changed.

It is interesting to consider similar results of simulations for the Fourier spectra $\langle C_{q_1,q_2}^n \rangle_M$ defined at different resolution rings $q_1 \neq q_2$. In Figures 7(a)–7(d) the amplitudes $|\langle C_{q_1,q_2}^n \rangle_M|$ calculated as a function of q_2 ($q_1 = 1.71$ Å^{-1}) are presented for a flat (a), (c) and curved (b), (d) Ewald sphere. The amplitudes $|\langle C_{q_1,q_2}^n \rangle_M|$ shown in Figures 7(a) and 7(b) were calculated in dilute limit from $N = 10$ clusters, using (33) in (5), and were averaged (see (10a)) over $M = 10^4$ realizations of the system. The spectra shown in Figures 7(c) and 7(d) were calculated for the case of coherent scattering from a cubic sample of 800 nm in size consisting of $N = 400$ clusters. In this case the coherently scattered intensities (see (12)) were calculated for each realization of the system, and then the two-point CCFs and their Fourier components were determined using (4a) and (4b) and averaged (see (10a)) over $M = 10^6$ diffraction patterns. A much larger number of diffraction patterns was considered in the latter case to average out the inter-particle contribution in the specra, which is absent in the former case. The difference between the spectra in Figures 7(c) and 7(a), and in Figures 7(d) and 7(b), is presented in Figures 7(e) and 7(f), respectively. It is easy to see from Figure 7, that the inter-particle contribution to the Fourier components $\langle C_{q_1,q_2}^n \rangle_M$ in the case of coherent scattering from $N = 400$ clusters is negligible in the entire q-range, except for a small region corresponding to $q_1 = q_2 = 1.71$ Å^{-1}. This is reflected in a sharp peak in the difference spectra presented in Figures 7(e) and 7(f). The width of this peak is related to the speckle size given by the sample size. As one can see from the insets in Figures 7(e) and 7(f), the magnitude of the inter-particle contribution at $q_1 = q_2$ is much higher than the structural contribution. This clearly shows, that in the case of coherent scattering from a dilute system of particles, the inter-particle contribution to the Fourier components $\langle C_{q_1,q_2}^n \rangle_M$ defined at $q_1 \neq q_2$ is negligibly small in contrast to the case $q_1 = q_2$, where this contribution can be very large [4–6].

Results presented here demonstrate the ability of XCCA to obtain information about the oxygen clusters with tetrahedral arrangement of oxygen atoms typical for water. The spectra $\langle C_{q_1,q_2}^n \rangle_M$ averaged for a large number of realizations of a system converge to a specific functional dependence of q, which one could hope to correlate with the local structure of the 3D system composed of many particles. Our

results show that small distortions of the oxygen pentamer do not lead to significant changes of the CCFs. This may help to identify an average local structure of water using experimentally determined CCFs. Our results demonstrate that the spectra $\langle C^n_{q_1,q_2}\rangle_M$ defined at $q_1 \neq q_2$ are more suitable for investigation of the structure of disordered system in the case of coherent scattering, than the spectra determined at $q_1 = q_2$. The Fourier components $\langle C^n_{q_1,q_2}\rangle_M$ can provide the structural information in the case when it is suppressed by a large inter-particle background in $\langle C^n_q\rangle_M$. Our results show that the Fourier components $\langle C^n_{q_1,q_2}\rangle_M$ can reveal structural information about local structure of a system with a larger number of particles, as compared to $\langle C^n_q\rangle_M$.

7. Conclusions

In summary, we showed here how XCCA can be applied to study the local structure of disordered materials. Important steps in this analysis are Fourier decomposition of CCFs and averaging over many realizations of the system. As it was shown in our work, statistical fluctuations of the Fourier components of CCFs, due to this averaging cancel out and information on the local structure of the system can be deduced. The most transparent results can be obtained in the analysis of the dilute 2D systems when contribution from the inter-particle correlations can be neglected. We demonstrated that in this specific case all Fourier coefficients of the angular decomposition of the coherently scattered intensity of a single unit can be completely recovered by the analysis of two- and three-point correlation functions. Importantly, not only amplitudes but also phases of this decomposition can be determined directly from the experimental data. Applying standard phase retrieval techniques to the recovered intensity gives the structure of a single unit in a disordered system.

The XCCA of dense systems shows more complicated features in the Fourier components of CCFs due to the presence of a substantial inter-particle contribution. It can be significantly reduced in some situations when CCFs are analyzed at different resolution rings. Our analysis showed that inter-particle correlations may strongly influence the values of Fourier components of CCFs defined at the same resolution ring. An extraction of a single-particle structure in dense systems by XCCA can be more complicated; however, it may provide information on the medium range order.

The generalization of this approach to 3D systems is still a challenge. Here scattering to high angles can be especially interesting due to the presence of a significant Ewald sphere curvature. Such scattering conditions allow to determine odd Fourier components of CCFs, which are inaccessible in a standard small-angle X-ray scattering experiment, and in this way explore hidden symmetries in the 3D system. This approach can be especially attractive for the study of atomic systems.

We foresee that the general approaches described here for the analysis of the structure of disordered systems based on the angular cross-correlation techniques will find their wide applications in future. They could become especially attractive at a newly emerging free-electron lasers and diffraction limited ultimate storage rings [48].

Acknowledgments

The authors are thankful to E. Weckert for a permanent interest and support of this project and to U. Lorenz and A. Singer for a careful reading of the paper. Part of this work was supported by BMBF Proposal 05K10CHG "Coherent Diffraction Imaging and Scattering of Ultrashort Coherent Pulses with Matter" in the framework of the German-Russian collaboration "Development and Use of Accelerator-Based Photon Sources" and the Virtual Institute VH-VI-403 of the Helmholtz Association.

References

[1] H. E. Stanley and N. Ostrowsky, Eds., *Correlations and Connectivity: Geometric Aspects of Physics, Chemistry and Biology*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1990.

[2] K. L. Ngai, *Relaxation and Diffusion in Complex Systems*, Partially Ordered Systems, Springer, New York, NY, USA, 2011.

[3] P. Wochner, C. Gutt, T. Autenrieth et al., "X-ray cross correlation analysis uncovers hidden local symmetries in disordered matter," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 28, pp. 11511–11514, 2009.

[4] M. Altarelli, R. P. Kurta, and I. A. Vartanyants, "X-ray cross-correlation analysis and local symmetries of disordered systems: General theory," *Physical Review B*, vol. 82, no. 10, Article ID 104207, 2010.

[5] M. Altarelli, R. P. Kurta, and I. A. Vartanyants, "Erratum: X-ray cross-correlation analysis and local symmetries of disordered systems: General theory," *Physical Review B*, vol. 86, no. 17, Article ID 179904, 2012.

[6] R. P. Kurta, M. Altarelli, E. Weckert, and I. A. Vartanyants, "X-ray cross-correlation analysis applied to disordered two-dimensional systems," *Physical Review B*, vol. 85, no. 18, Article ID 184204, 2012.

[7] R. P. Kurta, R. Dronyak, M. Altarelli, E. Weckert, and I. A. Vartanyants, "Solution of the phase problem for coherent scattering from a disordered system of identical particles," *New Journal of Physics*, vol. 15, Article ID 013059, 2013.

[8] D. K. Saldin, H. C. Poon, V. L. Shneerson et al., "Beyond small-angle X-ray scattering: Exploiting angular correlations," *Physical Review B*, vol. 81, no. 17, Article ID 174105, 2010.

[9] D. K. Saldin, H. C. Poon, M. J. Bogan et al., "New light on disordered ensembles: Ab initio structure determination of one particle from scattering fluctuations of many copies," *Physical Review Letters*, vol. 106, no. 11, Article ID 115501, 2011.

[10] R. A. Kirian, "Structure determination through correlated fluctuations in X-ray scattering," *Journal of Physics B*, vol. 45, no. 22, Article ID 223001, 2012.

[11] Z. Kam, "Determination of macromolecular structure in solution by spatial correlation of scattering fluctuations," *Macromolecules*, vol. 10, no. 5, pp. 927–934, 1977.

[12] Z. Kam, "The reconstruction of structure from electron micrographs of randomly oriented particles," *Journal of Theoretical Biology*, vol. 82, no. 1, pp. 15–39, 1980.

[13] W. Ackermann, G. Asova, V. Ayvazyan et al., "Operation of a free-electron laser from the extreme ultraviolet to the water window," *Nature Photonics*, vol. 1, pp. 336–342, 2007.

[14] P. Emma, R. Akre, J. Arthur et al., "First lasing and operation of an ångstrom-wavelength free-electron laser," *Nature Photonics*, vol. 4, pp. 641–647, 2010.

[15] T. Ishikawa, H. Aoyagi, T. Asaka et al., "A compact X-ray free-electron laser emitting in the sub-ångström region," *Nature Photonics*, vol. 6, pp. 540–544, 2012.

[16] M. Altarelli, "The European X-Ray Free-Electron Laser Technical design report," 2007, http://xfel.desy.de/technical_information/tdr/tdr/.

[17] S. R. Elliott, "Medium-range structural order in covalent amorphous solids," *Nature*, vol. 354, no. 6353, pp. 445–452, 1991.

[18] H. W. Sheng, W. K. Luo, F. M. Alamgir, J. M. Bai, and E. Ma, "Atomic packing and short-to-medium-range order in metallic glasses," *Nature*, vol. 439, no. 7075, pp. 419–425, 2006.

[19] M. M. J. Treacy, J. M. Gibson, L. Fan, D. J. Paterson, and I. McNulty, "Fluctuation microscopy: A probe of medium range order," *Reports on Progress in Physics*, vol. 68, no. 12, pp. 2899–2944, 2005.

[20] H. Shintani and H. Tanaka, "Frustration on the way to crystallization in glass," *Nature Physics*, vol. 2, no. 3, pp. 200–206, 2006.

[21] T. Kawasaki, T. Araki, and H. Tanaka, "Correlation between dynamic heterogeneity and medium-range order in two-dimensional glass-forming liquids," *Physical Review Letters*, vol. 99, no. 21, Article ID 215701, 2007.

[22] H. Tanaka, T. Kawasaki, H. Shintani, and K. Watanabe, "Critical-like behaviour of glass-forming liquids," *Nature Materials*, vol. 9, no. 4, pp. 324–331, 2010.

[23] C. A. Tulk, C. J. Benmore, J. Urquidi et al., "Structural studies of several distinct metastable forms of amorphous ice," *Science*, vol. 297, no. 5585, pp. 1320–1323, 2002.

[24] D. D. Klug, "Glassy water," *Science*, vol. 294, no. 5550, pp. 2305–2306, 2001.

[25] A. K. Soper, "Water and ice," *Science*, vol. 297, no. 5585, pp. 1288–1289, 2002.

[26] R. Pindak, D. E. Moncton, S. C. Davey, and J. W. Goodby, "X-ray observation of a stacked hexatic liquid-crystal B phase," *Physical Review Letters*, vol. 46, no. 17, pp. 1135–1138, 1981.

[27] N. A. Clark, B. J. Ackerson, and A. J. Hurd, "Multidetector scattering as a probe of local structure in disordered phases," *Physical Review Letters*, vol. 50, no. 19, pp. 1459–1462, 1983.

[28] P. Pusey and J. Rarity, "Calcul de la variation de mobilité des électrons dans PbTe type n entre 50 et 300 K," *Journal De Physique*, vol. 46, pp. 43–53, 1985.

[29] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, and J. Hajdu, "Potential for biomolecular imaging with femtosecond X-ray pulses," *Nature*, vol. 406, no. 6797, pp. 752–757, 2000.

[30] K. J. Gaffney and H. N. Chapman, "Imaging atomic structure and dynamics with ultrafast X-ray scattering," *Science*, vol. 316, no. 5830, pp. 1444–1448, 2007.

[31] M. M. Seibert, Tomas Ekeberg, R. N. C. Filipe Maia et al., "Single mimivirus particles intercepted and imaged with an X-ray laser," *Nature*, vol. 470, pp. 78–81, 2011.

[32] D. K. Saldin, V. L. Shneerson, R. Fung, and A. Ourmazd, "tructure of isolated biomolecules obtained from ultrashort X-ray pulses: exploiting the symmetry of random orientations," *Journal of Physics*, vol. 21, no. 13, Article ID 134014, 2009.

[33] D. K. Saldin, H. C. Poon, P. Schwander, M. Uddin, and M. Schmidt, "Reconstructing an icosahedral virus from single-particle diffraction experiments," *Optics Express*, vol. 19, no. 18, pp. 17318–17335, 2011.

[34] D. Starodub, A. Aquila, S. Bajt et al., "Single-particle structure determination by correlations of snapshot X-ray diffraction patterns," *Nature Communications*, vol. 3, Article ID 1276, 2012.

[35] R. P. Kurta, Y. Chesnokov, E. Weckert, and I. A. Vartanyants, "Cross-correlation analysis of X-ray scattering from oxygen clusters," *Journal of Physics*. In press.

[36] D. K. Saldin, V. L. Shneerson, M. R. Howells et al., "Structure of a single particle from scattering by many particles randomly oriented about an axis: toward structure solution without crystallization?" *New Journal of Physics*, vol. 12, Article ID 035014, 2010.

[37] A. V. Oppenheim, R. W. Sharfer, and J. R. Buck, *Discrete-Time Signal Processing*, Prentice Hall, New Jersey, NJ, USA, 1999.

[38] J. Goodman, *Speckle Phenomena in Optics: Theory and Applications*, Roberts and Company Publishers, Colorado, Colo, USA, 2007.

[39] V. Elser, "Three-dimensional structure from intensity correlations," *New Journal of Physics*, vol. 13, Article ID 123014, 2011.

[40] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Applied Optics*, vol. 21, no. 15, pp. 2758–2769, 1982.

[41] V. Elser, "Phase retrieval by iterated projections," *Journal of the Optical Society of America A*, vol. 20, no. 1, pp. 40–55, 2003.

[42] R. A. Kirian, K. E. Schmidt, X. Wang, R. B. Doak, and J. C. H. Spence, "Signal, noise, and resolution in correlated fluctuations from snapshot small-angle X-ray scattering," *Physical Review E*, vol. 84, no. 1, Article ID 011921, 2011.

[43] G. E. Walrafen, "Raman Spectral Studies of Water Structure," *Journal of Chemical Physics*, vol. 40, no. 11, Article ID 3249, 1964.

[44] M. Canpolat, F. W. Starr, A. Scala et al., "Local structural heterogeneities in liquid water under pressure," *Chemical Physics Letters*, vol. 294, no. 1–3, pp. 9–12, 1998.

[45] T. Head-Gordon and M. E. Johnson, "Tetrahedral structure or chains for liquid water," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 21, pp. 7973–7977.

[46] R. Ludwig, "Water: from clusters to the bulk," *Angewandte Chemie International Edition*, vol. 40, no. 10, pp. 1808–1827, 2001.

[47] P. Wernet, D. Nordlund, U. Bergmann et al., "The structure of the first coordination shell in liquid water," *Science*, vol. 304, no. 5673, pp. 995–999.

[48] M. Bei, M. Borland, Y. Cai et al., "The potential of an ultimate storage ring for future light sources," *Nuclear Instruments and Methods in Physics Research A*, vol. 622, no. 3, pp. 518–535, 2010.