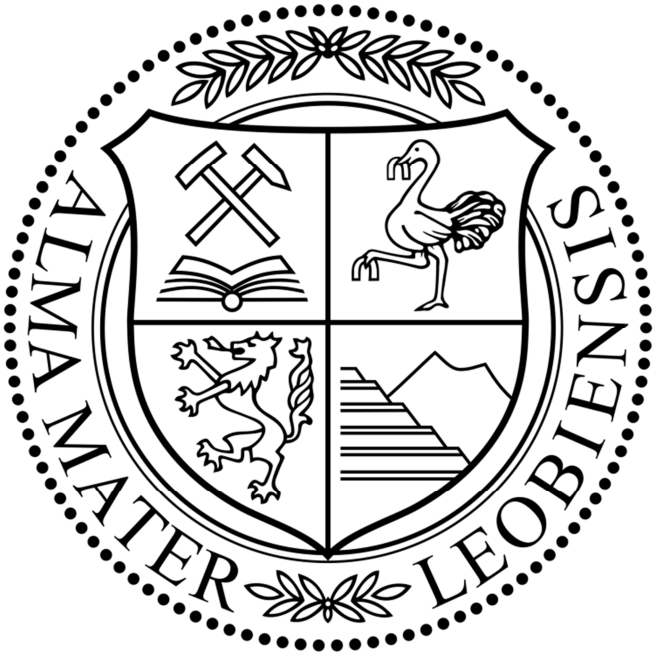
Diploma Thesis

Shape Retrieval of Inorganic Nanocrystals from SAXS-Data

Max Burian

Leoben, June 2014



Montanuniversität Leoben

Institute of Physics

Eidesstattliche Erklärung:

Ich erkläre an Eides statt, dass ich diese Arbeit selbständig verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubten Hilfsmittel bedient habe.

Affidavit:

I declare in lieu of oath, that I wrote this thesis and performed the associated research myself, using literature cited in this volume.

_____ _____

Datum Unterschrift Kandidat

Acknowledgement

First off, I thank my colleagues and friends from the Institute of Physics. I am grateful for the scientific and personal advice I received and the questions that were raised and expected to be answered.

I express my deepest gratitude to my advisor Dr. Rainer Lechner for introducing me to the world of scattering. His endless optimism kept me from throwing the towel numerous times, for which I am grateful - at least in retrospective. He provided me with an excellent atmosphere for doing research and was of great guidance throughout my work and studies.

I thank Prof. Oskar Paris for his highly valued scientific advice and teaching me to keep the eyes on the big picture. Furthermore, I am grateful for the trust that was placed in me from the first day I worked under his supervision.

Additionally, I thank Heide Kirchberger and Magdalena Ottrin for their support in the dullness of everyday life.

I thank my dear friends outside the institute who occasionally kept me from getting lost in reciprocal space - over and over again.

Most important I thank my parents for their unconditional support. Without it, these pages would be empty.

Abstract

The size and shape of inorganic Nanocrystals (NCs) have a fundamental influence on their unique physical properties. Small angle x-ray scattering (SAXS) is an established technique to obtain not only the mean size but also the mean shape of various colloidal NC systems. The *ab-initio* program DAMMIN has previously been developed to retrieve a three dimensional shape that best fits an experimental SAXS curve of monodisperse particles such as proteins. The use of this technique for inorganic NCs yields great potential to obtain the real mean shape in sub-nanometer resolution.

In order to obtain information on the sensitivity of this technique, numerous scattering curves of known theoretical models were computed and used as input for DAMMIN. To directly interpret the obtained three dimensional structures, several evaluation methods were developed and implemented. E.g., one of them reveals, that the averaging of multiple simulation leads to a three dimensional probability map which now can directly be correlated to the electron density in real space.

These developed analysis methods allow to derive quantitative parameters from the retrieved shape. Multiple series of simulations could be performed to reveal the influence of the main fitting parameters. Furthermore the influence of the information content provided from the experimental scattering curve is studied.

Additionally, the simulation was performed on systems with distinct size distributions. Finally an error of the values obtained by the developed methods is estimated that confirms DAMMIN as a powerful method for the shape retrieval of inorganic NCs with a polydispersity up to 10 %.

Kurzfassung

Größe und Form von anorganischen Nanokristallen (NCs) beeinflussen deren einzigartige physikalische Eigenschaften. Zur qualitativen und quantitativen Bestimmung der mittleren NC-Gestalt hat sich Röntgen-Kleinwinkelstreuung (SAXS) als Messmethode etabliert. Eine Möglichkeit der Interpretation von SAXS-Streukurven bietet das Programm DAMMIN. Dieses Programm ermöglicht die Bestimmung der detaillierten Gestalt von monodispersen Systemen, wie z.B. Proteine in Lösung, ohne dabei von einer Grundform auszugehen. Dabei wird ein dreidimensionales Model aus Kugeln erstellt, in dem die Streukurve des Models durch rekursives Abändern der Kugelkonfiguration an die experimentell bestimmte Kurve angepasst wird. Die Anwendung von DAMMIN auf kolloidale anorganische NCs bietet daher neue Möglichkeiten die mittleren strukturellen Eigenschaften der NCs ohne Annahmen zu bestimmen.

Um den Berechnungsprozess auf Sensibilität und Stabilität zu testen wurden theoretisch errechnete Streukurven von simplen Geometrien mit DAMMIN angeglichen. Um die resultierenden Modelle direkt interpretieren zu können wurden im Rahmen dieser Arbeit einige Evaluierungsmethoden entwickelt und implementiert. Dabei kann unter anderem die Besetzungswahrscheinlichkeit der Kugeln, welche durch die Mittelung mehrerer Modelle generiert werden kann, direkt mit der lokalen Elektronendichte der wahren Geometrie korreliert werden. Des Weiteren können mittels der entwickelten Methoden die erhaltenen Modelle quantitativ und qualitativ beschrieben werden, was den direkten Vergleich der Resultate verschiedener Simulationsreihen ermöglicht. Neben dem Einfluss der Hauptsimulationsparameter wurden außerdem die Auswirkungen der Verringerung des Messbereiches und damit des Informationsgehaltes der Streukurve untersucht.

Zusätzlich wurden Systeme mit einer künstlichen Größenverteilung simuliert, um Verschmierungseffekte in die Betrachtung miteinzubeziehen. Schlussendlich wird auf mögliche Fehlinterpretationen eingegangen, sowie die Abweichung bzw. der Fehler der von den Evaluierungsmethoden erhaltenen Werte abgeschätzt. Dadurch wird bestätigt, dass DAMMIN ein wirkungsvolles Programm ist um die Gestalt und Größe von inorganischen NC mit einer Größenverteilung von bis zu 10 % zu bestimmen.

Content

Abbreviations & Symbols

BA …… Boltzmann annealing

DA …… dummy atom

DAM …… dummy atom model

GA …… genetic algorithm

GIFT …… generalized indirect fourier transformation

HCP …… hexagonal closest packing

IFT …… indirect fourier transformation

MC …… Monte Carlo

NC …… nanocrystal

PDDF …… pair distance distribution function

R_G …… radius of gyration

SA …… simulated annealing

SA(X)S …… small angle (X-ray) scattering

SC …… Shannon channel

VDW …… Van-der-Waals

χ^2 …… chi-square

1 Introduction

One of the first definitions of the term *colloid* was stated by *Graham* in the 19th century which referred to the Greek meaning: “glue-like”. To him a colloid was unable to pass through a very fine membrane which introduces an important concept: the dimensions of the species are significant properties of the system. Even though definitions of *colloids* can be very refined, there is no strict distinction to other systems. A common denominator in various definitions is that colloidal materials consist of at least 2 phases where one is dispersed in the other. Therefore, one of the easiest ways to classify colloids is by the physical state of the phases [1]. But as stated, the size of the dispersed phase is a crucial phenomenon, where the colloidal size domain usually extends from a few nanometers to a few micrometers. Thus, colloids can be classified by their size distribution. If the particles of the dispersed phase are identical in size and shape the solution is considered to be perfectly *monodisperse.* If there is a certain size and shape distribution of the particles the solution is considered to be *polydisperse. If m*oreover this distribution occurs around not only one but two mean values, the solution is considered *bimodal* or even *multimodal* [1].

Over the last quarter century, colloidal nanocrystals (NCs) have been in the spotlight of various scientific communities since they show great potential for electronic and biomedical applications [2-5]. The widely used term *nanocrystal* (NC) is rather self-explanatory since it refers to a single- or polycrystalline body of some sort which shows its significant dimension in the nanometer scale [6]. Colloidal NCs are generally built out of an crystaline core that is covered by at least one layer of surfactants [7].

Recent work has shown that the unique physical properties of NCs can be related and influenced by their size and shape [8-11]. This is not only crucial for the understanding of the underlying mechanisms but it offers useful design parameters for the development of new materials. Therefore a lot of attention has been directed towards tailoring the NC shape during the synthesis [12-14].

Generally speaking, the equilibrium shape of a free particle is the shape that minimizes the surface energy, which is defined as the excess free energy per unit area. If we look at a (for example amorphous) material with an isotropic surface energy the resulting shape

will be a perfectly symmetric sphere. Materials with a certain atomic long-range-ordered structure have different surface energies in different directions, due to the regular placement of atoms in the crystal lattice. In other words: different lattice planes have different surface energies [13].

Compared to the isotropic case, the total free surface energy of an anisotropic particle can therefore be minimized by forming distinct planes, which are called *facets*. This principle was proven by Wulff in 1901 and this has established itself as the cornerstone of crystal shape analysis [15]. The essence of this theory is that by knowing the surface free energy of all possible facets of a particle one can predict its free equilibrium shape. Obviously, the number of possible facets of a crystal is infinite which makes it nearly impossible to predict the crystal shape of a given material *ab initio.* For most materials though the smallest free surface energy can be found for low-index crystal planes that show closest atomic packing [13].

With the gaining controllability of the NC-synthesis procedures the level of structural detail is increased which implies that some attention has to be directed towards the resolution of the measurement techniques used to study these systems. The standard method to retrieve the structure of single NCs is transmission electron microscopy (TEM) which yields in high-resolution images with a resolution down to a single Ångstrom. However, using this technique, only two-dimensional images of the focus plane can be obtained. For aligned NCs it is therefore impossible to see structures perpendicular to this focus plane which can lead to the misinterpretation of the real NC shape.

Using small angle x-ray scattering (SAXS) these draw-backs can be avoided. The main advantage of measuring colloidal NC systems using SAXS is the relatively speaking big volume of the sample which is irradiated (containing approx. 10^{10} NCs). This on one hand implies that the information obtained by this method is of high statistical value. E.g., compared to TEM not the shape of a single NC, but the mean shape of the entire system can be obtained. On the other hand due to this high number of randomly oriented particles the scattering occurs equally from all NC-directions. Therefore the three-dimensional mean structure of all NCs can be retrieved.

In the field of protein crystallography, SAXS has been accepted as an established method for the detailed shape retrieval of molecular structures [16]. The program package best known for this application is ATSAS, which was developed by D. Svergun. [17] The

centerpiece of this method is the program called DAMMIN. [18] It creates a three-dimensional probability map in real space that best fits a given SAXS-curve. For the use of the shape retrieval of proteins this map is only an intermediate step in which the known molecular structures are then fitted. [19] Up to this date, only very little effort has been made to directly interpret the probability map obtained from this *ab-initio* method. Furthermore, this method is rarely applied to different inorganic NC systems with narrow size distributions [20].

The objective of this work is the utilization of DAMMIN on theoretical models that are similar to inorganic NC-systems. For this purpose the scattering curves of idealized systems are calculated and used as input for the simulations. In a first step evaluation methods are developed and programmed to quantitatively describe the obtained models. Using the values obtained from these procedures, the stability and reliability of the retrieved model parameters for varying fitting parameters is to be determined. Furthermore the effects of smearing of the initial scattering curve, such as instrumental smearing or polydispersity, on the final models are studied. Finally an estimation of error of the obtained values is made.

This work is divided into seven sections. After this brief introduction, the underlying principles of colloidal nanocrystals, general x-ray scattering and numerical mathematics will be explained. The experimental section will deal with the handling of the software as well as with the computation of the theoretical scattering curves. In the following chapter the developed evaluation methods and the values obtained by them will be presented. Subsequently the results of the various simulation series will be presented and discussed. In the last chapter a summary of this work as well as an outlook for possible further investigations can be found.

2 Theory

2.1 Scattering Theory

The following treatment will be concerned with the scattering of x-rays only. First off, x-rays are photons, transverse electromagnetic waves, with a wavelength in the range of Ångström (10^{-10} m). Using Maxwell's equations it can be shown that an accelerated charge emits electromagnetic waves. This is the working principle of synchrotron research facilities where electrons are circulating in a storage ring and are used as the source of radiation. Maxwells' equations also imply that an electromagnetic wave exerts a force on an electric charge which as a result is accelerated and therefor radiates a new wave. In the classical description this phenomena is considered scattering [21].

2.1.1 General Scattering Theorem

When an electromagnetic wave interacts with matter it can either be absorbed, scattered or reflected. When comparing the incident and the scattered wave, the scattering is considered to be elastic if the energy of both waves is equal to each other. From this point on the presumption is made, that the incident plane-wave is monochromatic, polarized and only elastic scattering will occur.

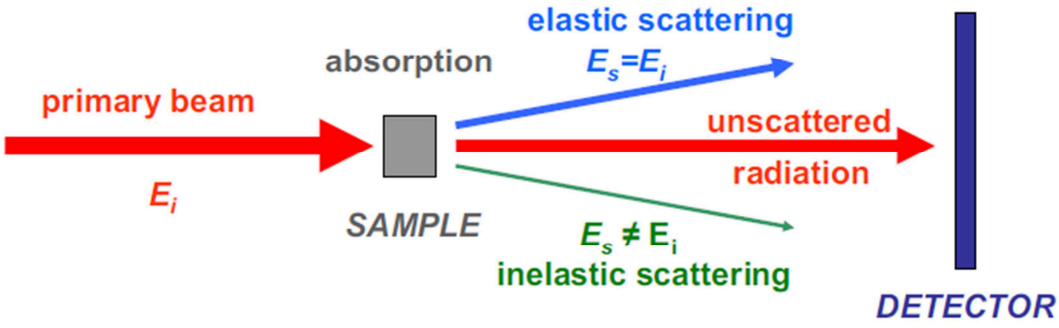


Figure 1: Illustration of the general phenomena of radiation interacting with matter including elastic and inelastic scattering and absorption. Taken from Ref. [22]

The signal measured by the detector in such a scattering experiment is the *differential scattering cross-section* which is defined by

$$\left(\frac{d\sigma}{d\Omega}\right) = \frac{I_{SC}}{\Phi_0 \Delta\Omega} \quad (1)$$

The number of scattered photons recorded per second by the detector is I_{SC} where the detector subtends a solid angle $\Delta\Omega$. The intensity of the incident beam is described by the flux $\Phi_0$which is the number of photons passing through a given area per second [21].

Since the intensity of the beam is correlated to the square of the electric field, the differential scattering cross-section can also be described in terms of the electric field of the incident and scattered wave which results in

$$\left(\frac{d\sigma}{d\Omega}\right) = \frac{|\boldsymbol{E}_{rad}|^2 R^2}{|\boldsymbol{E}_{in}|^2} \quad (2)$$

where R is the sample-detector distance. The incident plane wave is defined as

$$\boldsymbol{E_{in}}(\boldsymbol{r}, t) \equiv \boldsymbol{E_0} e^{-i(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t)} \quad (3)$$

with the electric vector $\boldsymbol{E_0}$ polarized perpendicular to the scattering plane, the propagation vector $\boldsymbol{k}$ and the angular frequency ω. The magnitude of the propagation vector is $\mathrm{k} \equiv |\boldsymbol{k}| = 2\pi/\lambda$ where λ is the wavelength of the wave in the medium [23]. If we evaluate Equation (2) for the case of scattering at a single electron in vacuum, it can be shown that [21]

$$\left(\frac{d\sigma}{d\Omega}\right) = {r_0}^2 \cos^2 \psi \quad (4)$$

where ψ is the angle between the propagation direction of the scattered wave and the observation point. The importance of this equation can be explained using the *Thomson scattering length of an electron*

$$r_0 = \left(\frac{e^2}{4\pi\varepsilon_0 mc^2}\right) = 2.82 \times 10^{-5}\text{Å} \quad (5)$$

which is defined only using natural constants. It is evident that the scattering cross-section of an electromagnetic wave by a free electron is independent of the energy of the incident wave [21].

Since in real experiments one is limited to finite resolution one will always generate a beam of incident waves that has a certain cross section. This means that scattering will occur at more than one charge. We define $\rho(\boldsymbol{r})$ as a number density which describes the charge distribution where $\boldsymbol{r}$ is the position vector of a single charge. It is to be noted, that the charges within this distribution do not interact and can therefore be seen as free charges. Furthermore we need to adopt the so called *Born approximation.* This approximation assumes that all charges interact equally with the incident wave. We neglect that the scattered wave influences neighboring charges and might be scattered again. In short: no multiple scattering occurs. [21]

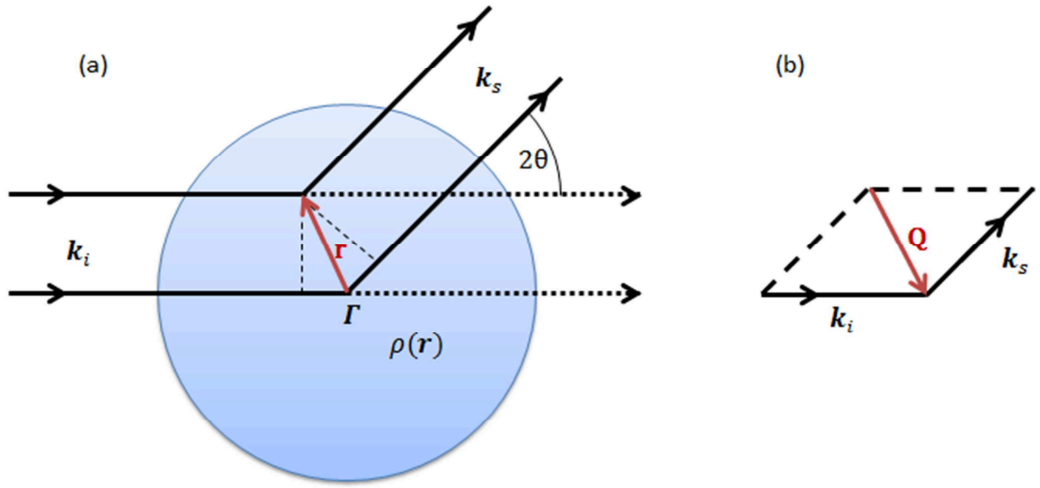


Figure 2: (a) Determination of the phase difference between waves scattered at the origin $\boldsymbol{\Gamma}$ and at $\boldsymbol{R}$ where $\boldsymbol{k}_i$ is the incident and $\boldsymbol{k}_s$ scattered wave. (b) Definition of the scattering vector $\boldsymbol{Q}$.

An example would be electrons in an electron-cloud. The entire scattered radiation can be described as the sum of superposition contributions from different volume elements of $\rho(\boldsymbol{r})$. To evaluate the scattered signal we need to look at the phase difference between the wave scattered at the origin and the one scattered at the position $\boldsymbol{r}$, as shown in Figure 2. The phase difference $\delta(\boldsymbol{r})$ is thus [21]

$$\delta(\boldsymbol{r}) = (\boldsymbol{k}_i - \boldsymbol{k}_s) \cdot \boldsymbol{r} = \boldsymbol{Q} \cdot \boldsymbol{r} \qquad (6)$$

$\boldsymbol{Q}$ is called the scattering vector. For elastic scattering, where $|\boldsymbol{k}_i| = |\boldsymbol{k}_s|$, with the scattering angle 2θ, Q is defined as [21]

$$Q = |\boldsymbol{Q}| = \frac{4\pi}{\lambda} \sin\theta \qquad (7)$$

where λ is the wavelength of the incident beam.

The elastic scattering amplitude of an assembly of N electrons can be written as [21]

$$A(\boldsymbol{Q}) = r_0 \sum_j^N e^{-i\boldsymbol{Q}\cdot\boldsymbol{r}_j} \quad (8)$$

If you look at the electron-cloud surrounding an atom the electrons can be seen as delocalized which results in a discrete density $\rho(\boldsymbol{r})$ describing the spherical vicinity of the nucleus. The scattering amplitude is then described as [21]

$$A(\boldsymbol{Q}) = \int \rho(\boldsymbol{r}) e^{-i\boldsymbol{Q}\cdot\boldsymbol{r}} d\boldsymbol{r} \quad (9)$$

It is to be noted that after all scattering occurs originates from changes in the electron density in a given sample. Therefore one needs to compare the electron density of the scatterer to the overall mean of the measured density. We will relate to this as $\Delta\rho(\boldsymbol{r})$.

As mentioned above, the measured intensity is equal to the squared amplitude of the wave. For the approximation that the electron density surrounding an atomic nuclei is homogeneous the intensity can be calculated using the amplitude $A(\boldsymbol{Q})$ and its *complex conjugate* $A(\boldsymbol{Q})^*$ as [21]

$$I(\boldsymbol{Q}) = A(\boldsymbol{Q}) * A(\boldsymbol{Q})^* = \left| \Delta\rho \int e^{-i\boldsymbol{Q}\cdot\boldsymbol{r}} d\boldsymbol{r} \right|^2 \quad (10)$$

2.1.2 The Debye-Formula

In the case of two electrons the intensity as a function of $\boldsymbol{Q}$ can be calculated using the following equation [21]

$$I(\boldsymbol{Q}) = 2r_0(1 + \cos(\boldsymbol{Q}\cdot\boldsymbol{r})) \quad (11)$$

From this equation it is obvious that in order to evaluate the scattering intensity of more than one charge an angle between $\boldsymbol{Q}$ and $\boldsymbol{r}$ has to be defined. In real systems $\boldsymbol{r}$ is often randomly oriented with respect to $\boldsymbol{Q}$. To account for this, a spherical average over all angles has to be performed which was done by P.Debye [24]. In a first step we define $\boldsymbol{\theta}$ as the angle between $\boldsymbol{Q}$ and $\boldsymbol{r}$ such that

$$\boldsymbol{Q}\cdot\boldsymbol{r} = Q\, r\, \cos\boldsymbol{\theta} \quad (12)$$

If we look at a system of two particles, one at the origin and one at the position $\boldsymbol{r}$, with the scattering amplitudes f_1 and f_2 the scattering amplitude can be written using Equation (8) as

$$A(\boldsymbol{Q}) = f_1 + f_2 e^{-i\boldsymbol{Q}\cdot\boldsymbol{r}} \quad (13)$$

If we now calculate the intensity using Equation (10) we get

$$\mathrm{I(Q)} = \mathrm{f_1}^2 + \mathrm{f_2}^2 + \mathrm{f_1 f_2 e^{iQ\cdot r}} + \mathrm{f_1 f_2 e^{-iQ\cdot r}} \quad (14)$$

For a fixed length $\boldsymbol{r}$ but a randomly oriented direction, the last two terms can then be combined such that

$$\langle \mathrm{I(Q)} \rangle = \mathrm{f_1}^2 + \mathrm{f_2}^2 + 2\mathrm{f_1 f_2}\langle \mathrm{e^{iQ\cdot r}} \rangle \quad (15)$$

Using Equation (12) above, the average of the phase factor can be written as

$$\langle e^{i\boldsymbol{Q}\cdot\boldsymbol{r}} \rangle = \frac{1}{4\pi} \iint e^{iQ\,r\,\cos\boldsymbol{\theta}} \sin\boldsymbol{\theta}\, \mathrm{d}\boldsymbol{\theta}\, d\varphi = \frac{\sin qr}{qr} \quad (16)$$

If we generalize the so called *Debye formalism* above to a system built up of N charges with the distance between every possible pair of $r_{jk} = |\boldsymbol{r}_k - \boldsymbol{r}_j|$ and the scattering amplitude of each charge described by f_i, one can calculate the scattering intensity of this system using

$$\langle I(Q) \rangle = \sum_{j=0}^{N} \sum_{k \neq j}^{N} f_j f_k \frac{\sin qr_{jk}}{qr_{jk}} \quad (17)$$

which is known as the *Debye-Formula* [21, 24].

2.1.3 Small Angle Scattering

The importance of small angle scattering (SAS), or in this case small angle x-ray scattering (SAXS), becomes obvious when looking at the reciprocal law of scattering processes. The classical approach to describe this is done by the *Bragg-equation*, which states the relationship between a characteristic geometry described by $\boldsymbol{d}$ and the scattering angle θ at a fixed wavelength as [21]

$$n\,\lambda = 2d \sin\theta \quad (18)$$

or by using Equation (7)

$$\frac{2\pi}{d} = Q \qquad (19)$$

As we want to look at dimensions that are bigger than 1 nm in size, the corresponding scattering angle, using a Cu-Kα source with λ = 0.1548 nm [21], decreases to values lower then ca. 4.5°. The information on e.g. the NC shape can therefore be accessed at the small angle-regime.

For dilute systems with N particles, each identical in size and shape with the volume of V_p, the scattering intensity can be written as [23]

$$I(Q) = N * V_p^2 * |F(\boldsymbol{Q})|^2 * S(Q) \qquad (20)$$

The term $F(\boldsymbol{Q})$ is called the formfactor of a single particle and is only depended on the mean shape of the particle. It is important to notice that as $Q \to 0, |F(\boldsymbol{Q})|^2 = 1$. The term $S(Q)$ is referred to as the *structure factor* and it describes the interaction between particles. For a dilute system, where the particles are on average widely spread from each other, no interaction occurs and so $S(Q) = 1$. The formfactor can then be calculated using the following equation[1] [21]

$$F(Q) = \frac{1}{V_p}\int \Delta\rho(\boldsymbol{r})\, e^{i\boldsymbol{Q}\cdot\boldsymbol{r}} dV_p \qquad (21)$$

For a given particle shape, one can try to calculate the formfactor analytically. This can easily be done for simple geometries by solving Equation 21 and perform a spherical averaging. For example, the formfactor of a sphere can be calculated as

$$F_{Sphere}(Q,r) = 3\,\Delta\rho(\boldsymbol{r})\left[\frac{\sin(Qr) - QRcos(Qr)}{(Qr)^3}\right] = \frac{3J_1(Qr)}{Qr}\Delta\rho(\boldsymbol{r}) \qquad (22)$$

where $J_1(x)$ is the spherical Bessel function of first kind.

A very powerful tool to gain information on the mean particle size is the *Guinier-analysis*. It can be shown, that at long wavelengths ($Qr \to \infty$) the scattering intensity of an arbitrary particle can be written as [21, 25]

$$I^{SAXS}(Q) \approx \Delta\rho^2 V^2 e^{-\frac{Q^2 {R_g}^2}{3}} \qquad (23)$$

[1] In the small angle scattering literature the formfactor is often defined as $P(Q) = |F(Q)|^2$. In other words it refers to the intensity and not the amplitude.

When the intensity is plotted on a logarithmic scale versus Q^2 in the regime of $Qr \ll 1$it will produce a line with the slope of ${R_g}^2/3$, where R_g is the radius of gyration (ROG) of the particle. This is a very fast and easy way to get a first estimate of the particle size and shape. For examples for spheres the radius can then be calculated by $R_s = \sqrt{5/3}R_g$. [21]

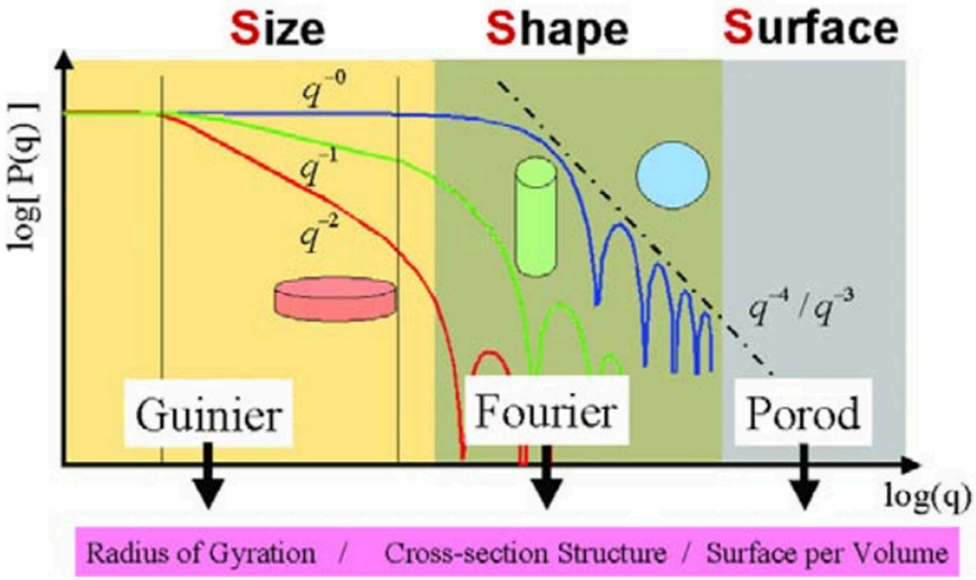


Figure 3: Overview of the different regimes of a particle-form-factor measured by SAXS. Taken from Ref. [26] © Anton Paar GmbH

Another very sensitive method to gain information on the particle shape is to look at the regime of $Qr \gg 1$ with the constraint that Q is still smaller than the inner-atomic spacings. This so called *Porod-analysis* is very sensitive to the relation between particle volume and particle surface. For the example of spheres with a surface S_p the scattering Intensity can be written as [21]

$$I^{SAXS}(Q) = \frac{2\pi\Delta\rho^2}{Q^4}S_p \qquad (24)$$

This is a commonly used criterion to ensure adequate background subtraction. In SAXS measurements of colloidal dispersion the sample (dispersed particles in a solvent) as well as the pure solvent have to be measured. Part of the data analysis is the correct subtraction of the background (scattering by the solvent) from the sample. To ensure that this is done properly, the slope of the corrected sample in the *Porod-Regime* Is compared to the ideal slope which is q^{-4} for point- and q^{-3} for slit-collimated systems [21, 26].

An overview on the methods mentioned above and which Qr regime yields what information can be found in Figure 3.

2.1.4 Modeling of SAXS-Data

For more complex geometries either Equation (21) or the spherical averaging cannot be performed analytically. The determination of the formfactor is therefore limited to numerical methods. In this case, one can consider the particle being built up out of N scatteres at the positions $\boldsymbol{r_i}$ and use a modification of the Debye-Formula. The intensity can then be calculated using the formfactor $F_i(r)$ of each scatterer as [21]

$$I(Q) = \sum_{j=0}^{N} \sum_{k \neq j}^{N} F_j F_k \frac{\sin qr_{jk}}{qr_{jk}} \qquad (25)$$

This is a very powerful method to compute the scattering from complex geometries since every structure can be reduced to an assembly of geometrically simpler ones. The difficulty with this approach is that it only works in one way: one can easily compute the scattering intensity in reciprocal space of a known geometry in real space but the calculation of the electron density from a known scattering intensity cannot be done. In the scattering-society this is related to as the *phase problem* [21].

Another approach to the problem of complex geometries was done by O.Glatter, who introduced the so called *Pair-Distance-Distribution-Function* (PDDF) $p(r)$ such that [27]

$$I(Q) = 4\pi \int_0^\infty p(r) \frac{\sin(Qr)}{Qr} dr \qquad (26)$$

with

$$p(r) = \Delta\tilde{\rho}^2(r) \cdot r^2 = \int \Delta\rho(\boldsymbol{r}_1)\Delta\rho(\boldsymbol{r}_1 - \boldsymbol{r}) d\boldsymbol{r}_1 \cdot r^2 \qquad (27)$$

where $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$ is the vector describing the distance between two radial points. From Equation (26) it can be seen that the scattering intensity is the *Fourier-transform* of the PDDF. Since this transformation works both ways, we can now easily calculate the PDDF from a measured intensity using

$$p(r) = \frac{1}{2\pi^2} \int I(Q) \cdot Qr \cdot \sin(Qr) \cdot dQ \qquad (28)$$

This makes the PDDF a very powerful method to gain information on the particle size and shape. We can summarize the findings above: the scattering amplitude $F(Q)$from a

particle depends on the electron density difference $\Delta\rho(\boldsymbol{r})$ and they are Fourier transformation pairs. The scattering intensity $I(Q)$ from a particle is the Fourier transform of the PDDF averaged over all directions in space [24].

2.2 Computational Methods

Up to this point, the presumption was made, that the electron density of a given system is known from which one can calculate the theoretical intensity using the equations above. In reality though one measures the scattering intensity and has to calculate the electron density from this. To retrieve the shape of a particle ensemble dispersed in solution, the orientational averaged intensity of all particles is measured. Due to this a significant amount of information is lost. The restoration of the particle shape is therefore an iterative process where in a first step the model intensity I^{mod} from a presumed initial model is calculated. In a second step the model intensity is compared to the measured intensity I^{exp} using the *chi-squared functional* [23]

$$\chi^2 \equiv \sum_{i=1}^{N} \left(\frac{I^{exp}(q_i) - I^{mod}(q_i)}{\sigma_i} \right)^2 \qquad (29)$$

with N as the number of measured data points and σ_i the uncertainty of the data-point q_i. The objective is to minimize χ^2 using numerical methods.

2.2.1 Simulated Annealing

The method of *Simulated Annealing* (SA) was introduced by S. Kirkpatrick and M. P.Vecchi in 1983. The idea behind SA can be compared to "[...] to the behavior of systems with many degrees of freedom in thermal equilibrium at a finite temperature." [28]. As a supposition one has to be able to calculate the "energy" E of a system with a given configuration C. The objective of SA is to change the configuration such that the energy finds a global minimum [29]. This is done by introducing $P(E,T)$, the *probability for acceptance* of a configuration, which depends not only on the energy of the configuration but also on the *temperature T of the system*. At this point it should be noted that the temperature of a system is only metaphor for a scalar which has to be adjusted to fit the energy range of given configurations as well as the unit of energy. There are several ways to calculate $P(E,T)$. One of the most common methods is to use the Boltzmann-distribution such that

$$P \sim e^{-\frac{E}{T}} \quad (30)$$

If Eq. 2.26 is used, the annealing process is called *Boltzmann annealing.* If one calculates P at very high temperatures, such that $E \ll T$, the probability of acceptance is equal for all configurations. By slowly decreasing the temperature according to a *temperature schedule* the energy of a configuration gains more importance. The advantage of the cooling procedure is that the system can jump out of local energy minima to find a better global solution. It was mathematically proven that for an infinite time the global minima is always found [28].

2.2.2 The Multipole Expansion

Since the SAXS function of a dispersed particles doesn't depend on the particle orientation, Stuhrmann used a mathematical representation of the scatterer $\rho(\boldsymbol{r})$ with rotational properties, namely *spherical harmonics* Y such that the scatterer in real space can be described as [30]

$$\rho(\boldsymbol{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \rho_{lm}(r) Y_{lm}(\theta, \varphi) \quad (31)$$

where the coefficients

$$\rho_{lm}(r) = \int Y_{lm}^{*}(\omega)\, \rho(\boldsymbol{r}) d\omega \,.\, d\omega = sin\theta \; d\theta \; d\varphi \quad (32)$$

are the *multipole components*. The *spherical harmonic conjugate* Y_{lm}^{*} is only dependent on the angle ω from a given direction and not on the distance from the center. With this knowledge it becomes clear from the equation above that a single multipole $\rho_{lm}(\boldsymbol{r})$ only describes the part of the scatterer for which the spherical harmonic is not zero. The scattering amplitude of a multipole can then be written as [30]

$$A_{lm}(r) = \sqrt{\frac{2}{\pi}} i^{l} \int \rho_{lm}(r)\; j_{l}(Qr) r^{2}\, dr \quad (33)$$

with j_l describing the *Bessel function of l-th kind*. The total scattering amplitude can then be calculated as

$$I(q) = \frac{1}{4\pi} \sum_{l=0}^{L} \sum_{m=-l}^{l} A_{lm}(r) A_{lm}^{*}(r) \quad (34)$$

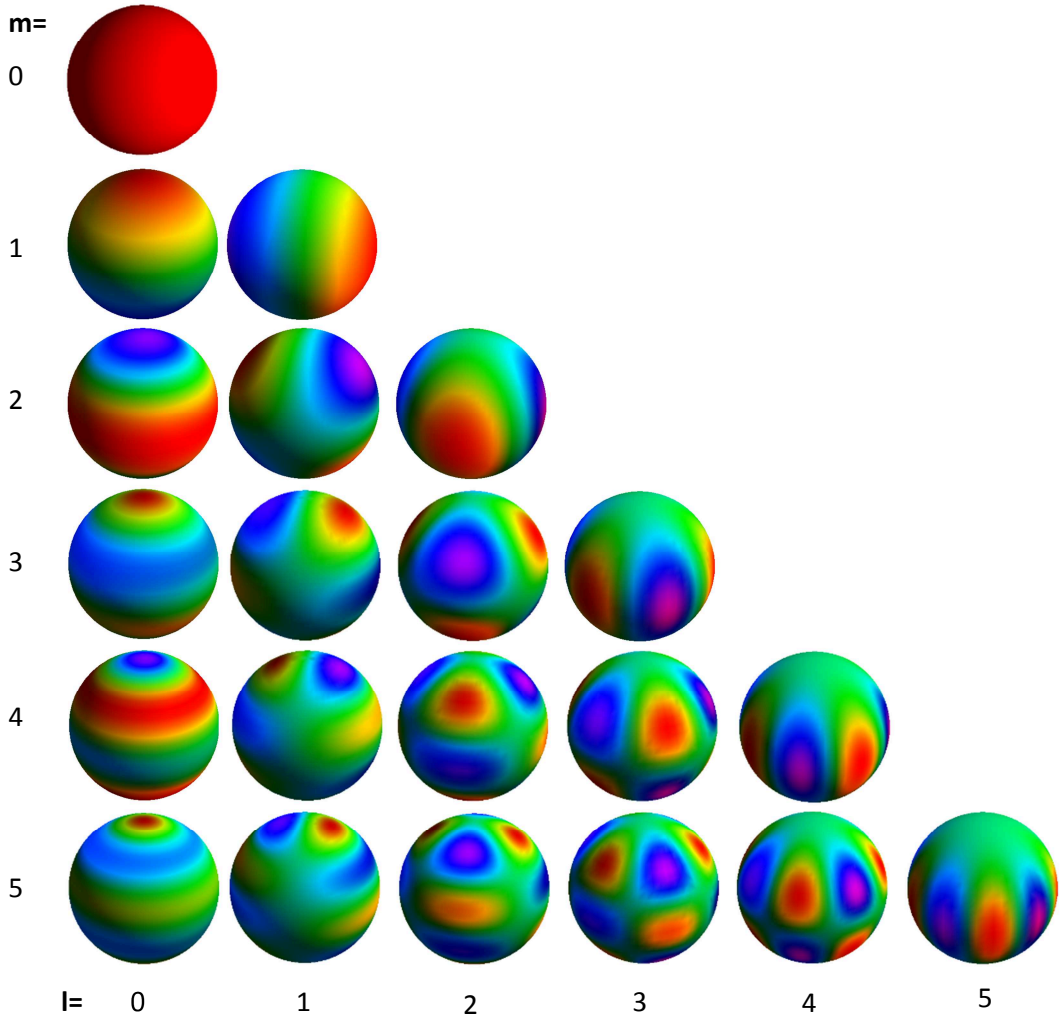


Figure 4: Visualization of the first 20 symmetrical spherical harmonics. The different colors relate to the different angular regions or also to the different *multipole components* (see Equation (32)). The green regions are not accessible angular region. DAs within these regions therefore do not contribute to the scattering intensity.

The first obvious improvement of Equation (34) compared to the Debye formula (Equation (25)) is that the two sums only have to be made over the number of spherical harmonics L and not over all particles N in a theoretical model, which safes a significant amount of calculations.

Furthermore, if one is in the iterative process of trying to fit a theoretical scattering to an experimentally measured intensity and therefore keeps changing the model, only the multipole of the altered section has to be recalculated. This speeds up the fitting process by orders of magnitude [31].

Obviously the *Stuhrmann approximation* is only an approximation. As seen in Equation (31) the scatterer can only fully be described using multipole components when the sum is performed if $l \to \infty$. In reality this is not feasible. If we compare this fact to Equation (34) we notice, that our summation limit is defined by the number of *spherical harmonics* L. This implies that L can be seen as a degree of convergence of the approximation.

Even though the equations above seem very complex, the initial idea is simple: the entire shape can be seen as split up into angular regions. A visualization of these regions can be found in Figure 4. If now a change in one of these regions is made to the model, only the corresponding *multipole* (see Equation (32)) has to be recalculated. One can therefore directly see the influence of a certain angular region $d\omega$ of the theoretical model on the calculated scattering curve.

2.2.3 Shape Retrieval Methods

In general the problem of shape retrieval can be solved using two different approaches. The one way is to retrieve data directly from the measured intensity. This idea was pursued by Glatter et al., who simply suggested a fourier-transformation as mentioned in Chapter *2.1.4 Modeling of SAXS-Data,* which results in the PDDF. Since the information gained in real experiments is limited by the obtainable q-range and smearing effects, this has to be done numerically. This was done by. Glatter who introduced the *indirect fourier transform* (IFT). [32, 33] It is to be noted that the IFT is a model free approach to extract information from the measurement data. In the following years more refined methods for increasingly complex systems were published that allowed, e.g. the determination of radial electron density profiles for inhomogeneous particles [34, 35].

The other approach to deduct information from a general scattering experiment needs some kind of a theoretical particle-model in the background for which an ideal scattering curve can be computed. This can be done using geometries with a known formfactor (see chapter *2.1.4 Modeling of SAXS-Data*). By optimizing the model parameters, χ^2 is minimized (see Equation (29)) until a final fit is obtained. Compared to the IFT, this approach requires some kind of *a priori* knowledge of the system, which in most cases is the case. This allows solving much more complex models. For example a big step towards the characterization of concentrated systems was made using the

generalized indirect fourier transform (GIFT) which assumes a given particle shape and as a result allows to determine the structure factor of the system. [27, 36-38]

From a computational point of view, these two methods are very distinct from their solving algorithms but are often used supplementary or even cross-linked. For this reason most software packages, such as SASfit [39], IRENA (IgorPro Plugin) [40] and GIFT [41], allow the computation of the PDDF as well as the fitting of geometrically simple models.

As the particle shape and structure gets more complex the scattering intensity has to be computed using the Debye-Formula (Equation (25)) or one of its approximations (e.g. Stuhrmann). This has been utilized for models that are built out of small spheres in software-programs such as foXs [42] or CRYSOL. [43] Another tool, PCG Singlebody [44], generated random points in a geometrically confined space and then calculates the scattering intensity as well as the PDDF. These methods are computationally a lot more time consuming. The shape retrieval from SAXS data before the 1990s was limited to the use of analytically solvable form factors. The development of shape retrieval methods is therefore closely linked to advances in computing technology.

During the 1980s the idea came up to use standard Monte Carlo (MC) methods to iteratively change the actual particle shape and minimize χ^2. These models are built up from N so called spherical *dummy atoms* (DA) of the same radius, described by a position vector, and are often referred to as *dummy atom models* (DAM). This idea was further developed by Chacón who used *genetic algorithms* (GA) to speed up the remodeling process [45]. Only one year later a similar software was published by Walther (Saxs3D) which was based on “Monte Carlo type reconstruction algorithm” [46]. The first successful implementation using the multipole expansion and simulated annealing mentioned above was done by Svergun in the same year with his software called *DAMMIN* [18]. In the following years a lot of similar programs have been released from Svergun’s group, mainly focusing on chainlike structures for biomolecular use such as MONSA [18] and GASBOR [47].

As this modeling software was, and still is, mainly used for the shape retrieval of organic structures, such as proteins, the theory behind this is based on the assumption that scattering occurs of particles that are all perfectly alike in shape and size.

3 Experimental

3.1 Creation of the Reference Formfactors

Since all available modeling software [45, 46, 48, 49] is based on numerical methods, at least some attention has to be directed towards the convergence of the results. In other words: a numerical solution will always to some degree deviate from the analytical result. In the case of *ab initio* shape determination this implies that the simulation results have to be compared to geometries for which the formfactor can be solved analytically. Even for monodisperse systems, this constraint significantly limits the choice of possible reference geometries.

3.1.1 Monodisperse Systems

An extensive overview of analytically solvable formfactors for systems of diluted particles is given by Pedersen [50]. At this point it should be noted, that the term "*analytically solvable"* has to be used with caution. Even though the formfactor amplitude itself is solvable for several geometries one still has to calculate the square of the formfactor to get the scattering intensity, according to Equation (20). This step can only be performed analytically for radial-symmetric shapes such as spheres, concentric shells, etc. For all other geometries that show some kind of shape or electron-density anisotropy the spherical average of the formfactor has to be made numerically. With todays' computational resources this can be done with a very high degree of numerical convergence [51]. We will therefore neglect errors resulting from this step.

For the systems used we assume that the particles are randomly oriented and strictly monodisperse [23]. For our purpose we focused on the geometries for which at least the formfactor-amplitude is analytically solvable: namely spheres, ellipsoids of revolution, cubes and core-shell spheres. The scattering curves were computed using the software GIFT [52], based on [36], which includes the numerical orientational average. The used formfactors can be found at [50]. The computed intensities using these form-factors as well as the three idealized dimensional models can be seen in Figure 5.

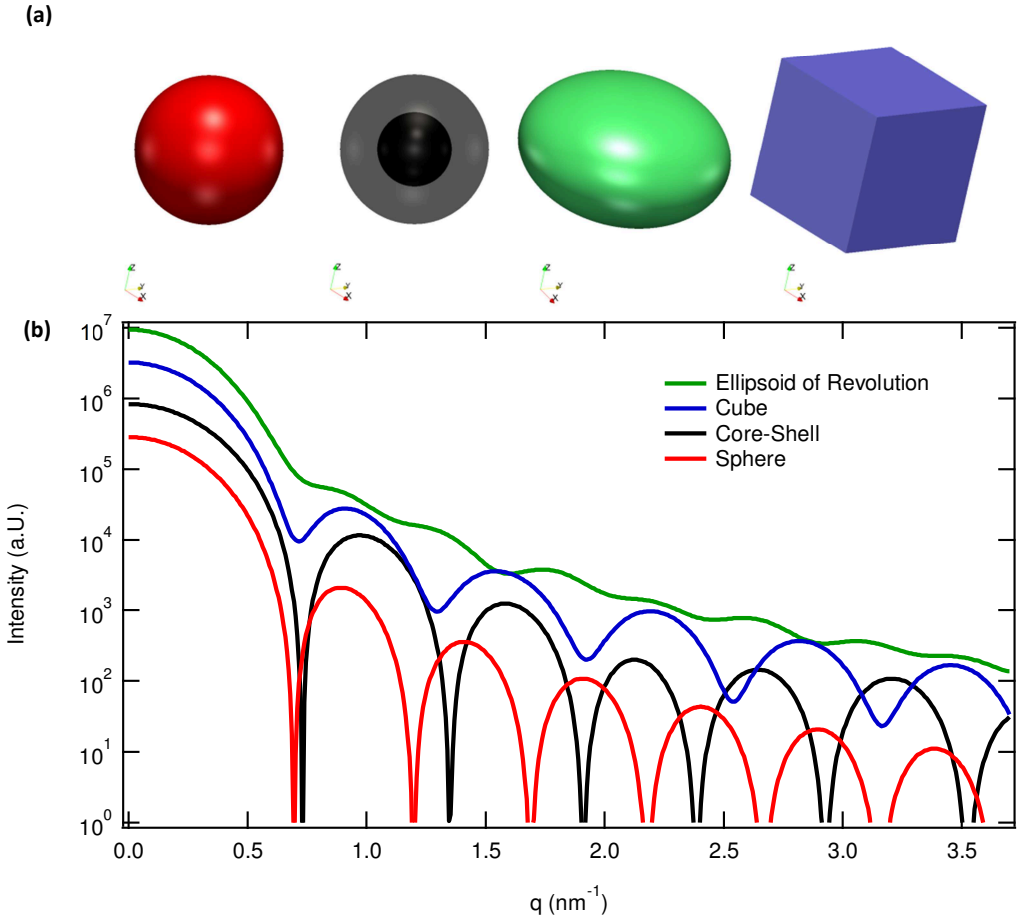


Figure 5: (a) Representations of the models used in the thesis. **(b)** Scattering intensities of the formfactors corresponding to the models shown above. The model dimensions were chosen, such that their radius of gyration is approx. 5nm, and can be found inTable 1. The visible q-range ensures at least 15 SCs.

Table 1: Dimensions of the models used in this thesis. The corresponding three-dimensional representations can be found in Figure 5(a)

model	**dimension**
sphere	$r = 7.5\ nm$
core-shell sphere	$r_{core} = 5\ nm, t_{shell} = 2.5\ nm$
ellipsoid of revolution	$a = b = 8.5\ nm, c = 5.25\ nm$
cube	$a = b = c = 16\ nm$

The dimensions of the reference geometries were chosen such that a q-range of most laboratory sources is adequate for this method. This links to a relevant question regarding the information content in a scattering curve. A general approach that describes the information that is carried by an arbitrary wave was introduced by

Shannon [53]. This lead to the introduction of so called *Shannon-channels* (SC) as a quantitative measure of how much information in a scattering wave can be found when looking only at a section of its Fourier-analysis. For SAXS the width of such a SC is defined as [54]

$$w = \frac{\pi}{D_{Max}} \qquad (32)$$

where D_{Max} represents the maximum dimension of the particle (e.g. the diagonal of cube, diameter of a sphere, etc.).

The difficulty of this topic becomes obvious, when paying attention to previous investigations. In his first studies on *uniqueness of results* [55] Svergun estimates that 15-20 SCs are necessary to ensure adequate information content for shape retrieval. In the same year he also published that the amount of SCs does not limit the number of parameters that can be retrieved from a scattering curve. One also has to take *a priori* knowledge of the measured system into account [56]. Only in 2013 a "Task Force on Small Angle Scattering" of the biomolecular community agreed on the following minimum requirements for shape retrieval from SA-data: "The minimum q value must be smaller than the first Shannon channel […] and it is suggested that four to five Shannon channels are covered [..]."[57]

To avoid misinterpretation all reference-data used in this thesis was computed using at least 15 SCs.

3.1.2 Polydispersity

First it is to be noted that in this case polydispersity only affects the size of given particles and not the shape.

Already in the early days of SAS the theoretical proof was made that it is not possible to deduce a particle shape and simultaneously the size distribution of the system from SAS data. [58] For example, the scattering behavior of an ellipsoid can be described using spheres with a very specific size-distribution. However, this size distribution is unique and degraded to such an extent, that no real system of spheres will have such a distribution.

This implies that shape retrieval of particles can only be done, if some kind of *a priori* knowledge on the size distribution is available. If, e.g., the assumption is made that the

distribution is monomodal, or in other words that it occurs around one maximum, both the shape and the polydispersity can now be retrieved.

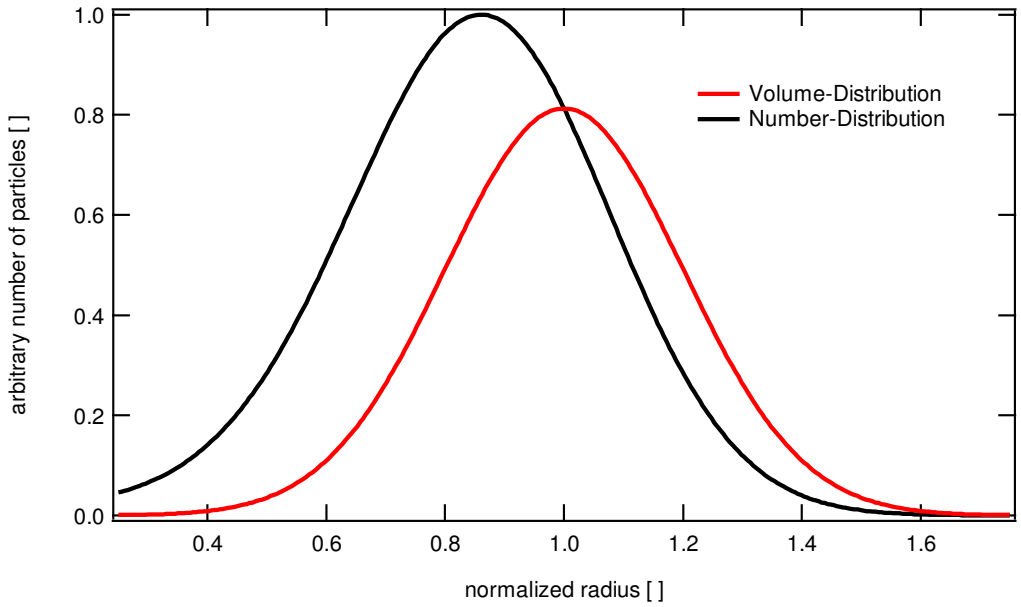


Figure 6: Gaussian volume-distribution of spheres (red) and its correlating number-distribution (black), normalized on the mean sphere radius. The particles of both curves occupy the same volume. A distortion of the number-distribution due to the volumetric effects can be seen.

Due to the Born-approximation, only single scattering occurs, which allows the superposition of waves that are scattered from single electrons and therefore from single particles. If we now look at a system built of two particles of different size, we can simply calculate the total scattering intensity using Equation (20) for each particle and adding up the resulting intensities. From this equation it follows that the scattering intensity is dependent on the square of the volume, and in the case of spheres, is therefore dependent on the 6^{th} power of the radius. As a result, a Gaussian number-distribution, would lead to an asymmetric Gaussian volume-distribution (3^{rd} power of the radius) and an even more asymmetric intensity-distribution. This effect can be observed in Figure 6, in which a Gaussian volume-distribution of spherical particles and its correlating number distribution are plotted.

The total scattering intensities of polydisperse systems used in this thesis were calculated using a Gaussian volume distribution. The distribution is represented using at least 11 points. At each point the corresponding formfactor was multiplied by the squared particle-volume and the arbitrary number of particles with this size. The resulting effect of polydispersity on a system of spherical particles can be seen in

Figure 7. The most significant effect is the well-known smearing of the minima. Also note the shift to smaller q-values with increasing polydispersity, which correlates to the fact mentioned above that larger particles scatter more than smaller ones.

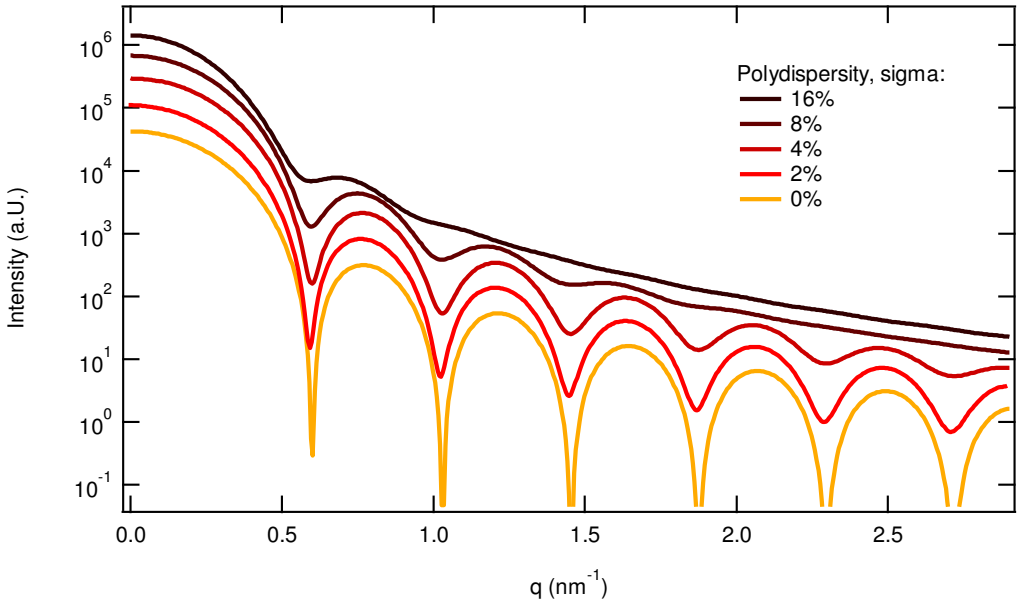


Figure 7: The effect of polydispersity on the scattering-intensity. For this illustration spheres of radius 7.5 nm were used. The sigma value corresponds to the *standard deviation* of the Gauss-distribution.

3.1.3 Smearing Effects

In every real scattering experiment the subject of *smearing* has to be addressed. In general, the resolution of measured data can be correlated to two factors. First, every detector has some kind of a sensor grid to measure the incoming flux at a certain position. This grid is divided into equally sized quadratic pixels where each has a certain size. Due to this finite pixel size one gets a different angular resolution for every sample to detector distance. Furthermore, the finite beam divergence can lead to a spread of the incident x-rays on the scanned physical pixels. The other factor is contributed to the fact that the beam focus is not perfectly spherical. Especially for asymmetrical beamshapes such as a e.g. rectangular slit profile strong smearing effects can occur [59, 60]. Additionally, experimental smearing is measured as a two dimensional detector signal. The actual effect on the one dimensional scattering curve is therefore strongly dependent on the averaging process of the detector data.

For these reasons, the actual smearing is usually measured for a given experimental setup. This can be done by measuring a periodic lattice with a lattice spacing that

produces sharp peaks in the SAS region. Such a system can be found in either turkey- or rattail-tendons [61, 62]. The resolution of our laboratory SAXS-setup is determined using rattail-tendons. The results obtained can be found in Table 2.

Table 2: Resolution-functions measured at the in-house NanoStar System by Bruker. The system contains a point-collimated beam with a 600 µm aperture and Cu-Kα source with λ=1.5418 Å. The peak-width was obtained by a Gauss-fit using a linear background subtraction.

sample-detector-distance $l \approx$	**peak-width (FWHM): $\Delta q =$**
261 mm	0.749 nm^{-1}
638 mm	0.075 nm^{-1}
1008 mm	0.050 nm^{-1}

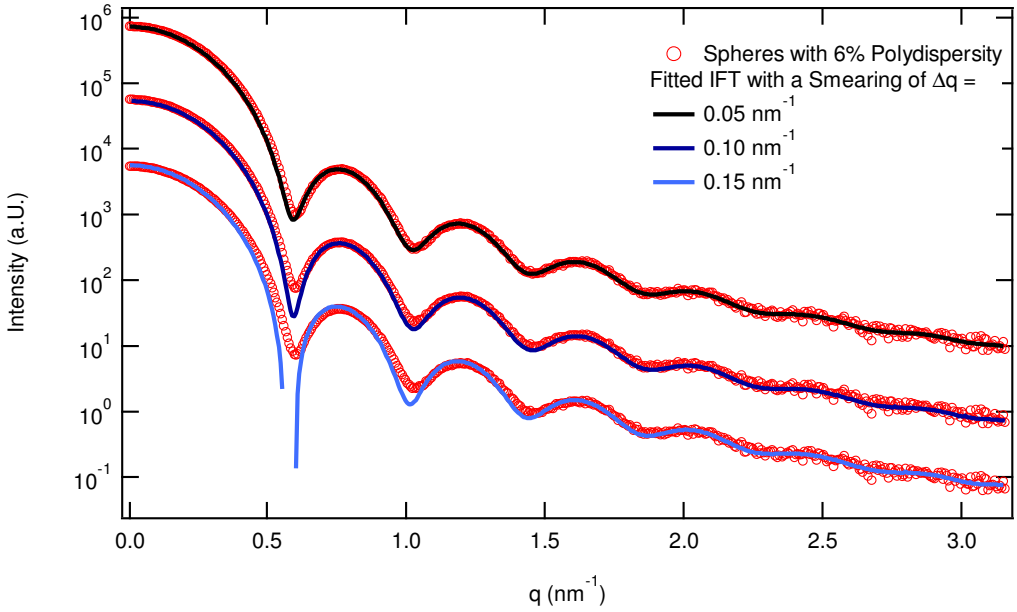


Figure 8: Illustration of smearing effects for the theoretical case of a measurement performed at our in-house NanoStar System at a sample-detector distance of 638 mm. The model-data was generated using GIFT [41]. The dimensions of the model were chosen to result in 10 SCs.

Neglecting smearing effects can lead to introduction of systematic errors. To avoid this, Glatter showed that applying the IFT on smeared data according to a certain sequence will result in a *desmeared* scattering curve [23]. Nevertheless, the recovery of *the real scattering curve* from smeared data can still lead to misinterpretation. To illustrate this, the scattering intensity of spheres with 6 % polydispersity was *desmeared* using values close to the ones obtained from our measurements (see Table 2). As it can be seen in Figure 8 a polydispersity can be misinterpreted as smearing effects since both, poyldispersity and smearing effects, causes a broadening of the minima. Thus, for

correct data analysis of real SAXS experiments the instrumental resolution function has to be determined and has to be taken into account for the shape retrieval.

3.2 GNOM

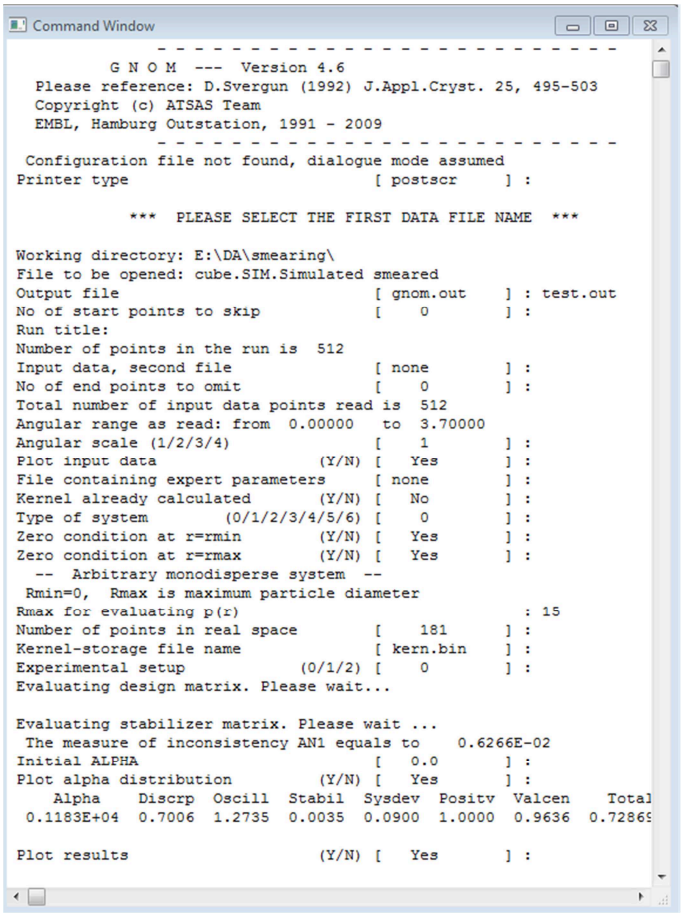


Figure 9: Exemplary data-input of the GNOM program. *Rmax* defines the maximum particle dimension in the PDDF. By changing the values under *Experimental setup* smearing effects can be defined.

GNOM is software written to compute the IFT of one-dimensional SAS curves. It is part of the ATSAS program-suite by Svergun, which is freely available for non-commercial use [17]. The actual fit is done the same way as proposed by Glatter [33] using the regularization method. [63] The significant novelty of this approach is the choice of the regularization parameter which in the case of GNOM is for the first time based on the quantitative description of *perceptional criteria,* such as oscillations of the PDDF etc. [64]. GNOM takes any background-corrected data file as input and computes either a PDDF for monodisperse systems or a size distribution for spherical polydisperse

systems. The output of this program is a single "*.out" file that includes the resulting PDDF, the initial scattering data, the desmeared scattering intensity of the corresponding PDDF and its deviation from the measured data.

All other programs of the ATSAS-suite used in this thesis require the GNOM-output file as an input, because all other theoretical models are fitted against the desmeared data from the deconvolution of the PDDF. Since all scattering curves used in this thesis are computed analytically and are therefore without smearing, the resulting desmeared data from GNOM is essentially the same as the input data. GNOM still need to compute the PDDF to obtain a maximum dimension as well as a scattering intensity with equidistant points. The program is operated by an interactive line input as can be seen in Figure 9.

3.3 DAMMIN

As mentioned above, DAMMIN is a program developed by Svergun that uses a simulated annealing procedure to fit a theoretical particle model against the scattering curve from deconvoluted PDDF which is obtained by GNOM [18]. For further use, this will be referred to as the *IFT-based scattering curve.* The underlying principles have already been explained under Chapter *2.2.3 Shape Retrieval Methods.*

The exact algorithm of the optimization procedure functions as follows: The program starts by randomly filling an enclosed spherical or elliptical volume, which can be defined by the user, with so called *dummy atoms* (DAs). It uses a closed-packed lattice to equally position the DAs. Using the Stuhrmann-approximation (see Chapter *2.2.2 The Multipole Expansion*), the scattering intensity of the randomized model is computed and compared with the desmeared curve according to the chi-squared functional (see Equation (29)). For now, we will refer to this as the *mean deviation*. In a next step, single DAs are either deleted or placed within the search volume. If the resulting mean deviation, including some kind of a *thermal term* (see Chapter *2.2.1 Simulated Annealing*), has improved, the change to the DAM is kept. After a certain amount of repetitions (user-defined value), this *thermal term* is decreased to a lower temperature. This procedure continues until one of two cases occurs: either the *thermal term* has been decreased for a certain amount of times, or the mean deviation is lower than a certain threshold value. For more details on the fitting-procedure see the initial

publication. [18] The result obtained from DAMMIN is the final configuration of DAs, also called *dummy atom model* (DAM), which yields the lowest mean deviation from the IFT-based scattering curve. An example of the results using the formfactor of monodisperse cubes can be seen in Figure 10.

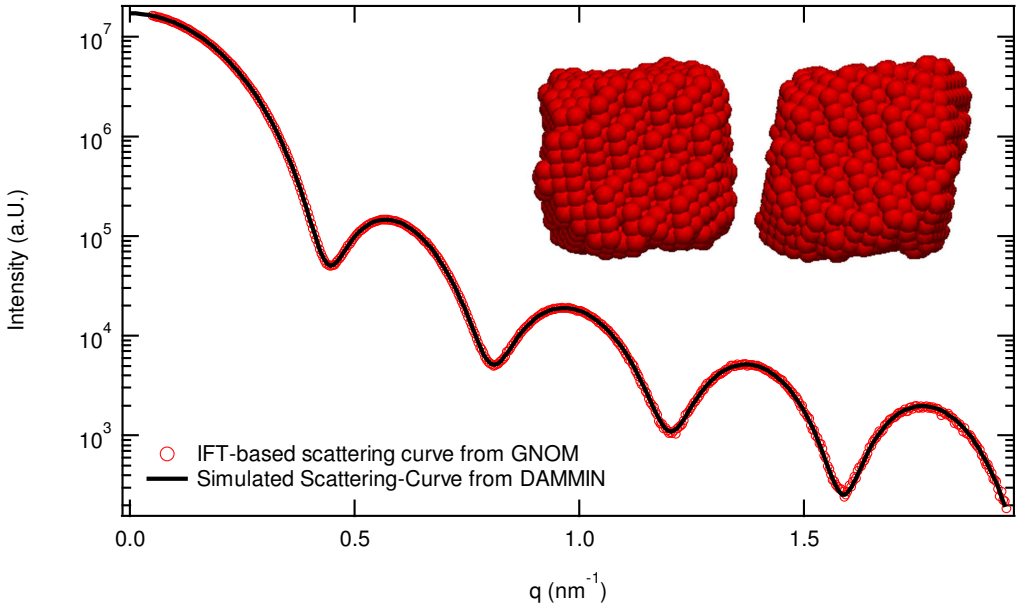


Figure 10: DAMMIN simulation using the formfactor of monodisperse cubes with a side-length of 16 nm. The final DAM is shown in the inset from two perspectives.

Additionally to the mean deviation, DAMMIN can also account for a so called *looseness criterion*. In any close-packed structure every DA has exactly twelve next-nearest-neighbors. When looking at a final DAM the mean number of next-nearest-neighbors can be seen as degree of looseness. In the fitting procedure, this mean number is multiplied by the *looseness penalty weight* (user-defined-value) and added to the chi-squared functional. Both of these terms together result in the function which has to be minimized. For all simulations in this thesis the looseness parameter was set zero, and was therefore neglected.

All simulations were performed on the SMMPMECH cluster of the Institute of Mechanics at the University of Leoben. The results presented in this thesis were obtained by using the default simulation-paramters as specified in Table 3 if not stated differently. The visualization of the DAMs is done using the open-source software ParaView. [65]

Table 3: Default DAMMIN simulation-parameters used in this thesis. All parameters not listed are kept at their default-values as suggested by the program.

parameter	default	reason
Number of knots in the curve to fit	124	A high number of splining points is chosen, to ensure the entire initial data curve is fitted.
Constant subtraction	-1	This procedure was skipped, since there is no background-scattering in idealized formfactors.
Order of harmonics	20	Since the goal of this thesis is to retrieve as much information as possible, the highest degree of convergence possible is chosen.
Initial DAM	S	A Spherical initial search volume is selected to neglect a biased initial shape.
Symmetry	P1	All symmetry effects are neglected.
Sphere diameter	Default	Suggested Rmax from the corresponding PDDF.
Packing radius of dummy atoms	Variable	This value is calculated from Rmax, such that approx. 5500 DAs are used.
Looseness penalty weight	Default	Default-value, suggested by the program.
Disconnectivity penalty weight	Default	Default-value, suggested by the program.
Peripheral penalty weight	Default	Default-value, suggested by the program.
Weight	2	Logarithmic curve weighting is chosen to ensure information in the entire q-range is treated equally.
Max # of annealing steps	250	This number is increased to ensure, the simulation stops when a minimal mean-deviation is found.

3.4 DAMAVER

To further analyze the models obtained by DAMMIN, a set of programs was written which handles multiple simulations of the same data. The first step of DAMAVER is to compare all given simulations and find the most probable one (program: *damsel* – more detailed description follows below). All models are then compared to the selected *mean model* and deleted, if they deviate from it more than a certain threshold level. The next step is to align all left over models to best fit the mean model (program: *damsup*). Both

programs damsel and damsup are based on another program called *supcomb* which superimposes two randomly oriented 3D structures on top of each other [66]. This is done by calculating the so called *normalized spatial discrepancy* (NSD). Simplified, the NSD is a measure for shape consistency and is minimized as two equal bodies are oriented the same way. During damsel, the NSDs between all models are calculated and finally the one model is chosen, for which the sum of the NSDs between it and all other models is a minimum. If the NSD between the selected model and any other one is then above a certain threshold level, the model is neglected for further use. In the last step all the aligned models are *averaged*. This is done by superimposing a new three dimensional lattice on top of the models. At every point of this new lattice, all aligned models are checked, if a given point of the new lattice is within the volume of a DA. This results in a *probability map* of all simulations, which Svergun denotes to as *occupancy.* The occupancy of a point in the averaged model therefore represents the number of DA that can be found in all superimposed models at this position. If, e.g., ten simulations are superimposed and there is one point of the new lattice where a DA can be found in all simulations, the occupancy of this point will be 10. An overview of the entire averaging process can be seen in Figure 11.

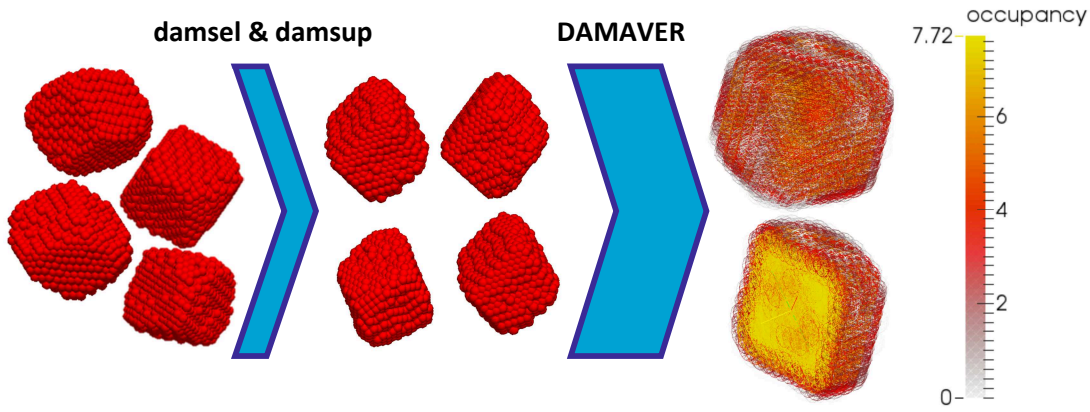


Figure 11: Overview of the averaging process of DAMAVER. After alignment of the models, they are superimposed on top of each other and a probability map is calculated. A high *occupancy* therefore means a high probability to find a DA in every model at the same position.

All simulations in this thesis were performed 10 times and averaged using DAMAVER.

4 Data Analysis

This chapter focuses on the numerical evaluation of the retrieved DAMAVER models. The resulting models are fully described by the position of the DAs using X_i and their corresponding occupancy Occ_i. Before any further data-treatment the models were repositioned such that the center of mass is in the origin of the Cartesian-system.

4.1 Weighted Radius of Gyration

As discussed in chapter *2.1.3 Small Angle Scattering* the determination of the radius of gyration (R_g) via the Guinier-Analysis is a common and powerful method to get an estimate of the mean size of the measured particles. Since x-rays are scattered by electrons the R_g determined from SAXS can be calculated for any particle volume as long as its spatial electron distribution can be described. The approach to do this comes from the field of classical mechanics such that the R_g is calculated in the same way by

$$R_{g,SAXS}^2 \triangleq R_{g,mech}^2 = \frac{\int \rho(r) r^2 dV}{\int \rho(r)\, dV} \qquad (33)$$

where $\rho(r)$ is either the mass density in mechanics or the electron density in condensed matter physics. The principle of this equation is that it directly correlates the electron density to a quantifiable number that can easily be measured and calculated. Using this equation the R_g of simple homogenous and radial symmetric, stepwise inhomogeneous geometries can be calculated analytically. The relationship between structural dimension and radius of gyration for geometries used in this thesis can be found in Table 4.

For a system built of point masses, the R_g can also be calculated using the finite version of Equation (33), namely as

$$R_{g,SAXS}^2 \triangleq R_{g,mech}^2 = \frac{\sum m_i {r_i}^2}{\sum m_i} \qquad (34)$$

In classical mechanics m_i is simply the mass which can be found at the distance r_i from the center of mass. For electron densities, m_i has to be replaced by the number of charges that can be found at the corresponding position.

Table 4: Formulas used for the calculation of the ROG for various geometries.

geometry	$\boldsymbol{R_g}$	**variables**
sphere	$R_g^2 = \frac{3}{5}R^2$	R … sphere radius
cube	$R_g^2 = \frac{W^2}{4}$	W … side length
ellipsoid	$R_g^2 = \frac{A^2 + B^2 + C^2}{5}$	A, B, C … axis of revolution
core-shell	$R_g^2 = \rho_{\text{core}}^* \frac{3}{5}R_C^2 + \rho_{\text{shell}}^* \frac{3}{5}\frac{R_O^5 - R_C^5}{R_O^3 - R_C^3}$	R_C, R_O … core, outer radius $\rho_{\text{core}}^* = \frac{\rho_{core}}{\rho_{core} + \rho_{shell}}$ $\rho_{\text{shell}}^* = 1 - \rho_{\text{core}}^*$

To ensure that during the entire scope of data evaluation, including fitting an IFT, performing a number of DAMMIN simulations and averaging all models via DAMAVER, no information was lost, the DAM was compared to the initial model by means of the R_g. To compute the R_g of the averaged model, the assumption is made that the *occupancy* is some kind of measure of a virtual electronic charge such that Equation (34) can be rewritten using $\|X_i\|^2$ as the norm, or distance from the center, as

$$R_{g,DAM}^2 = \frac{\sum Occ_i \|X_i\|^2}{\sum Occ_i} \qquad (35)$$

Before further investigations, the R_g was computed for all DAMs used in this thesis and compared to the initial model, from which the scattering intensities were calculated. Astonishingly, the values were without exception almost identical. The R_g of 15 different averaged models (cube, sphere, ellipsoid and core-shell-sphere) were calculated and can be found in Table 5. The mean deviation between ideal and obtained values is significantly less than 1 %.

From this it follows that the R_g of the DAM is equal to the R_g determined from the scattering curve. This implies that Equation (33) is valid which directly correlates the electron density of the particle with occupancy of the DAM. This is a significant finding

because it allows measuring and retrieving a three dimensional electron density map by interpreting the occupancy values of investigated systems.

Table 5: R_g from various obtained DAMs with dimensions according to Table 1. The ideal R_gs were calculated using the equations according to Table 4. Cube_1 was simulated using fewer harmonic. This will be discussed further in Chapter *5.1.3 Number of Harmonics.*

geometry	$\mathbf{R_g}$**from DAM [Å]**	ideal $\mathbf{R_g}$**[Å]**	**deviation [%]**
core-shell_1	57.697	57.239	0.80
core-shell_2	57.589	57.239	0.61
core-shell_3	57.589	57.239	0.61
core-shell_4	57.559	57.239	0.56
cube_1	78.893	80.000	1.38
cube_2	79.443	80.000	0.70
cube_3	79.419	80.000	0.73
cube_4	79.435	80.000	0.71
ellipsoid_1	58.059	58.031	0.05
ellipsoid_2	58.026	58.188	0.28
ellipsoid_3	58.581	58.662	0.14
ellipsoid_4	58.323	58.287	0.06
sphere_1	58.084	58.095	0.02
sphere_2	58.152	58.095	0.10
sphere_3	58.083	58.095	0.02

4.2 Linear Evaluation

When looking at characteristic lengths of the DAM (e.g. the radius of a sphere, length of a cube, etc.) several smearing effects occur. Since the DAM is built out of finite spheres with a radius r_{DA} the resolution of the model itself is limited by the DA size. Furthermore recall, that the final averaged DAM is a result of several single simulations which are statistically merged. The probability that all simulated models which are averaged are perfectly alike is very small. This means, that additionally to the resolution smearing, a statistical smearing will occur.

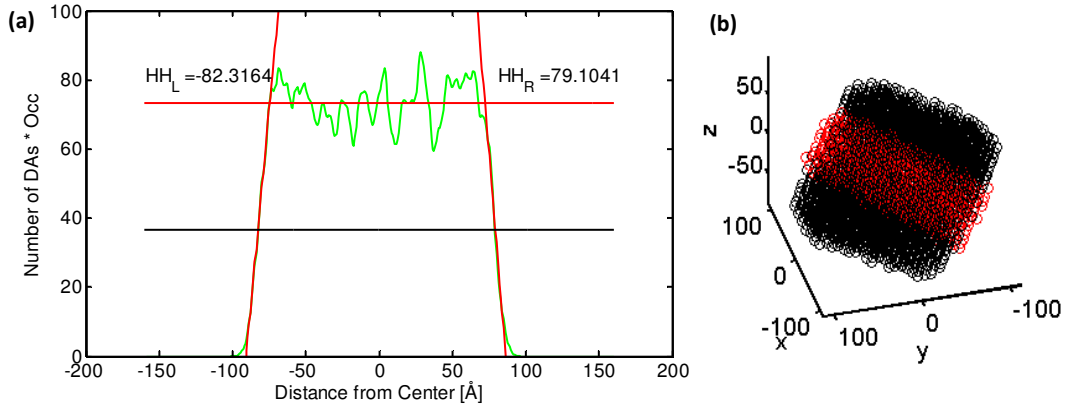


Figure 12: (a): Linear occupancy cross-section (green line) along the cut-through-cylinder with the corresponding linear fits (red lines) to retrieve the half-height-dimension (black line). The direction was chosen to measure along the short axis of the simulated cube with side lengths 16 nm. **(b):** 3D-scatter plot of the used model (black). The DAs used for the evaluation are highlighted in red.

The quantitative and exact determination of particle dimensions along specific directions is of great interest. To achieve this, a very long cylinder with a selectable radius is put through the center of the DAM. Afterwards, all DAs having a larger perpendicular distance from the cylinder axis than the cylinder radius are deleted. In the last step the cylinder is cut perpendicular to the axis into a number of equally sized slices. Now the mean occupancy of each slice is computed, by simply adding up the occupancies inside every slice. The result of this procedure is a linear cross section of the electron density along a certain direction. An example of this can be seen in Figure 12.

For homogeneous particles, the electron density, and therefore the occupancy, is constant on the inside of the model and can therefore be approximated by a line with slope zero. This constant value is referred to as the *mean occupancy*. The interesting part for the determination of model dimensions is obviously the border-region. This transition can be approximated by a simple tangent line, where the fitting region is defined manually. If now the tangent lines are intersected at half-height of the mean occupancy, the *half-height-dimension* of the given direction can be retrieved. In the example of Figure 12(a) the result is a half-height-dimension of 16.1 nm. Compared to the size of the real model of 16 nm (see Table 1) from which the scattering curve was calculated, this yields a deviation of less than one percent.

The linear evaluation becomes more difficult when the dimensions of rounded objects are to be retrieved. In the exemplary case above, the slices of the cut-through cylinder were always parallel to the facet of the cube. In the case of a sphere the transition-region will be distorted due to geometric reasons (see Figure 13(b)). It is therefore

advantageous to use the smallest cylinder-radius possible. This on the other-hand causes a new problem. For small cylinder-radii the probability of finding a DA in every evaluation slice gets very low which causes statistical fluctuations. For cylinder radii smaller than two DA diameters it is almost impossible to evaluate a half-height-dimension. This effect can be seen when comparing Figure 13(a) with Figure 12(a).

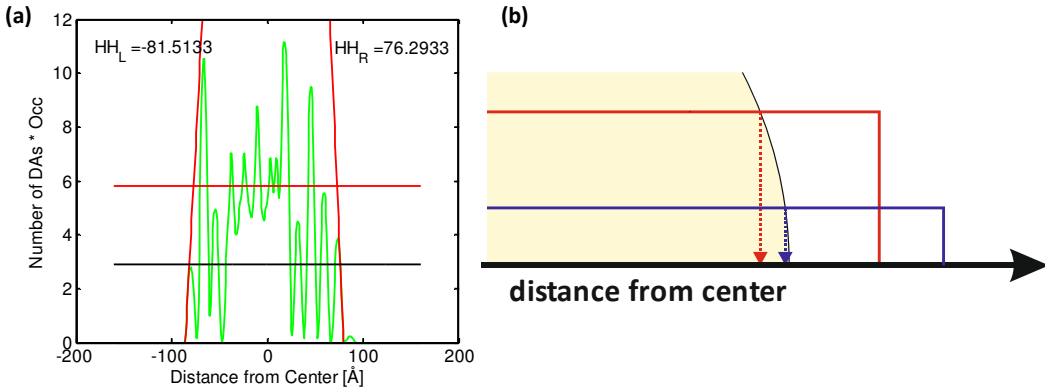


Figure 13: (a): Half-height-evaluation with a cylinder-radius of twice the DA diameter. The evaluation was performed on the same model and along the same direction as in Figure 12. Even though, the evaluation can still be done, there are already very high statistical fluctuations visible. **(b):** Illustration of the influence of the cylinder radius on the measured dimension of spherical models.

To avoid both of these problems, the direction of interest is measured multiple times with decreasing cut-through-cylinder-radii. If the obtained results are then plotted against the cylinder-radius, a linear relationship can be observed, which can be seen in Figure 14(a). As the real dimension can only be measured using an infinitely small cylinder, which for finite models is not possible, the linear relationship has to be extrapolated until it intersects the vertical axis. This is done by fitting a linear regression with the measured points and taking the offset of the line.

To graphically show how this is done, the method was applied to an ellipsoid of evolution model. The scattering curve was calculated for a homogenous ellipsoid of revolution with two long axis of 8.5 nm and a long axis of 5.25 nm. The measurements were taken along the long and the short axis, which can be seen in Figure 14(b). The cylinder radii were chosen to be multiples of the DA-diameter. This resulted in a short-axis radius of 5.3 nm and a long axis of 8.3 nm. Given that the DA-diameter is 0.4 nm, these measurements are within the spatial resolution of the DAM.

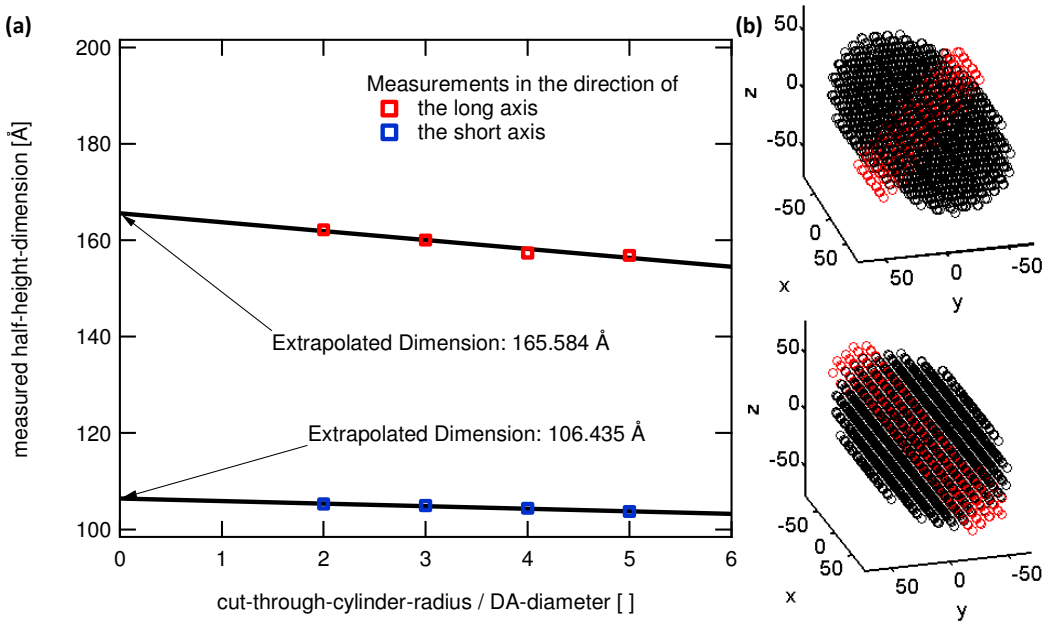


Figure 14: (a): Linear half-height-measurements taken along the same direction plotted against the cut-through cylinder radius. The measurements clearly yield a linear relation with the cylinder radius. After a linear regression, the line can be extrapolated to an infinitely small dimension, where the real model dimension can then be obtained. **(b):** 3D-scatter plot of the used model (black). The DAs used (red) for the measurement along the short/long axis is shown in the top/bottom view.

4.3 Averaged Dimension

The linear evaluation described above obviously needs some kind of decision regarding the measurement direction, which implies a subjective judgment on the model shape. This on one hand provides a powerful tool, if one single dimension is to be measured. On the other hand this can lead to a misinterpretation of the overall shape. To avoid this, the linear evaluation was fully automated. In a first step, the *occupancy density* of the cross-section is normalized. This is done by determining the average value between $-R_G/2$ and $+R_G/2$ and dividing the data by this. To determine the half-height boundaries, a high initial distance from the center is chosen and decreased step by step. The first time the normalized occupancy density exceeds 0.5, the boundaries are found.

As a result this method speeds up the evaluation process and makes it absolutely objective. It also allows choosing spherical random evaluation directions. If this is repeated for a very large number of random directions, a kid of *radial distance distribution* can be obtained. An example for the ellipsoid of revolution shown in Figure 14 can be seen in Figure 15. This histogram can simply be read as the number of

times a certain distance was measured. Since the direction of measurement is always centered in the origin, the values obtained are twice the *center to surface* distances. For simple geometries, an ideal distance-distribution can easily be derived. In the example of a monodisperse sphere, the distance distribution is represented by an infinitely sharp peak at the sphere-diameter. In the case of an ellipsoid of revolution the distribution is shown in the black curve of Figure 15(a).

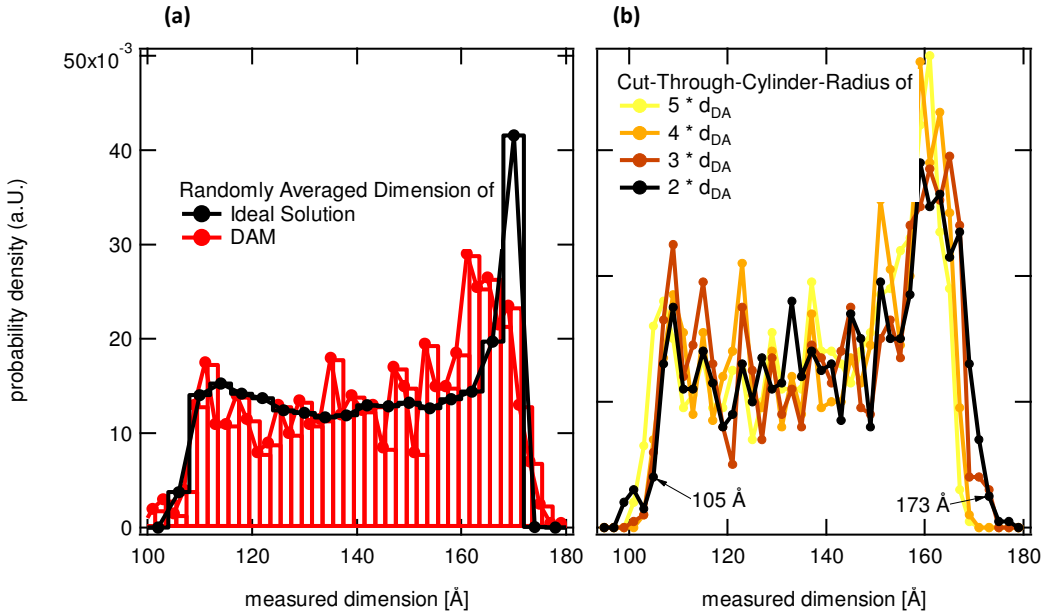


Figure 15: Various distance-distributions determined using spherically-random oriented linear evaluations. **(a):** The distance distribution of the DAM mentioned in the chapter above (red) is compared to the ideal solution which can be solved semi-analytical (black). The solution of the DAM was computed using a cut-through-cylinder-radius of 2∗DA-diameter.The DAM qualitatively is in good agreement with the ideal solution. **(b):** The same DAM of the simulated ellipsoid of revolution is measured using a decreasing cut-through-cylinder-radius. As expected, a shift of the measured dimension can be observed.

As mentioned in the chapter above, the radius of the cut-through-cylinder has a significant effect on the determined dimension. This is also true for the case of multiple measurements, as can be seen in Figure 15(b). At this point it is not clear how exactly to interpret the results. E.g., if the highest significant peaks of the smallest cylinder radius are chosen, values of 10.9 nm and 16.7 nm are obtained. Compared to the initial model values of 10.5 nm for the short axis and 17 nm for the long axis (according to Table 1) and considering a DA size of 0.4 nm, these measurements are in very good agreement with the initial values.

4.3 Spherical Average

As explained in chapter *2.1.2 The Debye-Formula*, the computation of the scattering intensity is based on an orientational average of a given particle symmetry. This implies that only spherical averaged information in the real space can be retrieved. With this in mind, an evaluation algorithm was written to retrieve characteristic structural information, such as edges, and determine a spherical averaged dimension. This method works for averaged models (DAMAVER) with the origin at their weighted center of mass.

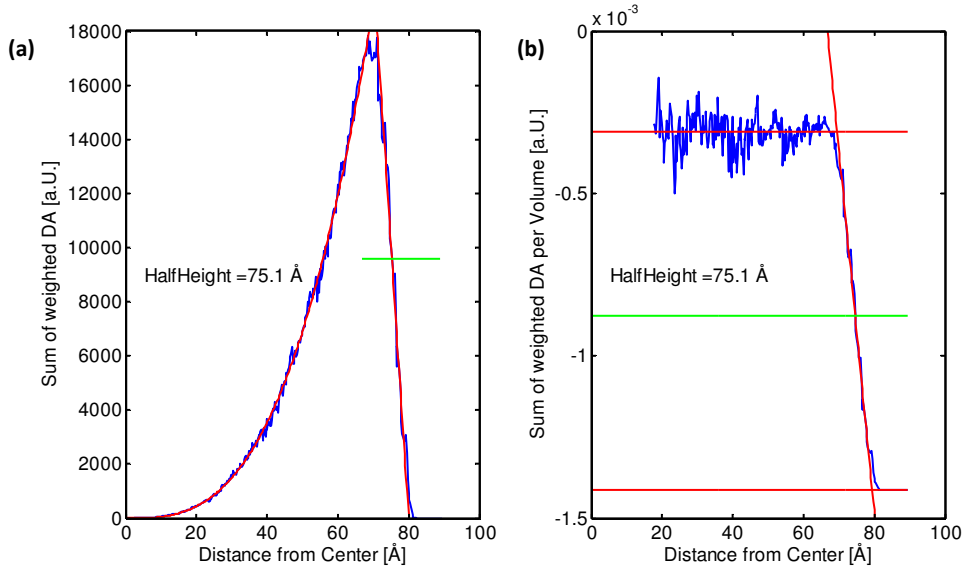


Figure 16: Results of the spherical-average-evaluation of a (almost) monodisperse sphere with diameter 15 nm. **(a)**: The weighted sum of all DAs within a spherical search volume is ploted against the corresponding radius which also is the distance-from-center. Using the half-height of the intersection of both fits, an almost perfect measurement can be retrieved. **(b)**: The density-difference of the DAM and ideal density is plotted over the distance-from-center. The density at the inside appears to be constant. As expected, a significant change can be seen at the outer DAM dimensio. By using a half-height fit a radial dimension can be obtained.

The averaging process is based on the "*onion-principle"* and consists of two procedures. They are both illustrated for the exemplary case of a sphere with 1 % polydispersity (numerical smearing necessary) and a diameter of 15 nm.

The first procedure simply sums up all DA that can be found in the volume of a sphere with an increasing radius and center at the origin. To account for occupancy-effects the summation is not performed of the DA themselves, but over their occupancy values. The result can be seen in Figure 16(a). The D*istance from Center* plotted on the horizontal axis corresponds to the radius of the evaluation sphere. As expected, starting from the

origin, the number of DA increases with a cuboid-power law. At one point (at approx. 70 Å) a deviation from this correlation occurs and the weighted number of DAs decreases linearly. To quantitatively evaluate this curve, a 3rd-degree-polynom is fitted in inner region and a simple tangent fit is done for the linear decrease. If now the half-height-value of the intersection of these two fits is taken, a distance to center of 75 Å is retrieved, which is in perfect agreement with the ideal model.

The second procedure uses a familiar principle. In this case, only the DAs within a spherical shell with a thickness of $0.005 * R_{Max}$ are looked at. Since the packing-density of the DAM is known (74 % for FCC) the ideal number of DAs can be calculated. If now both of these numbers are normalized to the volume and subtracted, we obtain a density difference, which can be seen in Figure 16(b). At the center of the model, the density is at a minimum and seems to be constant. As the distance-to-center is increased, a sudden change in the density-difference can be seen. This can be interpreted as an increasing deviation of the DAM-density from the ideal one. If this effect is fitted using two horizontal lines to obtain constant in- and outside values, and a tangent-fit to describe the density-difference increase, a half-height dimension can be obtained. Compared to the first procedure and the initial dimensions, this value is in perfect agreement.

This second method allows the interpretation of the spherical averaged occupancy of the entire model which, from a statistical point of view, is of high quality, since all DA of the model are being looked at. The disadvantage of this approach is that it needs a manual input of the fitting-regions. An automated method was tried, but failed, since for more complex structures, density fluctuations on the inside can occur. This will be investigated in the following chapter.

5 Results & Discussion

As stated in the introduction, one of the main objectives of this work is to determine the influence of the simulation parameters on the final result, or in other words, how reproducible and reliable these results are. The other objective was to develop evaluation methods that allow the direct interpretation of the DAM. The methods have already been introduced in detail in Chapter *4 Data Analysis.* A figurative overview of the evaluation methods and the obtained diagrams can be found in Figure 17. Part of the development of these evaluation techniques, and therefore also another objective of this work, is to investigate the stability and correctness of the obtained results. Both of these topics will be discussed in the following chapter.

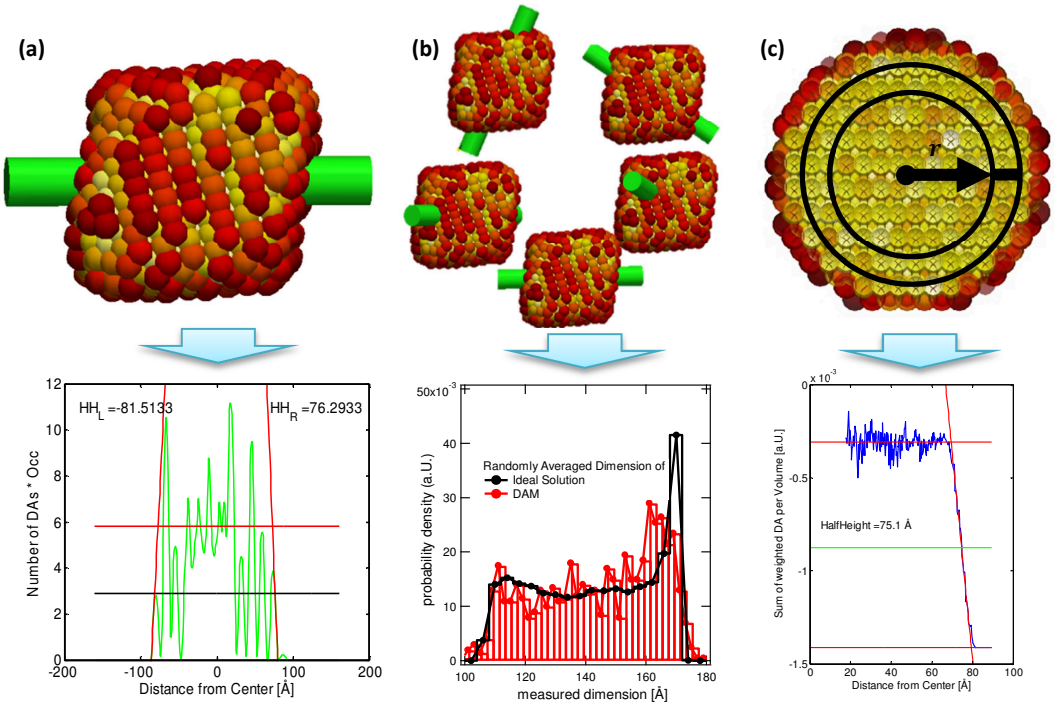


Figure 17: Overview of the used evaluation methods to directly interpret the retrieved DAM: **(a)** linear-evaluation, **(b)** averaged dimension and **(c)** spherical average evaluation. The detailed diagrams can be found in Figure 13(a), Figure 15(a) and Figure 16(b)

5.1 Stability of Results

From previous investigations it became obvious, that these parameters are *the number of DA per model, the number of Harmonics* and *the number of Shannon channels provided from the scattering data.* For every investigated parameter a series of ten equal simulations was performed and averaged. Unless described differently, the default parameters according to Table 3 were used.

5.1.1 Number of Dummy-Atoms

To investigate the resolution and stability of the results obtained the most obvious parameter to be changed is *the number of DAs*. For geometrical reasons, the resolution of the DAM is limited by the distance between two DAs. The first thought to increase the resolution is therefore to decrease the DA size. The number of DA correlates to the third power of the DA radius. Due to the computation of the double sum over all DAs to obtain the scattering intensity, the simulation time correlates to the square of the number of DAs. This is a significant drawback. In the example of a sphere with diameter of 15 nm a decrease of the DA diameter from 0.4 nm to 0.32 nm increases the number of DAs from approx. 6000 to 12000. The simulation time increased from approx. 16 h to 52 h. It is therefore of great interest to find a compromise between simulation time and resolution. To investigate the real effect of the DA-size, a series of calculations of a sphere was obtained using the specifications according to Table 6.

Table 6: Simulation parameters to investigate the influence of the number of DAs on the results. A sphere of diameter 15 nm and an artificial polydispersity of 1 % (the reason for this will be explained further below) were used to allow correct splining of GNOM.

name	radius of DA [Å]	number of DAs
sph_da_3	5	3043
sph_da_6	4	5895
sph_da_9	3.5	8781
sph_da_12	3.2	11590

As it can be seen in the results according to Figure 18(a-d) obtained from the spherical-average-evaluation all half-height-results are within the initial values ± the DA-diameter. When looking at the result obtained using the averaged-dimension method, all simulations fluctuate between 14.5 - 14.6 nm. The size of the cut-through-cylinder was chosen to be 2.5 nm to obtain good statistics. The real dimension will therefore be

underestimated which is the case as demonstrated in Figure 15(e). The important fact is that there is no evidence for a correlation between the number of DAs and the estimated particle size.

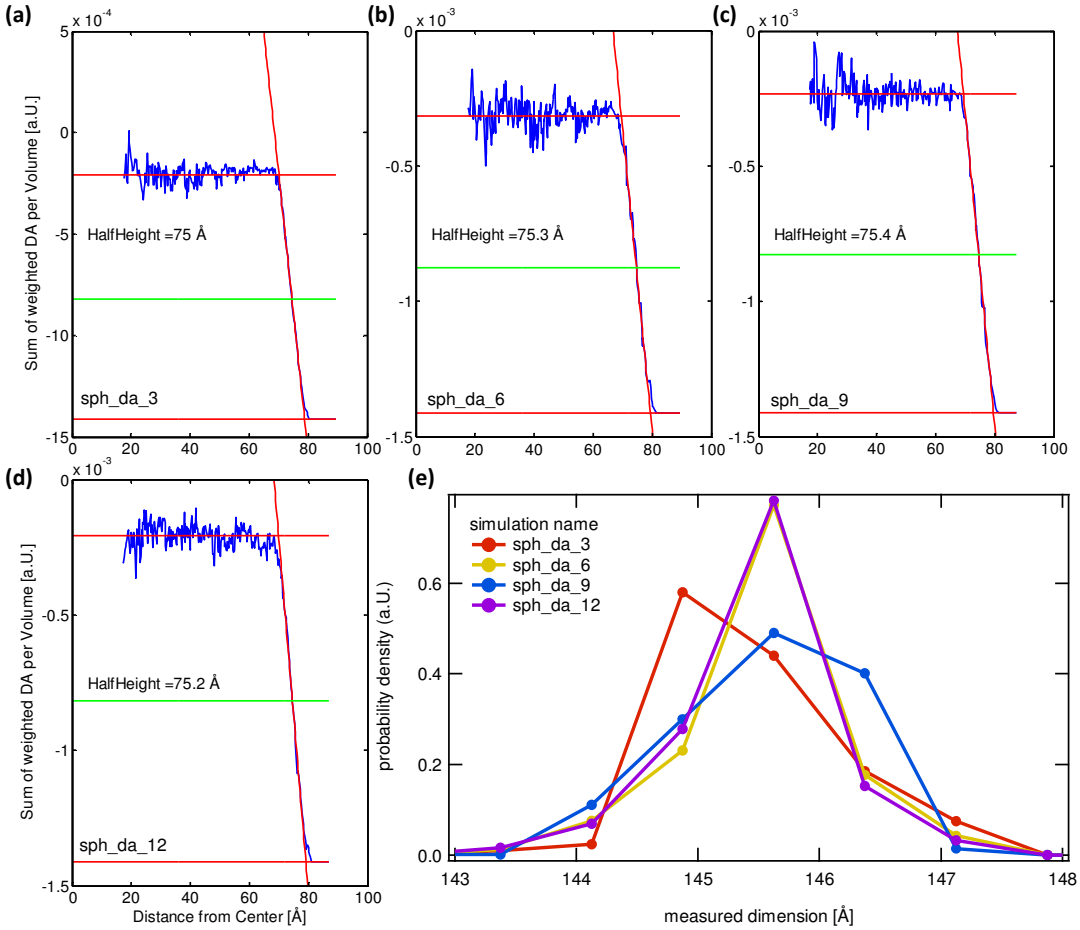


Figure 18: Spherical-average-evaluation of the simulations **(a)** sph_da_3, **(b)** sph_da_6, **(c)** sph_da_9 and **(d)** sph_da_12. **(e)** Averaged-dimension of the four sperical-DA-simulation-series using 1000 random directions and a cut-through diameter of 2.5 nm. Due to this rather large diameter, the obtained values will therefore underestimate the real dimension.

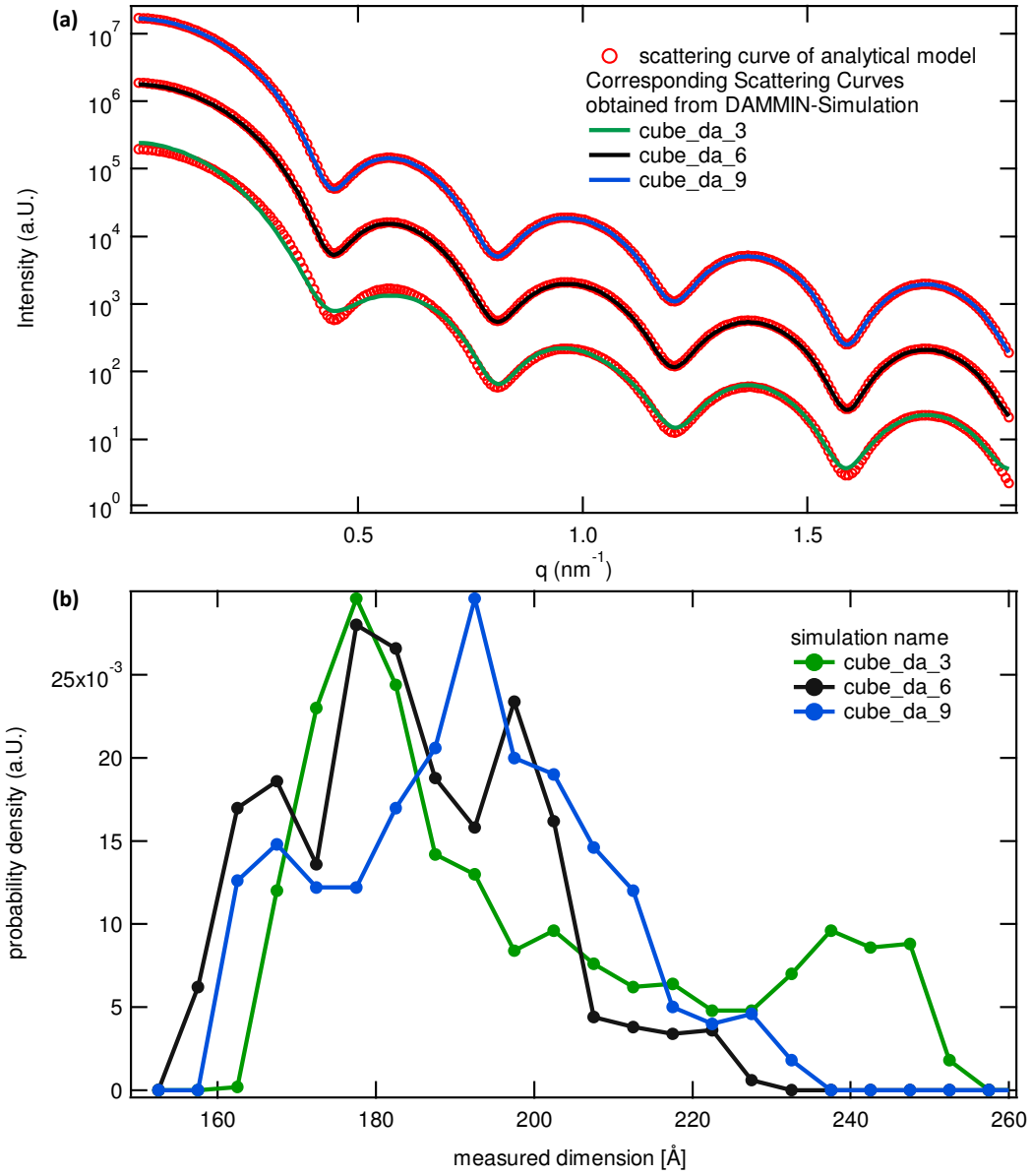


Figure 19: (a) DAMMIN results compared to the initially provided scattering curve of a cuboid with side length of 16 nm. The simulation with approx. 3000 DAs (cube_da_3) significantly deviates from the initial curve. **(b)** Averaged-dimension of the three cuboid-DA-simulation-series using 1000 random directions and a cut-through diameter of 3.87 nm. Also in this case, the simulation cube_da_3 clearly deviates from the other two as can be seen by a second maximum around 240 Å.

To further prove this point, a similar series of simulations was performed using a cuboid with side length of 16 nm. Compared to a sphere, a cuboid presents significant surface structures such as edges and planes. The parameters used can be found in Table 7. The interesting phenomenon here is that DAMMIN was not able to fit a model against the initial scattering curve with only approx. 3000 DAs. This can be seen in Figure 19(a). Especially in the low q-regime, the simulated curve clearly deviates from the initial one.

This effect also becomes visible when looking at the averaged-dimension-evaluation in Figure 19(b), where cube_da_3 exhibits a bimodal distance center around 240 Å. This corresponds to the overestimation of the scattering intensity at $I(0)$. All three simulations show some kind of minimum dimension at approx. 16 nm. If simulation cube_da_3 is neglected due to the reasons mentioned above, this minimum dimension is between 15.75-16.25 nm which is in perfect agreement with the initial dimension. An interesting fact is that the diagonal of the initial model is 27.7 nm but there is no distance measured in the cube_da_6 and cube_da_9 models higher than 24 nm. This implies that the corners of the cube cannot be resolved. The high cut-off value of the averaged dimension is somewhere around 22-23 nm which corresponds to the in-plane diagonal of the cube which is 22.6 nm.

Table 7: Simulation parameters to further investigate the influence of the number of DAs on the results. In this case, a cuboid with side-length of 16nm was used

name	radius of DA [Å]	number of DAs
cube_da_3	8.8	2988
cube_da_6	7.7	5594
cube_da_9	6.1	8853

In summary, the number of DAs is to be chosen to ensure that DAMMIN can describe all structural details. In case of a sphere, which is a body with no significant structural characteristics, the number of DAs has no effect whatsoever on the result. With increasing structural detail of the model the size of the DAs is to be decreased to ensure proper description of the structural singularities. Simulations with approx. 6000 DAs seem to be a good compromise between acceptable simulation time and retrievable detail. As long as the shape can fully be described the size of the DAs yields no significant effect on the obtained measurements.

5.1.2 Shannon Channels

As already discussed in Chapter *3.1.1 Monodisperse Systems*, the information content provided by the q-range of the scattering curve can be described using *Shannon Channels* (SCs). In real experiments the accessible q-range and therefore the amount of provided SCs is usually limited by the experimental setup. To see the effects of the amount of provided SCs on the final results a series of simulations with different geometries and density profiles was performed. To best illustrate these effects, a spherical core-shell model was chosen. The radius of the core is 5 nm with a 2.5 nm shell

(see Table 1). The electron density ratio of core to shell is 1:2. The resulting width of a Shannon channel is according to Equation (32) therefore 0.21 $nm^{-1.}$ The simulation-name with the corresponding Q_{max} can be found in Table 8. A decrease of the accessible q-region means a decrease of the number of SCs. To illustrate this effect, the scattering curves as well as fits from the corresponding DAMMIN simulations are shown in Figure 20. All simulations were successful. The DA-diameter was chosen to result in approx. 6000 DA per model.

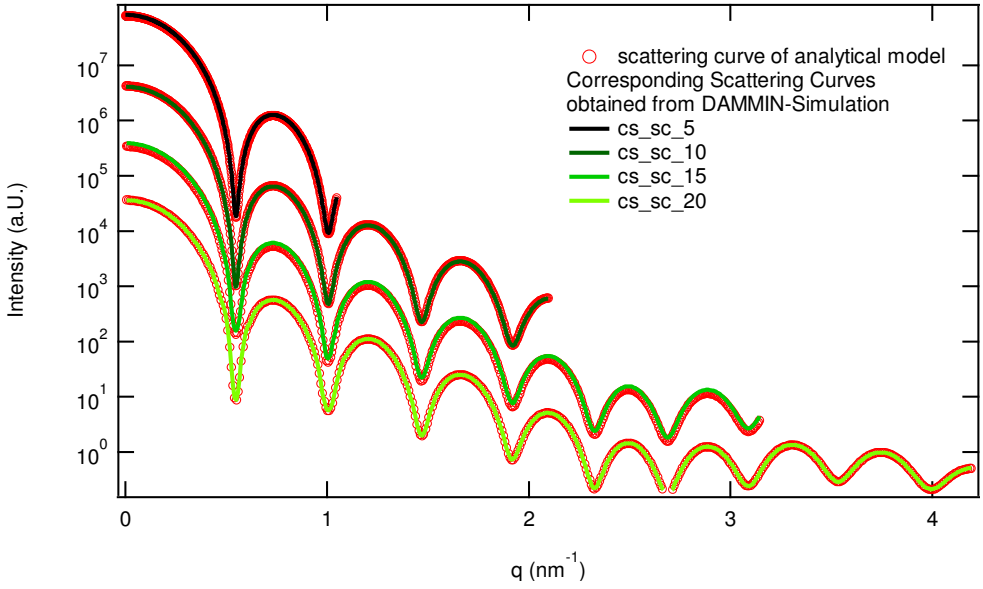


Figure 20: DAMMIN results compared to the initially provided scattering curve of with the corresponding q-ranges for the simulations listed in Table 8.

Table 8: Simulation parameters to study the influence of decreasing number of Shannon Channels, or in other words, the influence of a decrease of the accessible q-range.

name	Shannon Channels	q_{Max} [nm^{-1}]
cs_sc_5	5	1.05
cs_sc_10	10	2.10
cs_sc_15	15	3.15
cs_sc_20	20	4.20

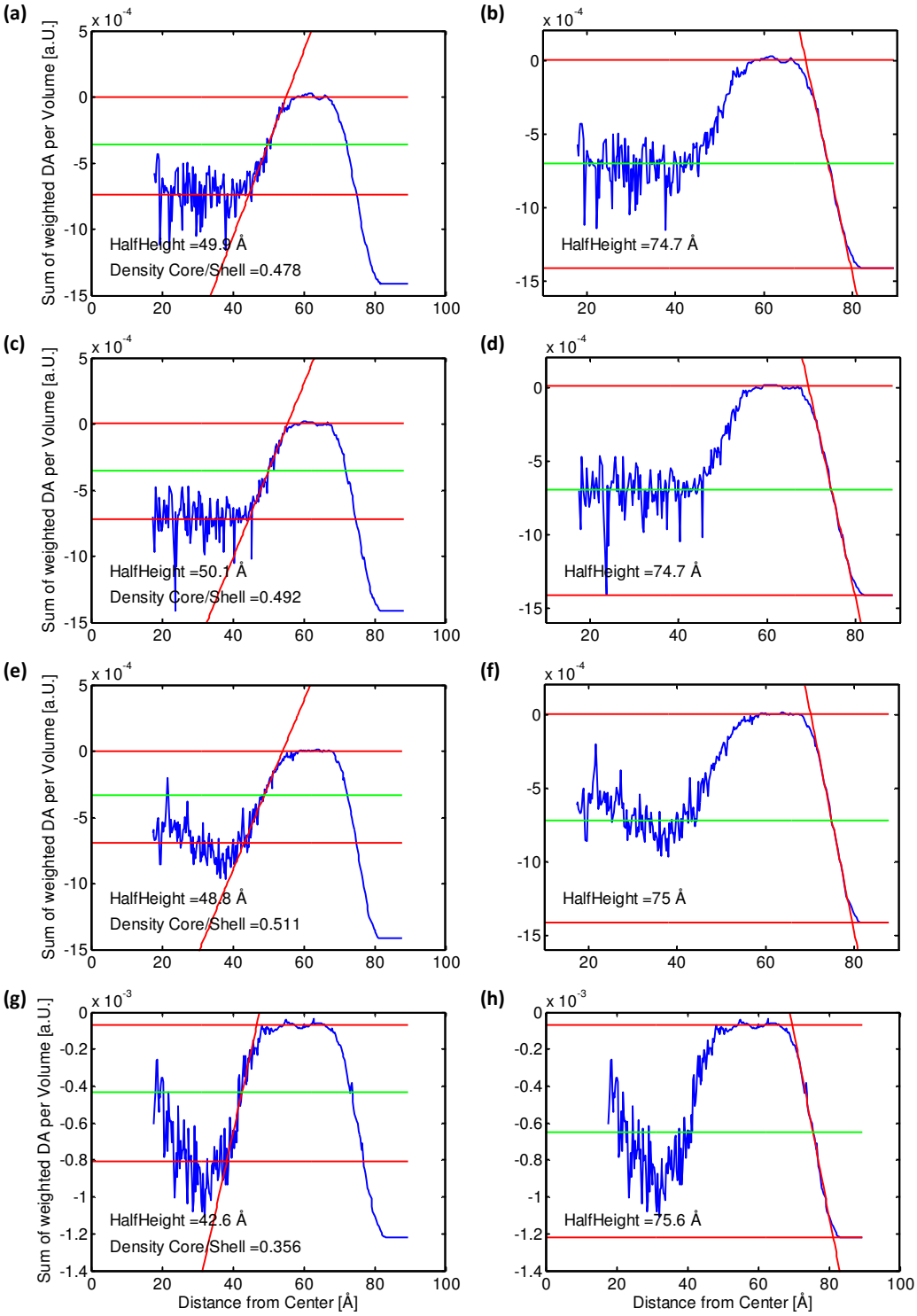


Figure 21: Spherical-average-evaluation of the core-shell model with 5 nm core and 2.5 nm shell with an electron density-ratio of core:shell = 1:2. **(a,c,e,g)** Determination of the transition dimension of core to shell. **(b,d,f,h)** Determination of the outer dimension. **Rows:** Provided scattering q-range with **(a,b)** 20, **(c,d)** 15, **(e,f)** 10 and **(g,h)** 5 SCs.

The first effect that becomes visible when looking at the spherical-average-evaluation presented in Figure 21 is the obvious change in electron-density difference throughout the shape. If the ratio between the core and shell plateau is calculated for cs_sc_20 (see Figure 21(a)) a core/shell-density-ratio of approx. 0.48 is obtained which is in good agreement with the initial values of the model. The half-height-evaluation provides a transition-dimension of approx. 4.99 nm and an outer radius (see Figure 21(b)) of 7.47 nm, which is also in good agreement with the initial values. A decrease of the SCs from 20 to 15 (see Figure 21(c,d)) shows no effect on the obtained values. Only after decreasing the SCs down to 10 the occupancy at the core is starting to fluctuate (see Figure 21(e,f)). This makes it harder to provide the boarders for the averaging of the core plateau and therefore influences the transition-dimension as well as the core/shell-ratio. This effect becomes even worse when the number of SCs is decreased to 5 (see Figure 21(g,h)) which makes it almost impossible to provide boarders for the tangent fit and the averaging of the core plateau. Furthermore, at this point the deviation of the obtained outer radius (see Figure 21(h) from the initial values becomes significant (more than the DA-diameter). The same effect is visible when looking at the averaged dimension of the simulations shown in Figure 22. The values of the simulations with 20, 15 and 10 SCs are in good agreement with the initial dimensions.

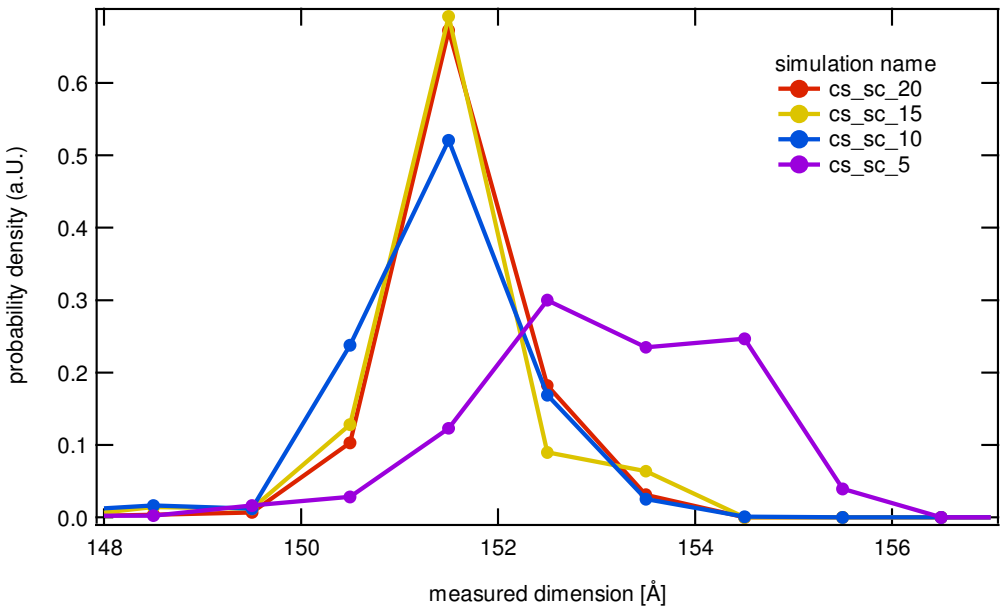


Figure 22: Averaged-dimension of the SC-simulation-series using 1000 random directions and a cut-through diameter of 2.5 nm. For 20, 15 and 10 SCs the result appears very alike with a maximum between 15.1-15.2 nm. There is a clear deviation in the case of 5 SCs. The parameters to the corresponding simulation can be found it Table 8.

When looking at Equation 20, it can be seen that Q_{Max} somehow correlates with the minimum dimension or resolution that can be retrieved from a given scattering curve. As it can be seen from this investigation a truncation of the measured q-range in a real experiment can lead to misinterpretation. Even if *a priori* knowledge is available (e.g. spherical symmetry of the shape and inner structure) the q-range should be at least 10 SCs to avoid any misconstruction. If detailed information on the inner particle structure such as the electron density is to be obtained, the range of accessible SCs should be kept as high as possible.

If this is the case, using the spherical-average-method the relative electron density can be obtained with an astonishing degree of detail. Besides this, the characteristic dimensions of the DAM can be determined, even for low numbers of SCs.

5.1.3 Number of Harmonics

As explained in Chapter *2.2.2 The Multipole Expansion* the number (or order) of harmonics used is a degree of convergence of the computation of the scattering curve from a given DAM. Since the simulation time correlates to the square of the used number of harmonics, Svergun suggests the use of 10 orders as a middle ground between resolution and simulation time. [67] Also for the investigation of the decrease of the number of harmonics the initial model of a cube with 16 nm side length and a DA-diameter of 7.2 nm was used (see Table 1). The simulation names with the effective duration of a single simulation can be found in Table 9. The calculations were performed on dedicated CPUs of the same model to ensure comparability of the time-results. All fits were successful as can be seen in Figure 23.

Table 9: Simulation parameters to study the influence of decreasing number of harmonics. The reader can also find the effective duration of a single simulation.

name	harmonics	simulation-time [h]
cube_h_20	20	15
cube_h_15	15	11.7
cube_h_10	10	8.7
cube_h_5	5	6.9

When looking at the duration in dependence of the number of harmonics it is evident, that there is no square-correlation between them but rather an asymptotic relationship. It is to be noted, that the simulation break criterion was set so a minimal threshold

mean deviation is reached (see Chapter 3.3 DAMMIN) to ensure solutions of equal qualities are found.

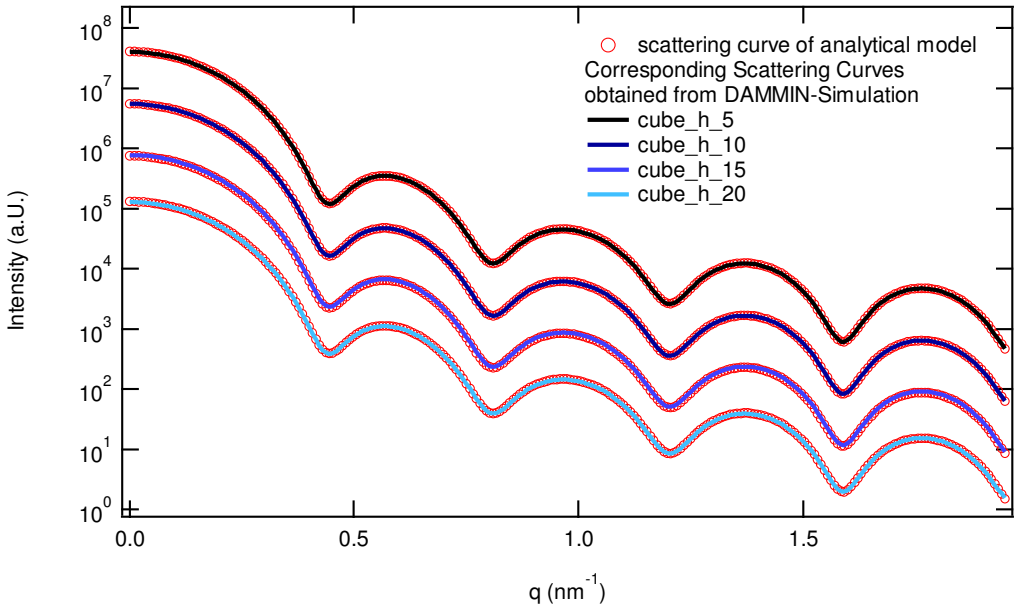


Figure 23: Analytical scattering curves of the cube used in the simulation series investigating the influence of the number of harmonics on the result compared to the fits of the DAMMIN simulation. All simulations resulted in a perfect fit.

The results of the spherical-average-evaluation of this series can be found in Figure 24(a,c,e,g). First off, there is no qualitative impact visible when the order of harmonics is decreased. There is a slight but continuous drift in the half-height values from 9.9 to 9.7 nm. This might be due to the fact that with decreasing resolution the DAs are on average placed further towards the center since the facets and edges cannot be resolved. This could therefore indicate some kind of structural loss. If this is compared to the size of a single DA with 0.72 nm, this drift is negligible. When directing the attention to the actual DAMs (see Figure 24(b,d,f,h)) more information can be gained. The models resulting from cube_h_20 and cube_h_15 show no evidence of an impact resulting from the decrease in harmonics. When the number is further decreased to 10 harmonics, the corners of the DAM are beginning to be flattened. In other words, there is a loss of the degree of surface detail. Furthermore the model is starting to become asymmetrically strained. This effect is visible even better when looking at cube_h_5. Here even the cuboid-planes are starting to become instable and the overall shape is starting to look *star-like*.

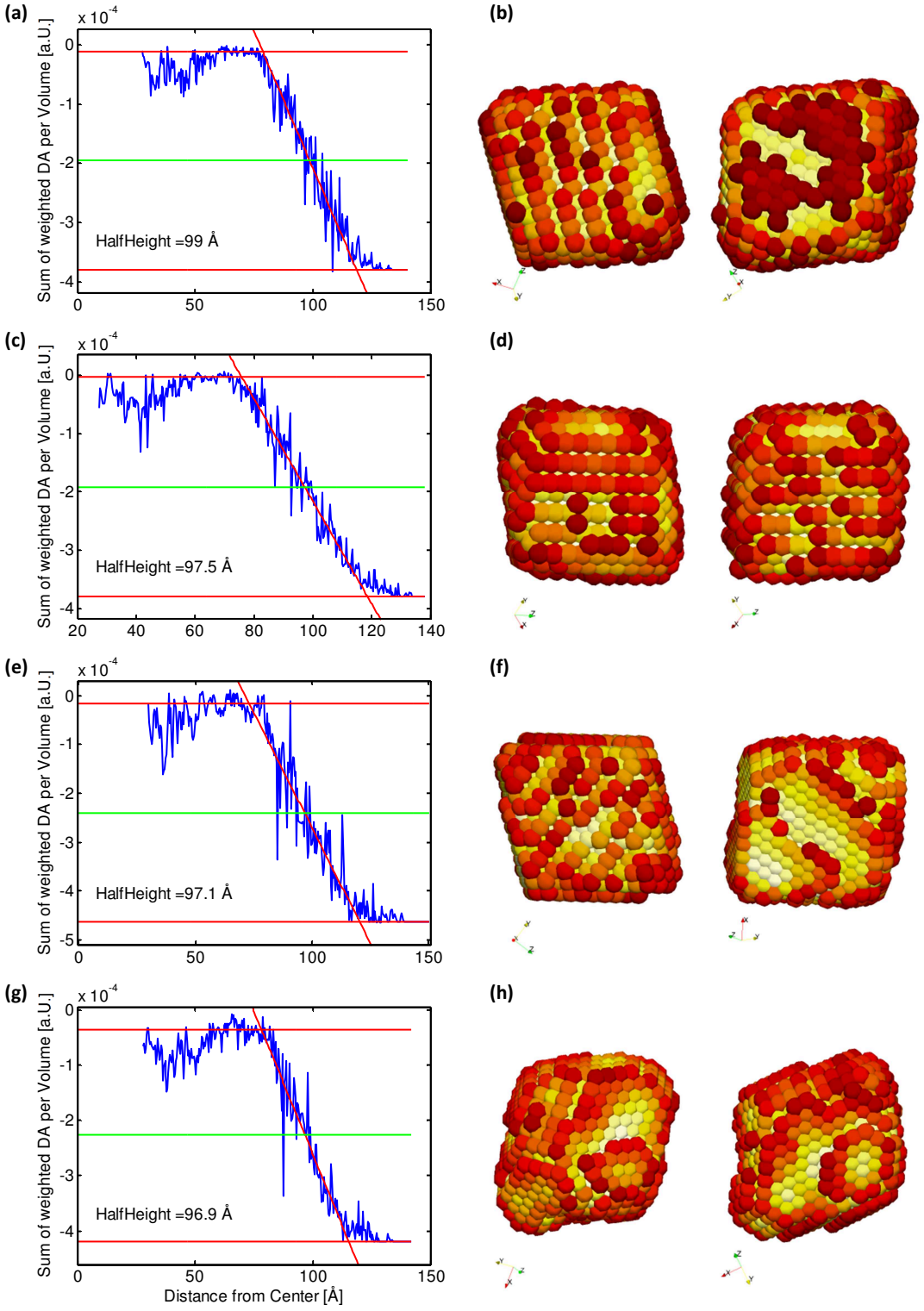


Figure 24: Results of the simulation series investigating the influence of decreasing the number of harmonics. **(a,c,e,g)** The spherical-average-evaluation of the DAMs shows no significant impact on the inner structure and the half-height-dimension. **(b,d,f,h)**: DAM shown from two perspectives. The color-gradient is a measure of the occupancy. **(a,b)** cube_h_20, **(c,d)** cube_h_15, **(e,f)** cube_h_10 and **(g,h)** cube_h_5.

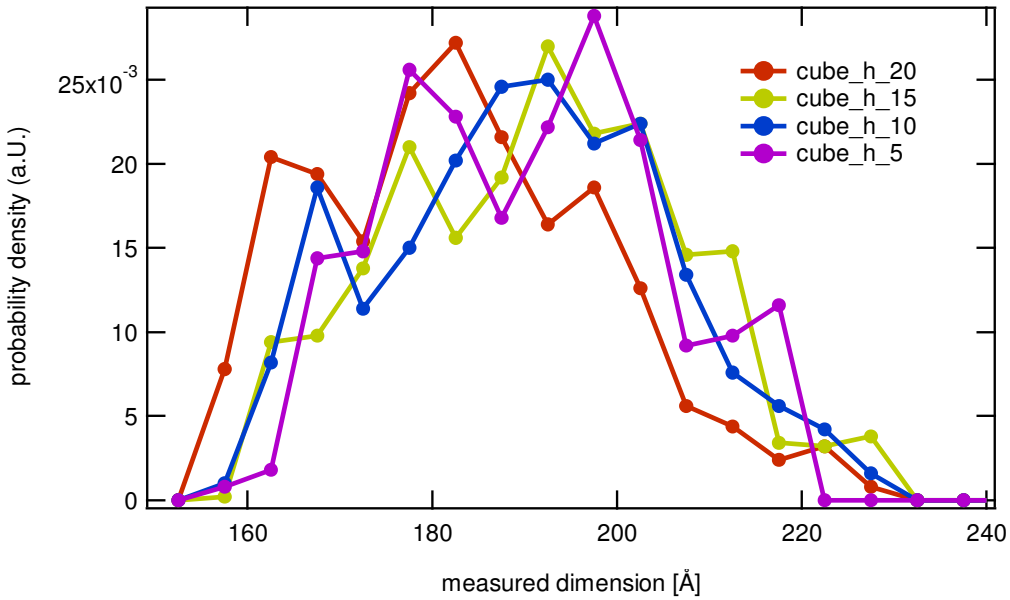


Figure 25: Averaged-dimension of the harmonics-simulation-series using 1000 random directions and a cut-through diameter of 3.875 nm. There is no strong evidence of an influence visible.

To quantitatively describe the surfaces of the different models, the averaged-dimension-evaluation was performed for which the results are visible in Figure 25. On the first look there is no strong evidence of a difference in the models. With a closer look towards the measured minimum and maximum dimension, a *narrowing* of the length-range can be noticed. For cube_h_5, the lower threshold value is increased from 15.7 to 16.3 nm whereas the upper threshold value is decreased from 22.8 to 21.8 nm. The overall narrowing is therefore 1.5 nm, which is approx. twice the DA-diameter. This corresponds to the drift of the half-height-values obtained from the spherical-average-evaluation.

Both observable phenomena of the cube_h_5 simulation indicate a *rounding* of the overall shape with a decrease of the number of harmonics. Nevertheless, all models suggest an overall cuboid shape. If absolute measurements are to be obtained from the DAM, it is recommended to use the maximum number of harmonics possible. Further, the use of less than 10 harmonics can also lead to the qualitative misinterpretation of the DAM. Models that are still retrieved with a low number of harmonics have to be interpreted very carefully. Since the computation time is not very costly nowadays, it is highly recommended to always use 20 orders of harmonics.

5.2 Polydispersity

The other main objective of this work is to investigate the influence of smearing effects on the simulation results. Real scattering experiments on inorganic NCs will always provide data, which are at least to some degree smeared by a finite instrumental resolution and some degree of polydispersity of the system. These two effects on the measured scattering intensity have to be distinguished from each other (see Chapter *3.1.3 Smearing Effects* for more details). Therefore in a first simulation series the effects of polydispersity on the DAM were studied for the simplest case possible, namely spheres. After this investigation, another series was conducted on ellipsoids of revolution. This chapter is therefore divided into two sections, each dealing with its dedicated geometry.

5.2.1 Sphere

Table 10: Names and corresponding polydispersity of the simulations used to investigate the influence of a finite size distribution on the results. The *polydispersity* is the standard deviation of the Gaussian-volume-distribution.

name	polydispersity [%]	name	polydispersity [%]
sphere_pd_1	1	**sphere_pd_8**	8
sphere_pd_2	2	**sphere_pd_10**	10
sphere_pd_4	4	**sphere_pd_15**	15
sphere_pd_6	6	**sphere_pd_20**	20

For this simulation series an initial model with diameter 15 nm was chosen. The simulation-names and the used polydispersity can be found in Table 10, whereat the polydispersity is the standard-deviation of the volume-distribution. It is not possible to use the formfactor of a perfectly monodisperse sphere since the minima are so distinct that DAMMIN cannot fit them. Therefore an artificial polydispersity of at least 1 % has to be used to successfully use DAMMIN for spheres. This is interesting, since a program that is based on the presumption of perfect monodispersity needs a threshold size deviation to actually work. The analytically computed scattering curves as well as the corresponding DAMMIN fits can be found in Figure 28.

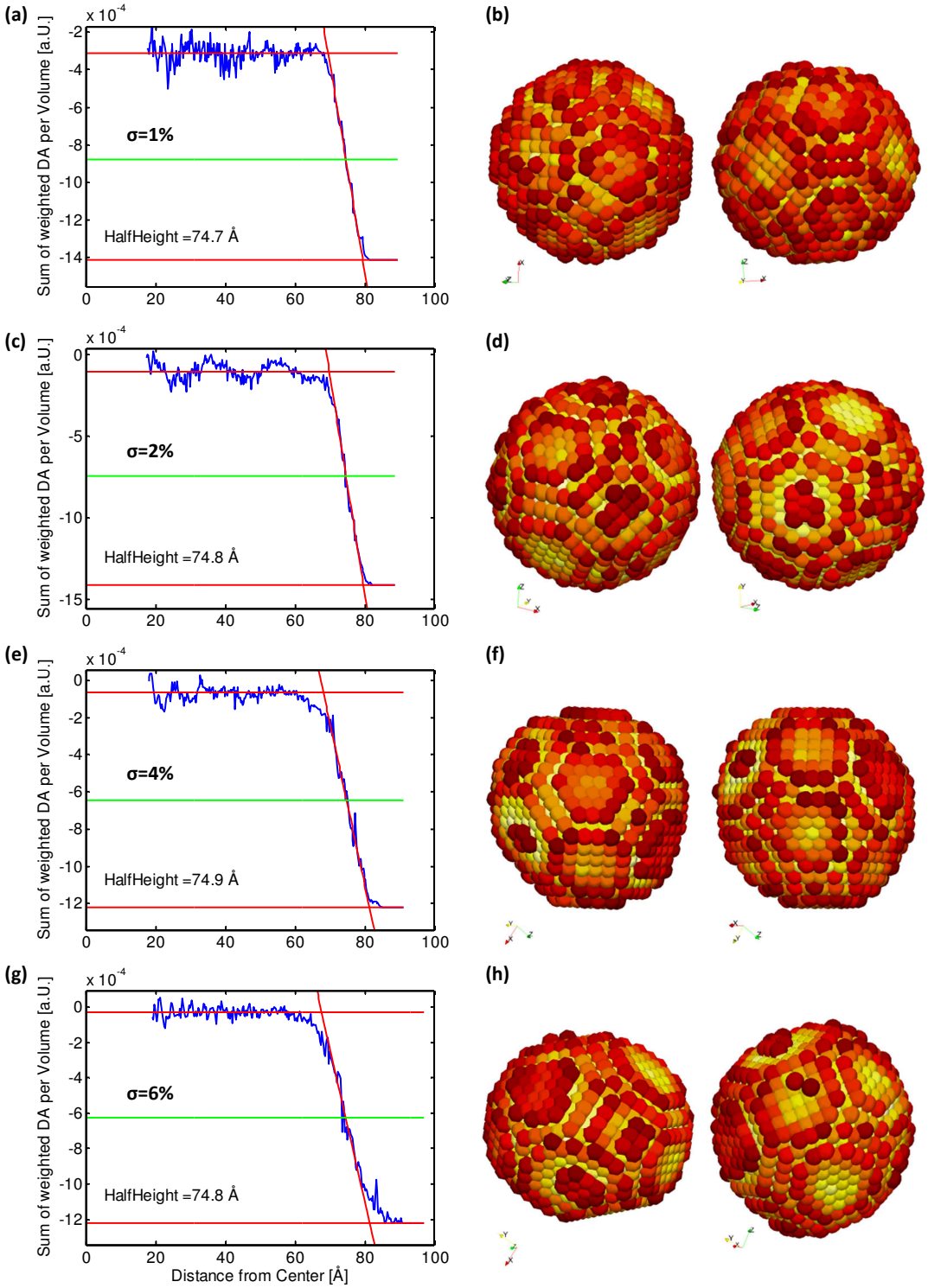


Figure 26: Results of the simulation series investigating the influence of increasing the polydispersity of the idealized system. **(a,c,e,g)** The spherical-average-evaluation of the DAMs shows no significant impact on the inner structure and the half-height-dimension. **(b,d,f,h)**: DAM shown from two perspectives. The color-gradient is a measure of the occupancy. **(a,b)** sphere_pd_1,**(c,d)** sphere_pd_2, **(e,f)** sphere_pd_4 and **(g,h)** sphere_pd_6.

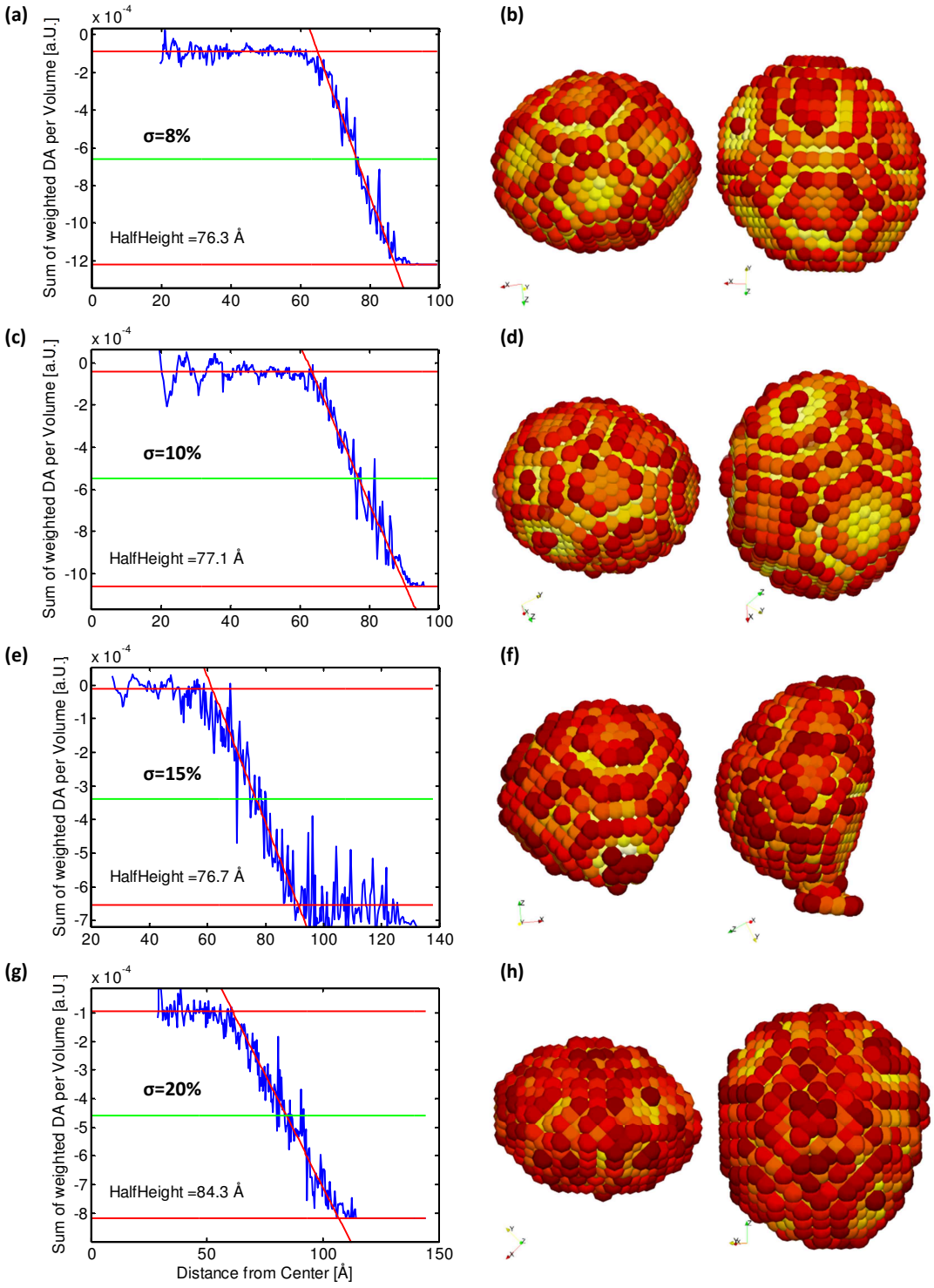


Figure 27: Continuation of **Figure 26**. **(a,c,e,g)** The spherical-average-evaluation of the DAMs a significant decrease of the slope of the half-height tangent. **(b,d,f,h)** DAM shown from two perspectives. Obviously at a certain threshold, the DAMs become unstable and start forming excrescence. **(a,b)** sphere_pd_8, **(c,d)** sphere_pd_10, **(e,f)** sphere_pd_15 and **(g,h)** sphere_pd_20.

Since throughout the simulation the initial search volume of DAMMIN had to be increased, the size of the DAs had to be adjusted to ensure that approx. the same number of DAs was used.

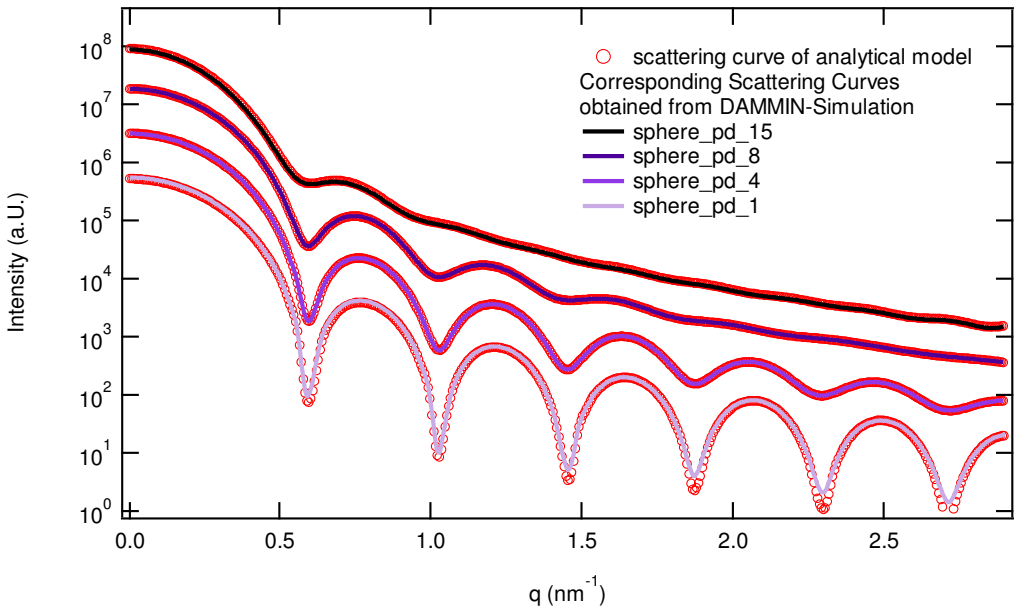


Figure 28: Some of the analytically computed scattering curves of systems of polydisperse spheres, used in the simulation series investigating the influence of polydispersity on the result, compared to the fits of the DAMMIN simulation. All simulations resulted in a perfect fit.

The results of the spherical-average-evaluation can be found in Figure 26(a,c,e,g) and Figure 27(a,c,e,g). As the polydispersity is increased there is no visible effect on the inner structure of the DAM. If however the obtained half-height-value is plotted against the polydispersity, a clear drift to higher dimensions is visible. This can be related to the fact, that the intensity correlates to the square of the particle volume (see the detailed explanation in Chapter *3.1.2 Polydispersity)*. The obtained values from the DAMs with a polydispersity below 8 % are almost identical. They deviate less than one percent from each other. A difference can only be seen when taking a closer look at the transition region between inner and outer particle occupancy where an increasing smearing occurs with increasing polydispersity (see Figure 26(a,c,e,g)). To visualize and quantitatively describe these two effects, the half-height-value and the half-height-tangent-slope are plotted against the polydispersity as shown in Figure 29. The behavior of the half-height-value affirms the suggestion made above. There is no significant deviation up to 6 % polydispersity.

If, however, the slope of the tangent line is calculated an inevitable change in exactly this region can be seen. Recall, that a decrease of this slope indicates a less abrupt transition of the mean occupancy, and therefore electron density, from DAM core to the surrounding. This implies that with increasing polydispersity this transition region is on average increasingly smeared out. If the models for which the obtained half-height-value was outside their error are neglected, a linear relationship between the slope and the polydispersity is suggested. Only at a polydispersity over 8 % a clear deviation from this assumption can be seen. Overall, the determination of this tangent slope might therefore yield a useful tool to estimate the polydispersity of a given system.

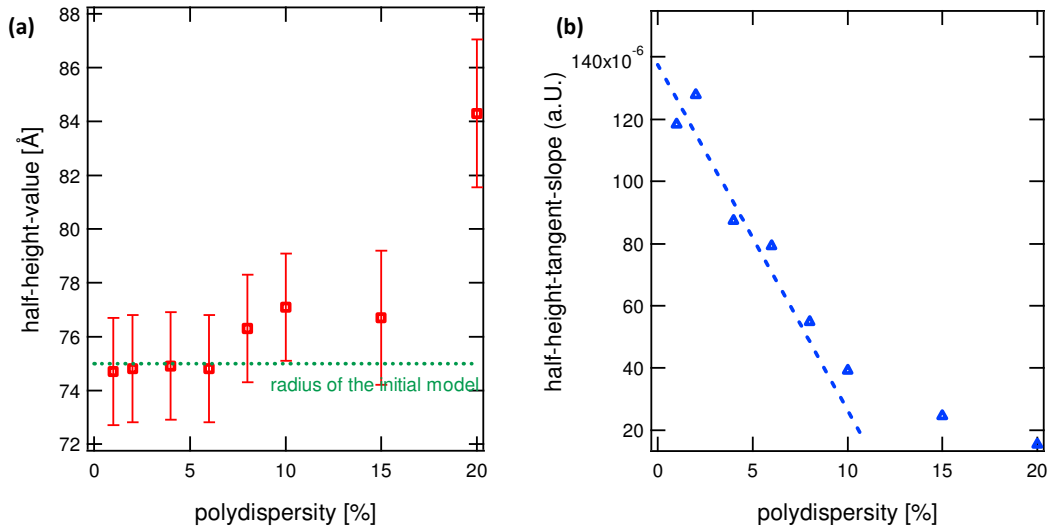


Figure 29: Results from the spherical-average- evaluation of the simulations investigating polydisperse spheres. **(a)** Obtained half-height-value plotted against the polydispersity. The green line indicates the original dimension of the model with diameter of 15 nm. The total error indicated is chosen to be the DA-diameter. **(b)** This shows the half-height-tangent-slope in dependence of the used polydispersity. The blue line indicates a linear correlation of the values up to a polydispersity of 8 %.

The visualized DAMs of all simulations can be found Figure 26(b,d,f,h) and Figure 27(b,d,f,h). For the first three models up to a polydispersity of 4 % there is no evidence of a deviation from the reference shape with 1 %. If the polydispersity is increased to 6 % (see Figure 26(g,h)) a one directional distortion of the initial shape is visible. This effect becomes more apparent at even higher polydispersities up to 10 %. (see Figure 27(a-d)). At this point the shape of the DAMs seems more similar to oblate ellipsoid than to a sphere. If the polydispersity is further increased to 15 and 20 % (see Figure 27(e-h)) the DAM becomes *unstable*.

In addition, all models were evaluated using the averaged-dimension method for which the results can be found in Figure 28. Especially for the simulations with a polydispersity

below 6 % a very distinct and clear peak is visible. The probability density of a perfect monodisperse sphere should look like an infinitely high streak with no width. Since DAMMIN is limited to finite dimensions, there will always be a peak broadening of some degree. When looking at sph_pd_1 in Figure 30 only 2 % of all measured dimensions are outside of 145±1 Å. Note that to gain good evaluation statistics, the size of the cut-trhough cylinder was chosen to be large which will therefore underestimate the real dimension. To quantitatively describe this peak, the FWHM was measured and subsequently, using the multiplier $2\sqrt{2\ln 2}$, a standard deviation of 1.2 % was calculated, which is very close to the initially used polydispersity of this system. This procedure was repeated for all simulations. The results can be found in Table 11. The standard deviation clearly corresponds to the polydispersity of the system and is up to 8 % in good agreement with the initial values.

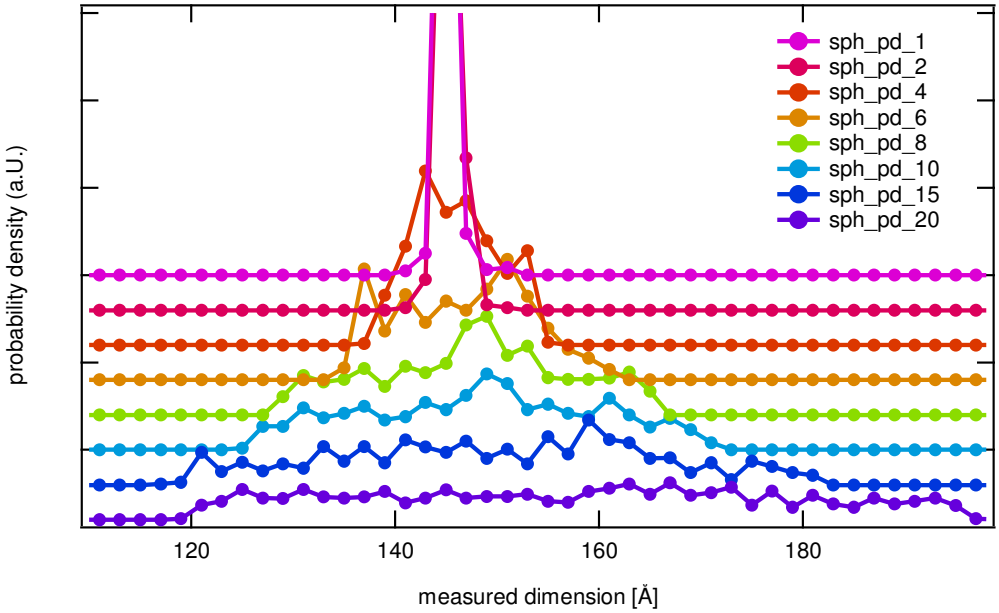


Figure 30: Averaged-dimension of the polydispersity-simulation-series of spheres using 1000 random directions and a varying cut-through diameter of 5 times the DA-diameter. The peak values of sph_pd_1 and sph_pd_2 were cut off to ensure good resolution for the high dispersed models.

From this investigation it can be seen that up to a polydispersity threshold level of 4 %, DAMMIN is able to fully retrieve the particle shape whereas there was no evidence of structural loss. Further, the methods presented in this thesis are very sensible to minimal deviations of the mean outer dimension. E.g. there is a strong correlation of the half-height-tangent-slope or the peak width of the averaged-dimension-evaluation and the size distribution of the system. For a polydispersity higher than 8 % the interpretation of the DAM is to be done with fortified skepticism to avoid

misinterpretation. In this case DAMMIN tends to form prolate and oblate ellipsoids of revolution to deal with the necessary distances to fit the scattering curve. This simulation series also shows that the presented evaluation methods, as well as the 3D-visualization of the DAM, should always be seen as complementary methods of interpretation and not as an exclusive one. Nevertheless, even for a size distribution of 10 % a sphere diameter of 15.4 nm is obtained which is a deviation of less than 3 % from the initial dimension of 3 %. If *a priori* knowledge of the particle system indicates a spherical shape, the spherical-average evaluation can therefore estimate particle dimension with a high degree of certainty, even for polydisperse systems.

Table 11: Results from the averaged dimensions peaks shown in **Figure 30**. The standard deviation σ was calculated from the FWHM using the scaling factor 2.355.

name	center [Å]	FWHM [Å]	σ [%]
sphere_pd_1	145	4	1.2
sphere_pd_2	146	6	1.8
sphere_pd_4	146	14	4.1
sphere_pd_6	147	20	5.8
sphere_pd_8	147	32	9.2
sphere_pd_10	148	40	11.5
sphere_pd_15	151	60	16.9
sphere_pd_20	158	70	19.9

5.2.1 Ellipsoid of Revolution

As it can be seen from the simulations performed on the system of polydisperse spheres, DAMMIN *deals* with high size distributions above 8 % by forming an ellipsoid of revolution. The question that arises from this phenomenon is, if the distinction between polydisperse ellipsoids of revolutions and polydisperse spheres can be made by DAMMIN. Therefore the geometry of an ellipsoid of revolution with a double axis of 8.5 nm and a single axis of 5.25 nm was chosen which corresponds to an aspect ratio of 1:1.6. The dimensions correlate with a sphere of 15 nm since they both have the same radius of gyration. The scattering curves for polydisperse ellipsoids were computed using a Gauss-curve with 21 steps (For more details see chapter *3.1.2 Polydispersity).* Since from the previous series on spheres no stable results were found above a polydispersity higher than 10 %, the attention was directed towards size distributions below this threshold. The simulation-names and the used polydispersity can be found in

Table 12. The analytically computed scattering curves as well as the corresponding DAMMIN fits can be found in Figure 31.

Table 12: Names and corresponding polydispersity of the simulations used to investigate the influence of a size distribution on the final results in the case of ellipsoids of revolution. The polydispersity is the standard deviation of the Gaussian-volume-distribution.

name	polydispersity [%]	name	polydispersity [%]
ell_pd_0	0	**ell_pd_6**	6
ell_pd_2	2	**ell_pd_10**	10

The results of the spherical-average-evaluation as well as the visualized DAMs of this series can be found in Figure 32(a,c,e.g) and Figure 32(b,d,f,h). In the result for ell_pd_0 a step in the transition region from core to outer dimension can be seen. The first deviation from the core-level is approx. at 5 nm which corresponds to the initial single axis. At approx. the half-height distance, a change or kink in the transition shape can be noticed. Due to this, the obtained half-height value cannot be related to any structural meaning. The corresponding DAM clearly shows an elliptical slope with values obtained from the linear-evaluation of 10.6 and 16.5 nm as already presented in Figure 14. When looking at the next model with a polydispersity increased to 2 %, no significant change in the density profile and the obtained DAM can be witnessed. Only when the size distribution is further increased to 6 and 10 %, a significant change can be witnessed. First of all, the density profile can now entirely be fitted with a single tangent which implies that the kink in slope mentioned above has disappeared. Furthermore, the obtained DAM now clearly deviates from the initial shape. Nevertheless, it is astonishing when looking at ell_pd_10 in Figure 31 that DAMMIN can still retrieve a shape that is similar to the initial model from this curve.

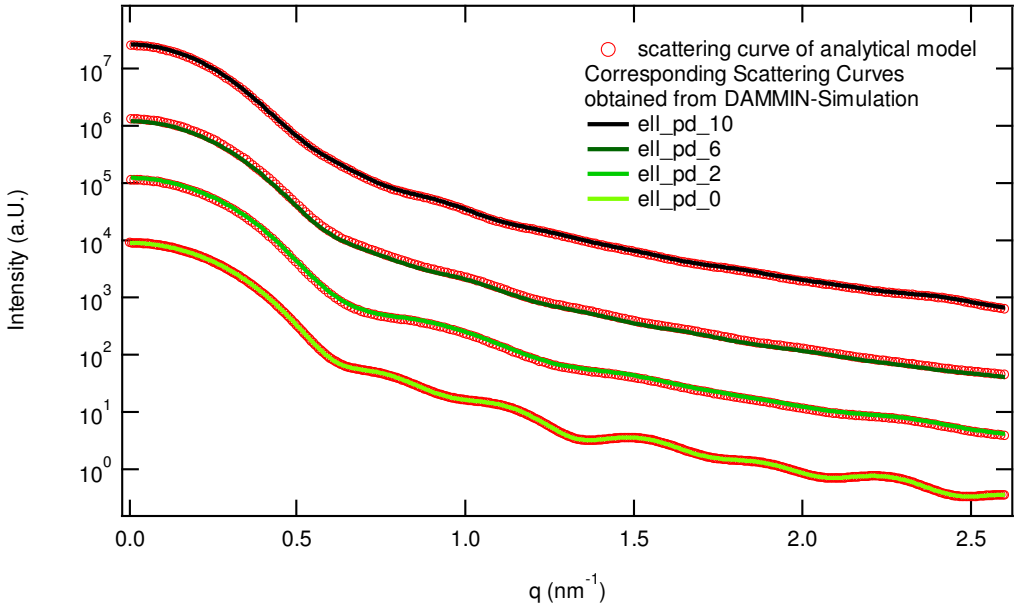


Figure 31: The analytically computed scattering curves of systems of polydisperse ellipsoids of revolution with an aspect ratio of 1:1.6, used in the simulation series investigating the influence of polydispersity on the result, compared to the fits of the DAMMIN simulation. All simulations resulted in a perfect fit.

Also for this series the DAM were analyzed using the averaged-dimension method where the results can be found in Figure 33. The minimum and maximum cut-off lengths of ell_pd_0 result in 10.4 and 16.2 nm which is in good with the initial values and the dimensions retrieved from the linear evaluation. An increase of the polydispersity to 2 and 6 % clearly broadens the range of the measured dimensions. Interestingly, the minimum cut-off dimension is totally stable and stays constant. The maximum cut-off dimensions increase with the polydispersity up to 6 %. Up to this point also the qualitative shapes of the distance-distributions look alike. Only at size-distributions of 10 % and more, the DAM becomes instable which is indicated by the fact, that no long dimensions of the DAM were measured.

The misinterpretation of such a scattering curve might therefore result in an error of up to 10 %. For the case of ell_pd_10 no relation of the maximum dimension to the previous models is observed. Anyhow, the qualitative shape as well as the minimum dimension of the distance distribution is constant and very stable, even for the case of ell_pd_10.

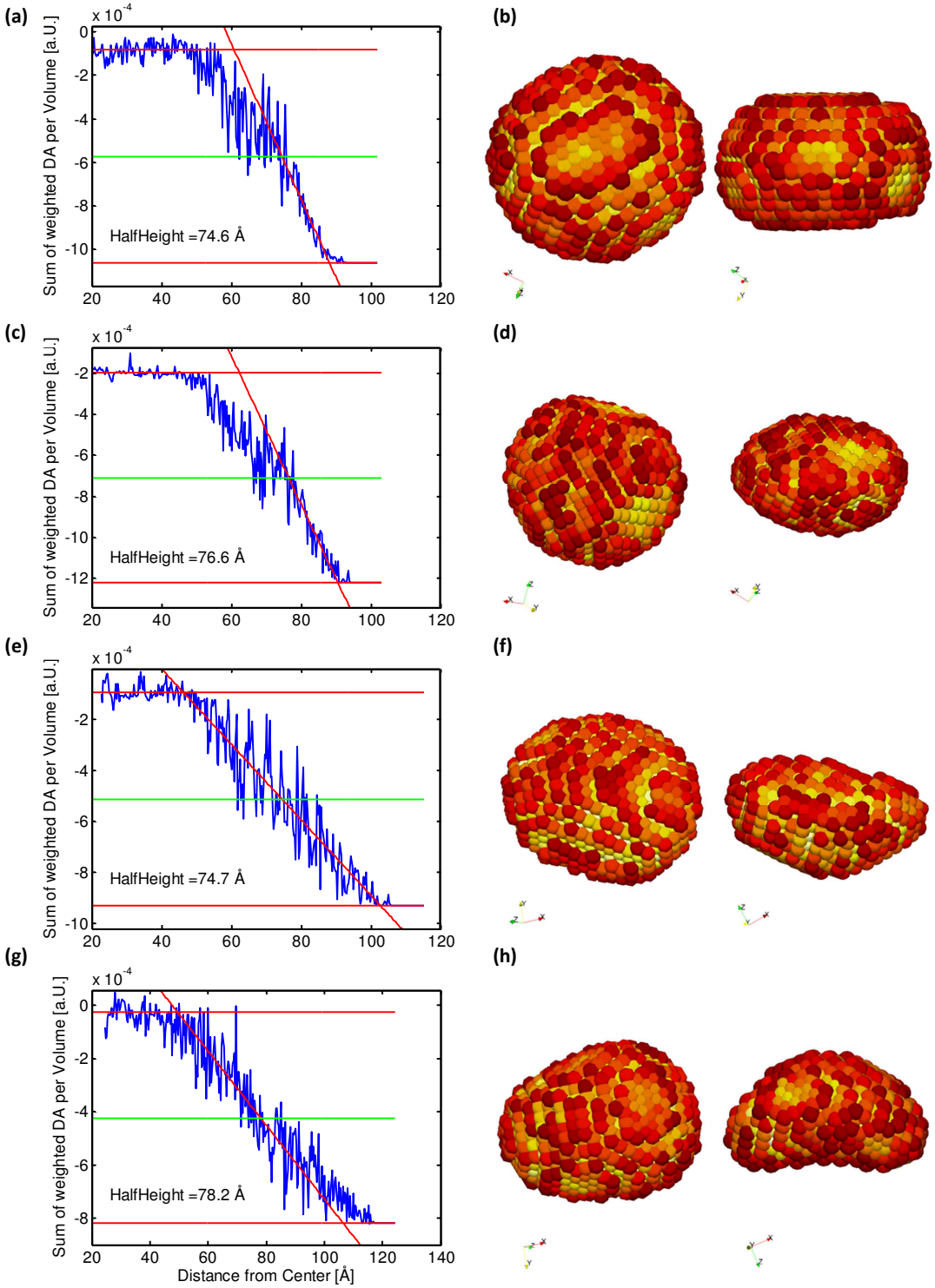


Figure 32: Results of the simulation series investigating the influence of increasing the polydispersity of a ellipsoids of revolution. **(a,c,e,g)** The spherical-average-evaluation of the obtained models. **(b,d,f,h)** DAM shown from two perspectives. The color-gradient is a measure of the occupancy. **(a,b)** ell_pd_0, **(c,d)** ell_pd_2, **(e,f)** ell_pd_6 and **(g,h)** ell_pd_10.

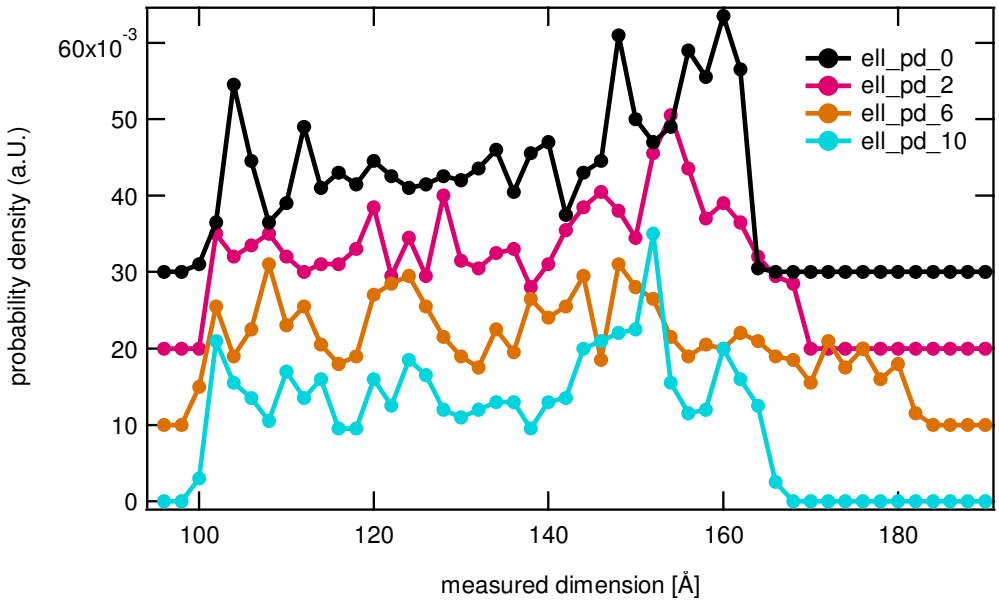


Figure 33: Averaged-dimension of the polydispersity-simulation-series of spheres using 1000 random directions and a varying cut-through diameter of 5 times the DA-diameter.

Overall, an increase of the size distribution to small values of 2 to 6 % still allows the qualitative interpretation of the obtained DAM. The values obtained by the mentioned methods should be treated careful. Even though the results of the perfectly monodisperse model are in very good agreement with the initial shape, the effects of size distribution produces a systematic error, that yields in the overestimation of the obtained dimensions. The same effect was witnessed with polydisperse spheres and can be contributed to the squared volume to intensity correlation.

Furthermore, the comparison of polydisperse spheres and polydisperse ellipsoids of revolution is in all cases possible when focusing on the transition region of the density profile as well as the averaged-dimension evaluation.

5.3 Estimation of Error

The following discussion is based on the assumption that all simulations were successful, or in other words that the error between the scattering curve of the DAM and the initial curve is negligible. From what can be seen above two sources of error can occur and where each one has to be discussed separately. First, as already mentioned, the simulation parameters as well as smearing effects due to polydispersity can introduce a

systematic error. Second, the described evaluation methods are accompanied by a statistical deviation whereas also a systematic error can occur. These two sources have to be discussed separately.

5.3.1 Evaluation Methods

When looking at the developed evaluation methods each of them has to be investigated separately concerning the resulting errors, since they have different algorithms behind them. The following discussion is based on the assumption that no *a priori* knowledge of the system is available.

First off, it is suggested to qualitatively interpret the obtained DAM and check it for success. This means that on one hand the initial scattering curve has to be compared with the fit resulting from the DAM. Further, the DAM should be visualized and checked for excrescent structures. If this is not the case the DAM can be described quantitatively according to the following steps.

To obtain an objective and quantitative description of the DAM the average-dimension-evaluation should be performed. This evaluation is based on the same principles than the singular dimension evaluation with the draw-back that the computation of a probability density in shape of a histogram limits the dimensional resolution: Since the obtained histogram simply states how often values in-between an upper and a lower boundary were measured (width of a histogram bar), the final result is degraded by these boundaries or in other words by the width of the histogram bars. Even though it is generally recommended to use at least 1000 random orientations, the use of more than 50 steps in a histogram strongly distorts the result by an increasing noise to signal ratio Also with this method, it is highly recommended to use the tangent-extrapolation-method even though, the determination of the dimension from the histogram may be accompanied by a higher error than by measuring along this direction using the linear-evaluation-method. The final error is therefore mainly subject to the step width, and cannot be estimated for the general case. Nevertheless, it does not exceed the DA-diameter.

If the dimension-distribution obtained from the average-dimension-evaluation as well as the visualized DAM show some kind of characteristic dimensions (e.g. long and short axis of an ellipsoid, facet-to-facet distance, etc.) this dimension can be measured using the linear evaluation method. The computation of the linear-dimension-evaluation is the most sensitive method presented in this thesis. As already mentioned in Chapter *4.2*

Linear Evaluation the choice of the cut-through-cylinder-radius has a strong effect on the obtained results. On one hand, too big radii lead to the introduction of size effects and therefore to the underestimation of the real dimension which is per definition a systematic error. But, due to the high amount of DAs that can be found in the evaluation volume the fits are very stable, which results in a low statistical error of less than 1 % for a radius of 5∗DA-diamter (standard deviation). On the other hand, a decrease of the cut-through radius to 1∗DA-diamter can lead to a standard deviation of up to 5 % but also a decrease of the systematic error. It is therefore recommended to always use the linear extrapolation method presented in Figure 14. Using this method, the systematic error is compensated which leaves a statistical error that is now subject of the linear regression. In the case of e.g. Figure 14 the error of the intersection-value of tangent and vertical axis is 3.15 %, calculated using standard error propagation.

If the previous evaluations, as well as *a priori knowledge,* suggest a spherical shape, the spherical average evaluation can be performed. For, all other geometries this technique should only be used to determine some kind of significant change in structure, such as the estimation of the smallest dimension.

If the obtained model really is spherical, the results from the half-height-method are very stable. Since the fitting boundaries for the tangent are subject to manual input, the error depends mainly on the user. The coefficient of determination of the linear regression is in all cases above 99 % percent and therefore this error is negligible. The determination of the ratio of occupancy, or core-shell-electron-density-ratio, (see results from Chapter 5.1.2 Shannon Channels) the results are strongly subject to high fluctuations of the core density. This can easily be explained when recalling that the number of DAs in a sphere volume correlates to the third power of the volume. Therefore, for small search volumes at the core of the model the mean occupancy can only be calculated over a relatively small number of DAs which automatically results in bad statistics. The obtained density ratio should therefore always be expressed with a margin of approx. 10 %.

5.3.2 Simulation parameters

The investigations above point out that the choice of wrong simulation parameters can easily lead to the misinterpretation of the real shape. As already mentioned all of the evaluation methods used in this thesis should be used to qualitatively check if the obtained model is plausible by comparing it to *a priori* data.

A method to check if the obtained models are stable is the variation of the simulation parameters for single simulations and not the averaged model. If e.g. the number of DA is increased the shape of the DAM should not significantly differ from the previous model. If it does, a higher resolution might be needed since some structural detail cannot be retrieved. It is therefore recommended to at least run a single simulation with a high number of DAs to check for any effects.

The series on polydisperse particles shows that smearing of data can lead to a distortion of the DAM. To be sure that all of the instrumental smearing effects are accounted for, the IFT has to be done using the experimental smearing data. To retrieve the actual instrumental resolution calibration measurements of the experimental setup have to be made. Still, especially with inorganic NCs, there will always be some kind of a finite size distribution which will also lead to biased model. The averaged-dimension-evaluation is an advised method to see if there are still characteristic dimensions observable or not.

Another topic about the experimental setup has to be addressed: all of the investigations above were made using $q_{min} = 0$, which in real experiments cannot be achieved. On the other hand chapter *5.1.2 Shannon Channels* shows that effects can occur when the information content in the scattering curve is decreased by lowering q_{max}. This leads to a conflict of interest when it comes to the experimental setup where on one hand q_{min} is to be minimized and on the other hand q_{max} is to be maximized. Anyhow, since the limits of the experimental setup can only in be changed in rare cases, the setup is to be chosen to keep q_{min} within the first SC. If the accessible q-range is now smaller than 10 SCs, the interpretation of inner structural details should be done very careful.

6 Conclusion and Outlook

In the present work the *ab initio* shape retrieval method DAMMIN, which was originally developed for the determination of strictly monodisperse systems such as proteins, is used on theoretical models of inorganic nanocrystal systems. DAMMIN is based on the refinement of a three-dimensional bead-model from which a scattering curve is calculated and fit against an experimental curve using the *multipole expansion.* Since any arbitrary shape can be described by an assembly of spheres, the presented method needs no *a priori* knowledge of the shape of the measured system.

To explicitly describe the obtained models several evaluation methods were developed and tested, at which all of them allow the interpretation of models from single simulations as well as from averaged simulations. Nevertheless, the main focus is set on the retrieval of structural details from the averaged results, since the averaging process introduces the so called *bead occupancy.* By comparing the radius of gyration of the obtained model with the initial input, this *bead occupancy* can be correlated to the electron density. The averaged results are therefore three-dimensional electron density maps in real space.

The first objective of this work was to check the simulation process for the stability of the results. This could only be done by utilizing the presented evaluation techniques, which allowed to qualitatively and quantitatively describe the obtained models. As shown in the presented investigations, the simulation results are very stable. Even, if only the minimal recommended parameters are used, the particle shape can be retrieved qualitatively. If, however, the detailed dimensions are to be determined, the number of DAs as well as the number of used harmonics is to be kept as high as possible. When looking at inner details of the shape, such as variations of the electron density (e.g. core-shell-particles), the accessible q-range is to be maximized. However, a qualitative interpretation can still be made if only 2-4 minima in the scattering curve are visible. An overview of all models used in this work, their scattering curve-fits and the corresponding three-dimensional models from the DAMMIN simulations are shown in Figure 34. The obtained dimensions are in great agreement with the initial values.

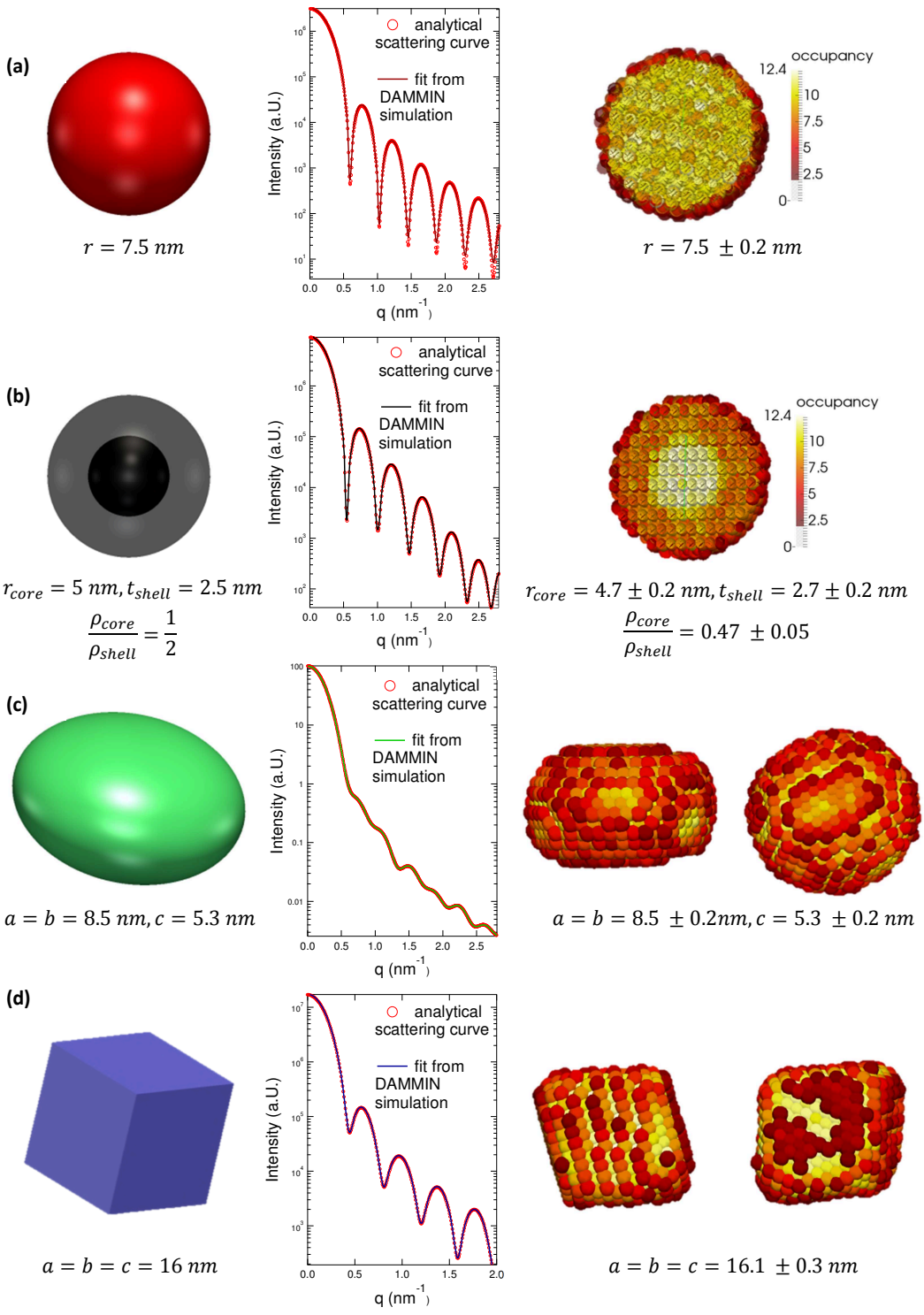


Figure 34: Conclusive overview of the best results from the DAMMIN simulations of the models used in this thesis. In the first column the initial model with its dimensions is presented. In the second column the analytical scattering curve is compared to the DAMMIN simulation. The third column shows the 3D DAM as well as the retrieved dimensions. The results were taken from simulations **(a)** sp_pd_1, **(b)** cs_sc_20, **(c)** ell_pd_0 and **(d)** cube_h_20.

The second objective was to determine the influence of smearing effects such as instrumental smearing and polydispersity on the ability to retrieve the particle shape. Even though, DAMMIN was originally developed for strictly monodisperse systems, the size and shape of the obtained models did not change up to Gaussian size distributions of up to 4 %. Using the presented evaluation techniques, characteristic dimensions such as e.g. the-smallest-distance-measured or a deviation from the spherical shape can be retrieved from systems with size-distributions of 10 % and even higher. Interestingly, for the case of spheres, the initial scattering curve of perfectly monodisperse spheres had to be artificially smeared which was done using a size-distribution of 1 %.

Last but not least, an estimation of the error was made. If the retrieved model is interpreted as suggested (see *Chapter 5.3.1 Evaluation Methods)* the obtained particle dimensions can be determined within the range of one *dummy atom diameter.* For most applications this results in an error of approx. 2 Å. By using DAMMIN together with the presented evaluation techniques, the detailed shape of nanocrystals can be obtained close to the atomic resolution.

Future investigations will focus on the application of the presented methods on real NC-systems.

Another suggestion is to directly implement the atomic crystal lattice into the DAMMIN simulations. This would allow the construction of an idealized atomistic mean shape of the NCs. As a second possibility, the reconstruction of an atomistic model from from the obtained DAMs

References

1. Cosgrove, T., *Colloid science : principles, methods and applications*. 2nd ed. 2010, Chichester, West Sussex: Wiley. xiv, 375 p.
2. Talapin, D.V., *Nanocrystal solids: A modular approach to materials design.* Mrs Bulletin, 2012. **37**(1): p. 63-71.
3. Piliego, C., et al., *5.2% efficient PbS nanocrystal Schottky solar cells.* Energy & Environmental Science, 2013. **6**(10): p. 3054-3059.
4. Sun, Y.G. and Y.N. Xia, *Shape-controlled synthesis of gold and silver nanoparticles.* Science, 2002. **298**(5601): p. 2176-2179.
5. Yin, Y. and A.P. Alivisatos, *Colloidal nanocrystal synthesis and the organic-inorganic interface.* Nature, 2005. **437**(7059): p. 664-670.
6. Burt, J.L., et al., *Beyond Archimedean solids: Star polyhedral gold nanocrystals.* Journal of Crystal Growth, 2005. **285**(4): p. 681-691.
7. Yin, Y. and P. Alivistatos, *Colloidal Nanocrystal synthesis and the organic-inorganic interface.* Nature, 2005. **437**.
8. Dayal, S., et al., *The Effect of Nanoparticle Shape on the Photocarrier Dynamics and Photovoltaic Device Performance of Poly(3-hexylthiophene):CdSe Nanoparticle Bulk Heterojunction Solar Cells.* Advanced Functional Materials, 2010. **20**(16): p. 2629-2635.
9. Gao, L. and Q.H. Zhang, *Effects of amorphous contents and particle size on the photocatalytic properties of TiO2 nanoparticles.* Scripta Materialia, 2001. **44**(8-9): p. 1195-1198.
10. Kelly, K.L., et al., *The optical properties of metal nanoparticles: The influence of size, shape, and dielectric environment.* Journal of Physical Chemistry B, 2003. **107**(3): p. 668-677.
11. Bratlie, K.M., et al., *Platinum nanoparticle shape effects on benzene hydrogenation selectivity.* Nano Letters, 2007. **7**(10): p. 3097-3101.
12. Lee, S.M., S.N. Cho, and C. Jiwoo, *Anisotropic Shape Control of Colloidal Inorganic Nanocrystals.* Advanced Materials, 2003. **5**(15).
13. Tao, A.R., S. Habas, and P.D. Yang, *Shape control of colloidal metal nanocrystals.* Small, 2008. **4**(3): p. 310-325.
14. Huang, M.H. and P.H. Lin, *Shape-Controlled Synthesis of Polyhedral Nanocrystals and Their Facet-Dependent Properties.* Advanced Functional Materials, 2012. **22**(1): p. 14-24.
15. Hunter, R.J., *Foundations of colloid science*. 2nd ed. 2001, Oxford ; New York: Oxford University Press. xii, 806 p.
16. Rambo, R.P. and J.A. Tainer, *Super-Resolution in Solution X-Ray Scattering and Its Applications to Structural Systems Biology.* Annual Review of Biophysics, Vol 42, 2013. **42**: p. 415-441.
17. Svergun, D. *Data analysis software ATSAS 2.5.2*. 2014 [cited 2014; Available from: http://www.embl-hamburg.de/biosaxs/software.html.

18. Svergun, D.I., *Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing.* Biophysical Journal, 1999. **76**(6): p. 2879-2886.
19. Petoukhov, M.V. and D.I. Svergun, *Global rigid body modeling of macromolecular complexes against small-angle scattering data.* Biophysical Journal, 2005. **89**(2): p. 1237-1250.
20. Shtykova, E.V., et al., *Structure and properties of iron oxide nanoparticles encapsulated by phospholipids with poly(ethylene glycol) tails.* Journal of Physical Chemistry C, 2007. **111**(49): p. 18078-18086.
21. Als-Nielsen, J.M.D., *Elements of modern X-ray physics*. 2001, New York: Wiley.
22. Erko, M., *Water Properties in Confined Geometry*, in *Institute of Physics*. 2012, University of Leoben: Leoben.
23. Lindner, P. and T. Temb, *Neutrons, X-Rays and Light - Scattering Methods Applied to Soft Condensed Matter*. 2002, Amsterdam: Elsevier.
24. Debye, P., *X-ray dispersal.* Annalen Der Physik, 1915. **46**(6): p. 809-823.
25. Glatter, O. and O. Kratky, *Small angle x-ray scattering*. 1982, London ; New York: Academic Press. 515 p.
26. Schnablegger, H. and Y. Singh, *The SAXS Guide: Getting acquainted with the principles*. 2011, Austria: Anton Paar.
27. Weyerich, B., J. Brunner-Popela, and O. Glatter, *Small-angle scattering of interacting particles. II. Generalized indirect Fourier transformation under consideration of the effective structure factor for polydisperse systems.* Journal of Applied Crystallography, 1999. **32**: p. 197-209.
28. Kirkpatrick, S., C.D. Gelatt, Jr., and M.P. Vecchi, *Optimization by simulated annealing.* Science, 1983. **220**(4598): p. 671-80.
29. Ingber, A.L., *Simulated annealing: Practice versus Theory.* Mathl. Comput. Modelling, 1993. **18**(11): p. 29-57.
30. Stuhrmann, H.B., *Interpretation of Small Angle Scattering Functions of Dilute Solutions and Gases.* Acta Cryst., 1970. **26**: p. 297-306.
31. Svergun, D.I., *Solution Scattering from Biopolymers: Advanced Contrast-Variation DataAnalysis.* Acta Cryst., 1994. **50**: p. 391-402.
32. Glatter, O., *Data Evaluation in Small-Angle Scattering - Calculation of Radial Electron-Density Distribution by Means of Indirect Fourier Transformation.* Acta Physica Austriaca, 1977. **47**(1-2): p. 83-102.
33. Glatter, O., *New Method for Evaluation of Small-Angle Scattering Data.* Journal of Applied Crystallography, 1977. **10**(Oct1): p. 415-421.
34. Glatter, O., *Convolution Square Root of Band-Limited Symmetrical Functions and Its Application to Small-Angle Scattering Data.* Journal of Applied Crystallography, 1981. **14**(Apr): p. 101-108.
35. Glatter, O. and B. Hainisch, *Improvements in Real-Space Deconvolution of Small-Angle Scattering Data.* Journal of Applied Crystallography, 1984. **17**(Dec): p. 435-441.
36. BrunnerPopela, J. and O. Glatter, *Small-angle scattering of interacting particles .1. Basic principles of a global evaluation technique.* Journal of Applied Crystallography, 1997. **30**: p. 431-442.

37. Bergmann, A., G. Fritz, and O. Glatter, *Solving the generalized indirect Fourier transformation (GIFT) by Boltzmann simplex simulated annealing (BSSA).* Journal of Applied Crystallography, 2000. **33**: p. 1212-1216.
38. Fritz, G., A. Bergmann, and O. Glatter, *Evaluation of small-angle scattering data of charged particles using the generalized indirect Fourier transformation technique.* Journal of Chemical Physics, 2000. **113**(21): p. 9733-9740.
39. Kohlbrecher, J. and I. Bressler, *SASfit*. 2014, PSI: Villigen, CH.
40. Ilavsky, J. and P.R. Jemian, *Irena: tool suite for modeling and analysis of small-angle scattering.* Journal of Applied Crystallography, 2009. **42**: p. 347-353.
41. Fritz-Popovski, G., et al., *GIFT*. 2011, TU Graz: Graz, Austria.
42. Schneidman-Duhovny, D., et al., *Accurate SAXS Profile Computation and its Assessment by Contrast Variation Experiments.* Biophysical Journal, 2013. **105**(4): p. 962-974.
43. Svergun, D., C. Barberato, and M.H.J. Koch, *CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates.* Journal of Applied Crystallography, 1995. **28**: p. 768-773.
44. Popovski, G., *PCG Singlebody*. 2013: Leoben.
45. Chacon, P., et al., *Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm.* Biophysical Journal, 1998. **74**(6): p. 2760-2775.
46. Walther, D., F.E. Cohen, and S. Doniach, *Reconstruction of low-resolution three-dimensional density maps from one-dimensional small-angle X-ray solution scattering data for biomolecules.* Journal of Applied Crystallography, 2000. **33**: p. 350-363.
47. Svergun, D.I., M.V. Petoukhov, and M.H.J. Koch, *Determination of domain structure of proteins from X-ray solution scattering.* Biophysical Journal, 2001. **80**(6): p. 2946-2953.
48. Franke, D. and D.I. Svergun, *DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering.* Journal of Applied Crystallography, 2009. **42**: p. 342-346.
49. Svergun, D.I., *Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing.* Biophys J, 1999. **76**(6): p. 2879-86.
50. Pedersen, J.S., *Analysis of small-angle scattering data from colloids and polymer solutions: modeling and least-squares fitting.* Advances in Colloid and Interface Science, 1997. **70**: p. 171-210.
51. Buss, S.R. and J.P. Fillmore, *Spherical averages and applications to spherical splines and interpolation.* Acm Transactions on Graphics, 2001. **20**(2): p. 95-126.
52. Fritz, G. and O. Glatter, *Structure and interaction in dense colloidal systems: evaluation of scattering data by the generalized indirect Fourier transformation method.* Journal of Physics-Condensed Matter, 2006. **18**(36): p. S2403-S2419.
53. Shannon, C.E., *A Mathematical Theory of Communication.* Bell System Technical Journal, 1948. **27**(3): p. 379-423.
54. Moore, P.B., *Small-Angle Scattering - Information-Content and Error Analysis.* Journal of Applied Crystallography, 1980. **13**(Apr): p. 168-175.
55. Volkov, V.V. and D.I. Svergun, *Uniqueness of ab initio shape determination in small-angle scattering.* Journal of Applied Crystallography, 2003. **36**: p. 860-864.
56. Svergun, D.I. and M.H.J. Koch, *Small-angle scattering studies of biological macromolecules in solution.* Reports on Progress in Physics, 2003. **66**(10): p. 1735-1782.

57. Trewhella, J., et al., *Report of the wwPDB Small-Angle Scattering Task Force: Data Requirements for Biomolecular Modeling and the PDB.* Structure, 2013. **21**(6): p. 875-881.
58. Mittelbach, R. and G. Porod, *Zur Röntgenkleinwinkelstreuung kolloider Systeme.* Acta Physica Austriaca, 1962. **15**: p. 25.
59. Paris, O., *Innere Oxidation von Cu-Fe Legierungen untersucht mittels Röntgenstreumethoden am Drehanodengenerator und am Synchrotron*, in *Formal- und Naturwissenschaftliche Fakultät*. 1992, Universität Wien: Wien.
60. Wignall, G.D., *Instrumental Resolution Effects in Small-Angle Scattering.* Journal of Applied Crystallography, 1991. **24**: p. 479-484.
61. Fratzl, P., et al., *Mineral Crystals in Calcified Tissues - a Comparative-Study by Saxs.* Journal of Bone and Mineral Research, 1992. **7**(3): p. 329-334.
62. de la Cuesta, F.B., et al., *Collagen imaged by Coherent X-ray Diffraction: towards a complementary tool to conventional scanning SAXS.* Xiv International Conference on Small-Angle Scattering (Sas09), 2010. **247**.
63. Tikhonov, A.N., *On the stability of inverse problems.* Comptes Rendus De L Academie Des Sciences De L Urss, 1943. **39**: p. 176-179.
64. Svergun, D.I., *Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria.* Journal of Applied Crystallography, 1992. **25**: p. 495-503.
65. Kitware, *ParaView*. 2014, Open Source.
66. Kozin, M.B. and D.I. Svergun, *Automated matching of high- and low-resolution structural models.* Journal of Applied Crystallography, 2001. **34**: p. 33-41.
67. Svergun, D. *DAMMIN Manual*. 2009 [cited 2014 May]; Available from: http://www.embl-hamburg.de/biosaxs/manuals/dammin.html.