Simulation Data Comparison of GISAXS measurements

Felix Hoffmann, University of Wuerzburg, Germany

MiNaSX P03 beamline PETRA III

Supervisor: Dr. Matthias Schwartzkopf

September 4, 2019

Abstract

An Algorithm is designed that automatically compares a measured GISAXS image with a database, to find the best correlation. Therefore a process chain is developed that prepare the images for the comparison. Different methods of comparison where implemented and tested, as well as different methods to probe the database for the best correlation in a efficient way. To test the system the growth kinetics of Au on Si during the sputter process is investigated.

Contents

1 Introduction

Grazing Incidence Small Angle X-ray Scattering (GISAXS) provides a method to probe the surface morphology with high resolution on the nano scale, as well as a time resolution of milliseconds. Nanostructured surfaces open a wide field of new or enhanced properties, such as metalpolymer nanocomposites [5], plasmonic and magnetic applications or ultradense magnetic data storage [9]. A deep understanding of the growth processes of these nanostructures is necessary to make further progress in this fields. The MiNaXs P03 group at PETRA III performs nano and micro focused GISAXS measurements as static and real time investigation [3]. To increase the information that can be taken from GISAXS measurements, an Algorithm is designed that compares experimental data with a database of simulated data to get a more detailed understanding on the surface morphology.

2 Theory

The method of grazing incidence small angle x-ray scattering was introduced by J.R. Levine and J.B. Cohen [6] as a tool for surface sensitive thin film investigations of Au on SiOx. Also the first GISAXS real time investigation of a growth process was performed on Au on Si [1].
(GISAXS) allows insight into the form and in-plane arrangement of the particles. Nowadays, such experiments produce a huge amount of data which cannot be evaluated manually anymore. Hence, they are looking for a way to automatise the evaluation process. The scattering images are the convolution of the form factor F (describing the particle shape) and the structure factor S (describing the in-plane arrangement of the particles). The form factor is the Fourier transform of the particles electron density correlation function and the structure factor is the Fourier transform of the particle correlation function. Additionally refraction and reflections have to be taken into account. The analysis of GISAXS experiments requires the so-called distorted-wave Born approximation (DWBA). Hence, a simple inversion of the Fourier transform is not possible. Moreover, the involved physics and the resulting equations are very complex so that an analytical evaluation is hardly possible for low correlated systems. The current workflow comprises three steps: GISAXS experiment, GISAXS simulation(s) and the use of other real space surface probing techniques that might give a clue about the surface geometry. Nevertheless, experimental GISAXS data is mainly evaluated by performing simulations with software packages like IsGISAXS [8] or BornAgain [7] and by varying the simulation parameters as long as a suitable match is found. Unsurprisingly, this is a rather tedious and time consuming task when done manually.

3 SiDaCo (Simulation Data Comparison)

In order to ease the process, a database with structure and form factors was established. This database provides an easy way to cut down the simulation times as one only needs

to perform n simulations with varying structure and form factors to gain access to n^2 GISAXS intensity outputs. This is due to the fact that the Intensity I is proportional to the product of the structure and form factor in the Fourier space. One may write the measured intensity in the so-called decoupling approximation

$$I(q) = S(q) \cdot |F(q)|^2 \tag{1}$$

with F being the form factor of an average nanoparticle taking into account refraction and reflections (see fig 1). Hence, a database with each n structure and form factors (obtainable from IsGISAXS simulations) gives rise to n^2 combinations and, therefore, GISAXS intensity distributions. The aim of this report is to use this database and compare it to measured GISAXS pattern to find the most suitable match, which is describing the system.

3.1 Database

The database contains data sets of the form factor F and structure factor S simulated with IsGISAXS, for Au particles on Si. The form factor is simulated for hemispherical geometry of the particle, for a given average radius R and deviation of the radius σ, as the FHWM of a Gaussian distribution. The structure factor is related to the average distance of two particles D and the FHWM of the Gaussian particle distance distribution ω. The particles are arranged on a hexagonal lattice, with an average lattice distance D. Therefore every simulation is characterized by its unique combination of R, σ, D, ω.

	R	D	σ/R	ω/D
range	0-7.5 nm	0-15 nm	0-0.5	0-0.5
step size	0.1 nm	0.2 nm	0.01	0.01

Table 1: Particle and lattice parameter for the given database.

Furthermore the database contains information about the beam and detector properties. Since the database only contains simulation for the system Au on Si, all beam and detector parameter a listed in tab. 3.1. Where λ, the wave length of the incoming x-ray

λ	α_i	$2\theta_{min}$	$2\theta_{max}$	$\alpha_{f,min}$	$\alpha_{f,max}$	n_1	n_2
0.0953 nm = 13 keV	0.4 deg	0 deg	5 deg	0 deg	5 deg	500	500

Table 2: Beam and detector parameter for the given database.

beam, and α_i, the angle of incidence of the incoming beam relative to the surface plane, describing the beam parameter. $2\theta_{min,max}$, the minimal and maximal (respectively) in plane angle, $\alpha_{f,min,max}$, the minimal and maximal (respectively) out of plane angle and

n_1 and n_2, the number of pixel, describing the detector parameter.
The database is build as a list of 2768 objects, where each object contains one form and one structure factor, as well as the corresponding parameters R, D, σ and ω, which a listed in tab.3.1. Each object is addressed by a key. At the beginning the database has to be converted, so that the keys can be accessed by a set of given parameter, rather then that a set of parameter is accessed by a set of keys. Therefore an four dimensional array is created, where each axis is described by one parameter and each element of the array contains two keys, one for the structure factor and one for the from factor. here show two pictures of ff and sf each for different parameter and corresponding gisaxs pattern.

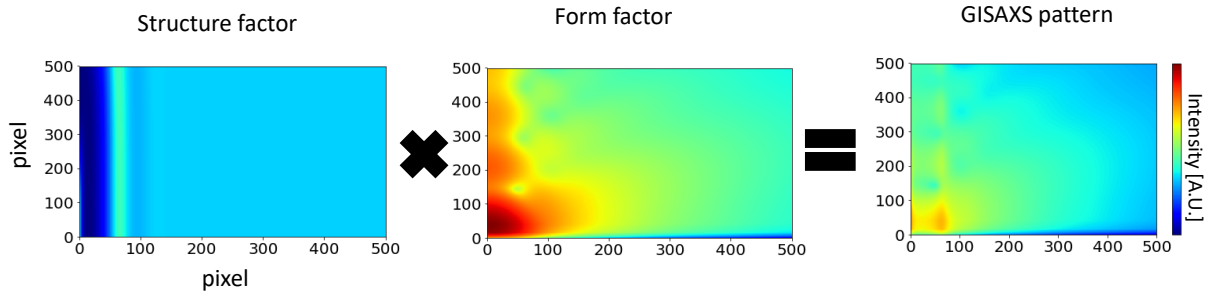

Figure 1: left: structure factor, middle: form factor, right: GISAXS intensity pattern. Simulated for hemispherical Au cluster on Si, with $R = 4.6\,\text{nm}$, $D = 10.2\,\text{nm}$, $\sigma = 0.37$, $\omega = 0.23$.

3.2 Program Layout

The program consists of three major parts. Firstly (i) the preparation of both, simulated and measured data sets, which is done by

- crop data sets to same size in q-space
- bin the data sets to same pixel size
- mask and normalize simulation.

Secondly (ii) the data sets have to be compared by defining a quantity which represents the correlation among these data sets and thirdly (iii) a method have to be created, which simulation is compared next, depending on the correlation of the previous comparison. In fig.2 the layout of the program is visualized. It shows the order in which it handles the input data and how it is processing it. The operations within the orange-dashed box do account to the image preparation. Here the mask is also considered as exp. data, since it has to processed the same way as the experimental GISAXS data. The operation outside this box belong to the image comparison as well as the movement in parameter space. All important steps will be explained in the following chapters.

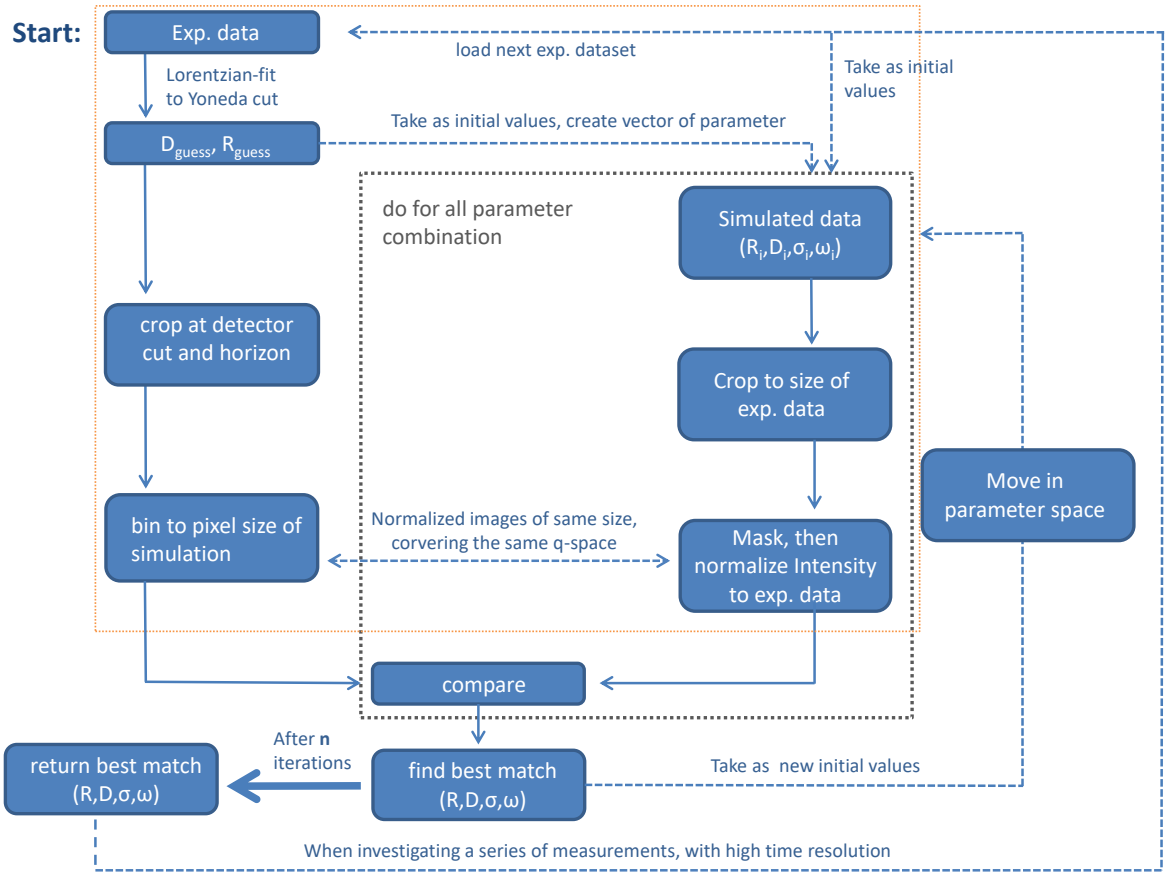

Figure 2: Layout and process chain of SiDaCo.

3.3 Prepare images

Before the simulations can be compared with the measured GISAXS pattern, they need to be prepared for a quantitative comparison. The cropping is done by the *crop_to_axes*() function. The inputs are the experimental data (dat), the mask-file (mask), which is of the same size as the experimental data, the form factor (sim1), the structure factor (sim1). Since the GISAXS pattern is symmetric to the detector cut and only the region above the horizon is simulated, the experimental data and mask is cropped at these axes (see fig 3a, red dashed lines). The simulation is then cropped to the same size in q-space. Now the experimental data and simulation needs to have the same size (size of the corresponding array). By comparing the size of the arrays, the algorithm decides which array needs to be binned. By calling the *bin_array*() function, the input *array_to_bin* is binned to the size of the input *array_bin_to*. Since the fraction of pixel sizes of simulation and experiment is usually a non integer, the binning is weighted, depending on the interval one pixel of each, experimental or simulation, represents in q-space or deg-space.
Now the simulation is masked. The intensity of experimental data is in units of counted

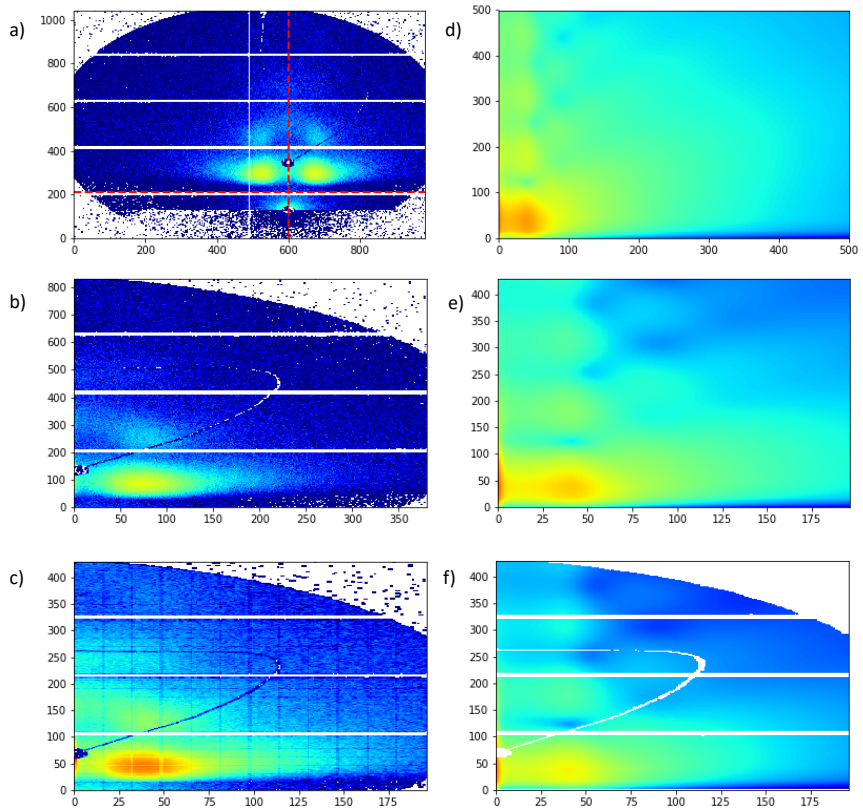

Figure 3: Overview of image preparation steps. Left column: measured data, Right column: simulated data. a) raw experimental data with crop position at detector cut and horizon (red, dashed lines). b) exp. data after cropping. c) exp. data after binning. d) raw simulated data. e) simulated data after cropping. f) normalized, masked and cropped simulated data.

photons, whereas the intensity of simulation is is normalized to 1. Therefor the simulation is normalized to the Intensity of the experimental data by

$$I_{sim}(q_y, q_z) = I_{sim}(q_y, q_z) \cdot \frac{\sum_{q_y,q_z} I_{exp}(q_y, q_z)}{\sum_{q_y,q_z} I_{sim}(q_y, q_z)}. \tag{2}$$

3.4 Compare images

At this stage the comparison is done by comparing the Yoneda-cut and the off-detector-cut at the $q_{y,max}$ position, where the Yoneda peak is. To do this, a one dimensional array representing the Yoneda-cut is computed, by calling the function $make\_horizontal\_cut()$. The input is a 2d array, as well as the cut position in pixel ($cut\_position\_x$) and the linewidth, given the range of integration in q_z direction. Note that a linewidth of n

pixel result in a integration in q_z of n pixel above and n pixel below *cut_position_x*, resulting in an integration over $2n + 1$ pixel. This is done for both, experimental data and simulation.
For vectors describing the Yoneda and off-detector cut the Pearson coefficient

$$r_{xy} = \frac{\sum_i (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_i (x_i - \overline{x})^2}\sqrt{\sum_i (y_i - \overline{y})^2}} \tag{3}$$

is calculated. A Pearson coefficient of 1 yields a total positive linear correlation, 0 is no linear correlation, and 1 is total negative linear correlation. The total correlation of a simulation and the experimental data is the product of Pearson coefficient at the Yoneda cut and off detector cut.
Due to binning effects and the fact that the detector is built of smaller modules, the Yoneda cut data is further smoothed with a *SavitzkyGolay* filter with a default window length of 3 pixel.

3.5 Move in parameter space

Since the database contains $2768^2 = 7667361$ combination of simulated GISAXS pattern, it is necessary to probe the parameter space in such a way, that one finds the best correlated simulation with as less comparisons as possible. To archive this, a good initial guess is important. This is done by fitting the Yoneda linecut with an Lorentz's curve and determine the peak position $q_{y,max}$, known as the cluster correlation. The the cluster correlation peak [1] is related with the average center-center particle distance by

$$D \sim \frac{2\pi}{q_{y,max}}, \tag{4}$$

which is taken as initial guess for D_{init}. R_{init} is guessed to be $\frac{D}{2} - 0.5$. The assumption $R < D/2$ is full filled, when the deposited material does not form a layer, but forms particles [1]. Which does also depend on the sputter rate [2]. The effective thickness, that needs to be deposited to form a closed layer is known as percolation threshold. The percolation threshold for Au on Si is $\delta = (6.8 \pm 0.1)\,\mathrm{nm}$ [1] performing Volker Weber growth[2]. σ_{init}/R and ω_{init}/D are guessed to be 0.3, which turns out to in the range of reasonable values for sputter deposited Au on Si that is investigated later. To probe the parameter space, one of two major modes to move in the parameter space can be performed for an iteration. The order of which mode is taken for each iterations can be given by the user.

Snake search A random vector in the parameter space is created. The vector crosses the point (see fig 4b p0) in parameter space which is given by the initial guess $(R_i, D_i, \omega_i, \sigma_i)$. Every element described by the vector will be compared to the experimental data. The element (see fig 4 b) p1-p3) of the vector that results in the best correlation will also be an element of the vector created in the following iteration.
Probe nearest neighbors It is also useful to only probe the parameter space around one point in parameter space. For example when the best correlation did not change for

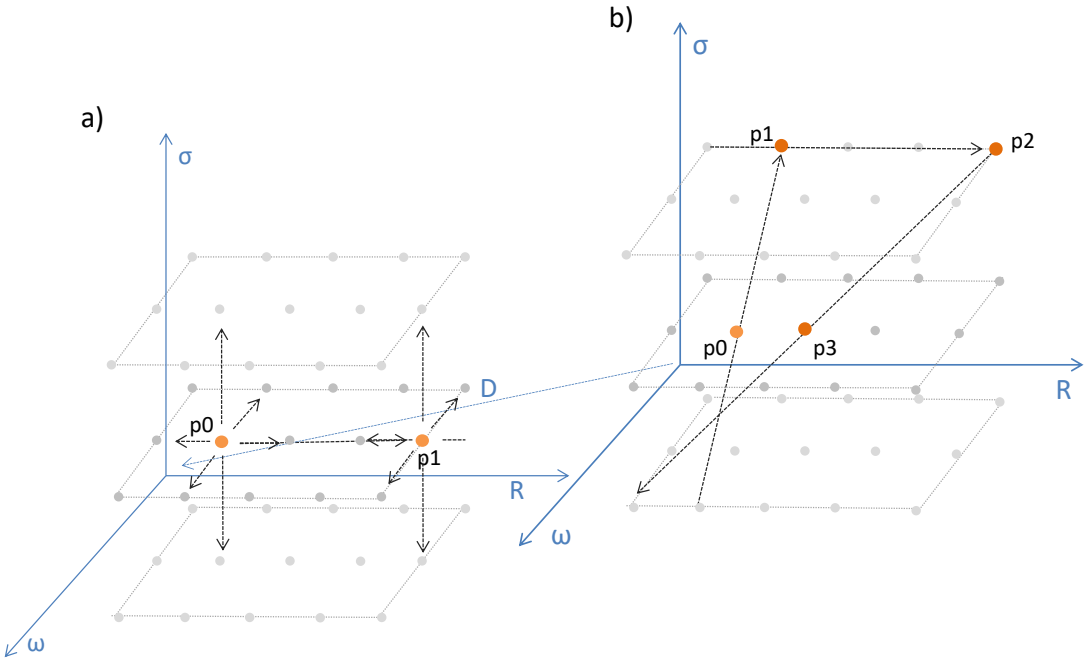

Figure 4: Illustration of "Probe nearest neighbors" (a) and "Snake search" (b) searching modes to move in parameter space .

a couple of iterations. Then the measurement will be compared to the nearest neighbor and the third nearest neighbor (see fig 4a p1). When a new best correlation is found this is done again. The the best correlating element of the vector, again, is also element of the vector created for the next iteration

A lot of GISAXS experiments are performed by taking images at a rate up to 67 Hz [1] to perform real time investigations during industrial-style sputter deposition. In this case the first image of the measurements is compared to the database with a high number of iterations (i.e $n = 30$), to obtain a higher probability to find the absolute maximum of correlation. When the best fitting simulation is found, the next image of the measurement series is compared with the database with a lot lower number of iterations. The maximum of correlation from the previous image is taken as initial guess. This can be done when the changes of parameter (R, D, ω, σ) from one measurement to the next, is only in the range of the step size of the given database.

4 Testing the algorithm

The algorithm is tested with data set of GISAXS measurements of gold deposition on a Si (100) substrate performed at the MiNaXS beamline P03 [3]. A sputter power of 3 W is applied, resulting in a deposition rate of 0.04 nm/s. GISAXS measurements where done with a rate of 20 Hz, using a PILATUS 1M detector with a pixel size of 172 μm. The DC sputter is closer described elsewhere [4]. The data set contains measurements of effective thicknesses ranging from 0 nm to 8 nm. The wavelength of the incoming beam is $\lambda = 0.0953$ nm and an angle of incidence of $\alpha_i = 0.41$ deg, which is slightly above the angle of incidence used for the simulation, with $\alpha_{i,sim} = 0.4$ deg.
The algorithm is fed with five GISAXS measurements done at layer thicknesses between 0.5 nm and 8 nm. As described in section 3.5, since the layer thickness difference from one measurement to the next is 0.5 nm to 4 nm, every measurement is compared independent from the previous one. Also at this stage the comparison was only done by comparing the Yoneda cut, whereas the off-detector cut is not included. The movement in parameter space is only performed in the "search snake" mode (3.5). Fig. 5 shows the

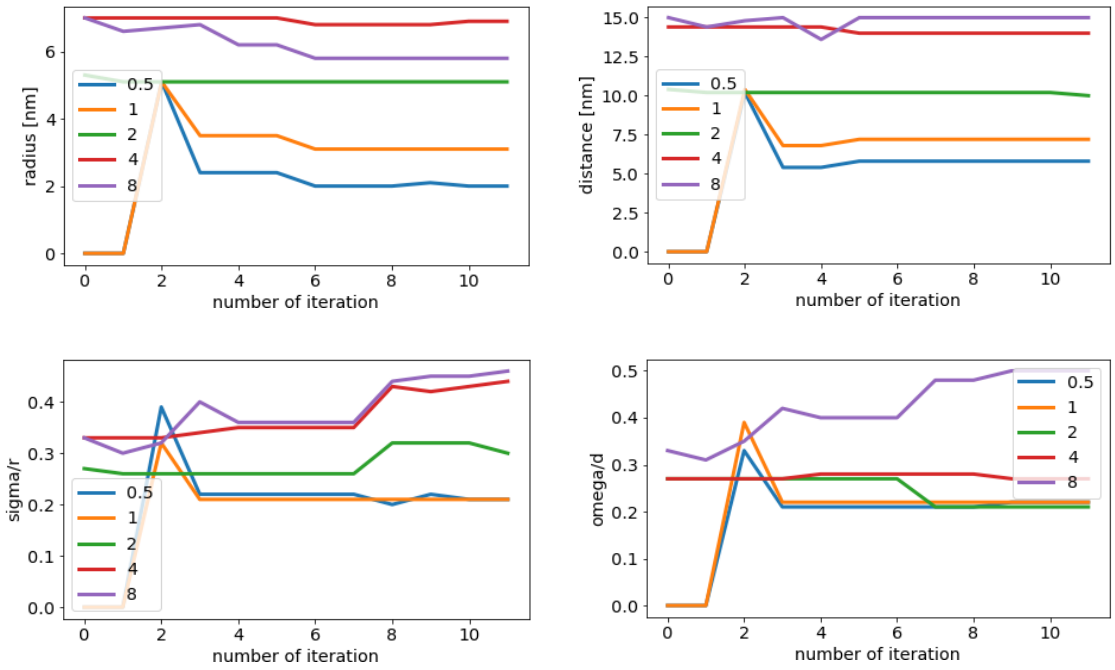

Figure 5: R, D, σ, ω development over iterations for different thicknesses.

evolution of parameter over the number of iteration. Note that one iteration means a comparison with all (up to 71) elements of the parameter vector. The colors of the curve represents a different layer thickness. For the comparison of thickness 0.5 nm and 1 nm, each parameter (R, D, ω, σ) changed from 0 to a finite value after three iterations, after the third iteration they stay more or less constant. For the thicknesses greater than 1 nm, the changes of parameter over iterations are significantly smaller. This impose that the initial guesses of R, D, ω and σ where better for thick layers. For all thicknesses R and D do not much vary for higher iterations, whereas ω and especially σ do change for higher iteration. This caused by the fact, that R and D are related to the peak position, which have a big influence on the Pearson coefficient, whereas σ and ω manifest themself more in the broadening and smearing of the peaks, which does influence the Pearson coefficient less. Note that after the 11th iteration the radius of the 8 nm layer is smaller then the of the 4 nm layer. This behaviour is contradicting, since the particle are suppose to grow with the amount of deposited gold. Fig. 6 shows the

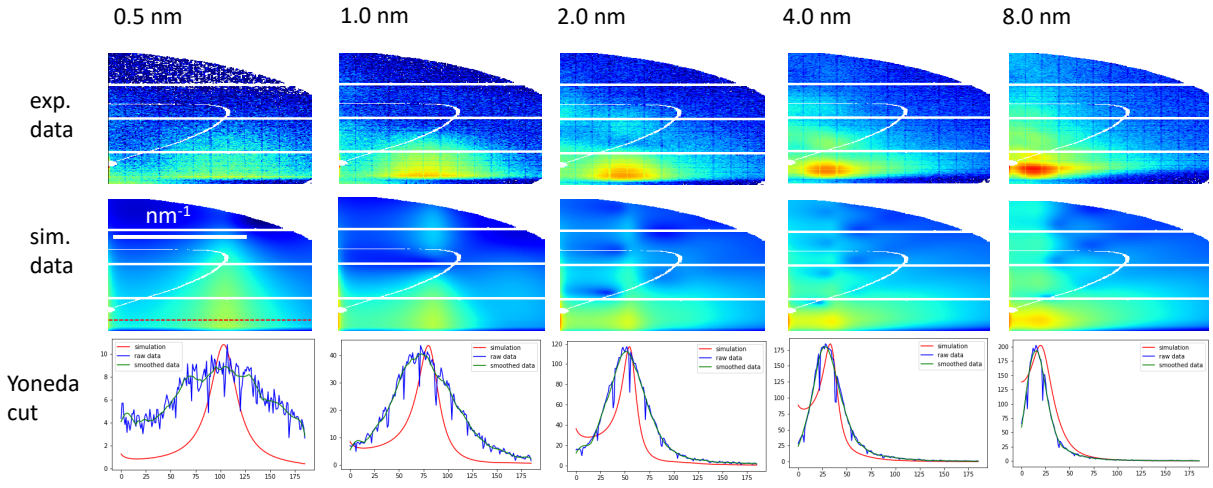

Figure 6: Measured (top) and simulated (middle) GISAXS pattern for 0.5 nm to 8 nm Au on Si. Also the corresponding Yoneda cuts are presented in the lower panel.

measured GISAXS pattern and the corresponding simulations and yoneda cuts which resulted in the best correlation, for different thicknesses. For each thickness one finds a good agreement of the measured and simulated Yoneda peak position. The broadening of the peak does not fit well, especially for low thicknesses. ω and σ seems to be underestimated for thickness 0.5 nm to 2.0 nm.

Tab. 4 contains the best parameter of the best fitting simulation for each measurement. One needs to consider that the data bank simulates just up to $D = 15$ nm. As seen in fig. 5, the distance does not change after five iteration and stick at $D = 15$ nm, which indicates that the real distance is greater than 15 nm. The quantity $2R/D$ in tab. 4. It is one at the percolation threshold, further more it is supposed to increase with time

	R	D	σ/R	ω/D	$2R/D$
0.5 nm	2.0 nm	5.8 nm	0.20	0.21	0.69
1.0 nm	3.1 nm	7.2 nm	0.21	0.21	0.86
2.0 nm	5.1 nm	10.0 nm	0.0.3	0.21	1.02
4.0 nm	6.9 nm	14.0 nm	0.44	0.27	0.99
8.0 nm	5.8 nm	15.0 nm	0.46	0.5	0.77

Table 3: Best matches found for the first test.

and don't further decreased once this threshold is reached. Looking at $2R/D$ in tab. 4 shows a strong increase and a crossing up to $\delta = 2\,\text{nm}$, then it decreases and crosses the percolation threshold again. This indicates that the obtained parameter in tab. 4 do not describe the system well.

For the second test the off detector cut is included and contribute to the correlation coefficient as described in 3.4. Also the "probe nearest neighbours" searching mode is included, and runs alongside the "search snake" mode, as described in 3.5. Here the comparison of the 8 nm thick layer is not performed, since the database does not include the expected parameter describing this system. Fig. 7 shows the best correlated simulation for each thickness (2nd row). The both lower rows show the comparison of Yoneda cut and off detector cut. comparing the Yoneda cut of the 0.5 nm thick layer, one finds a better correlation of simulation and experiment than seen in fig. 6. This test results in average higher values of σ and ω then the first test. Excluding the the 0.5 nm

	R	D	σ/R	ω/D	$2R/D$
0.5 nm	1.5 nm	4.8 nm	0.38	0.39	0.63
1.0 nm	3.1 nm	7.2 nm	0.22	0.22	0.86
2.0 nm	4.6 nm	10.2 nm	0.37	0.23	0.90
4.0 nm	6.9 nm	14.4 nm	0.49	0.27	0.96

Table 4: Best matches found for the second test.

thick layer in tab. 4 and tab. 4, D does not change from one test to the other. But the radius does tend to smaller values. The radius is also the height of the hemispherical particle (assumption in simulation), it manifests itself in the off detector cut at the cluster correlation peak. Since the off detector peak is taken into account calculating a correlation coefficient, the obtained values for R should be more reliable than from the first test. $2R/D$ in tab. 4 does increase continually and does not cross the percolation threshold for layer smaller than 4 nm. It shows that with including the "probe nearest neighbours" searching mode and the off detector cut the comparison could of simulation and experimental data could be improved.

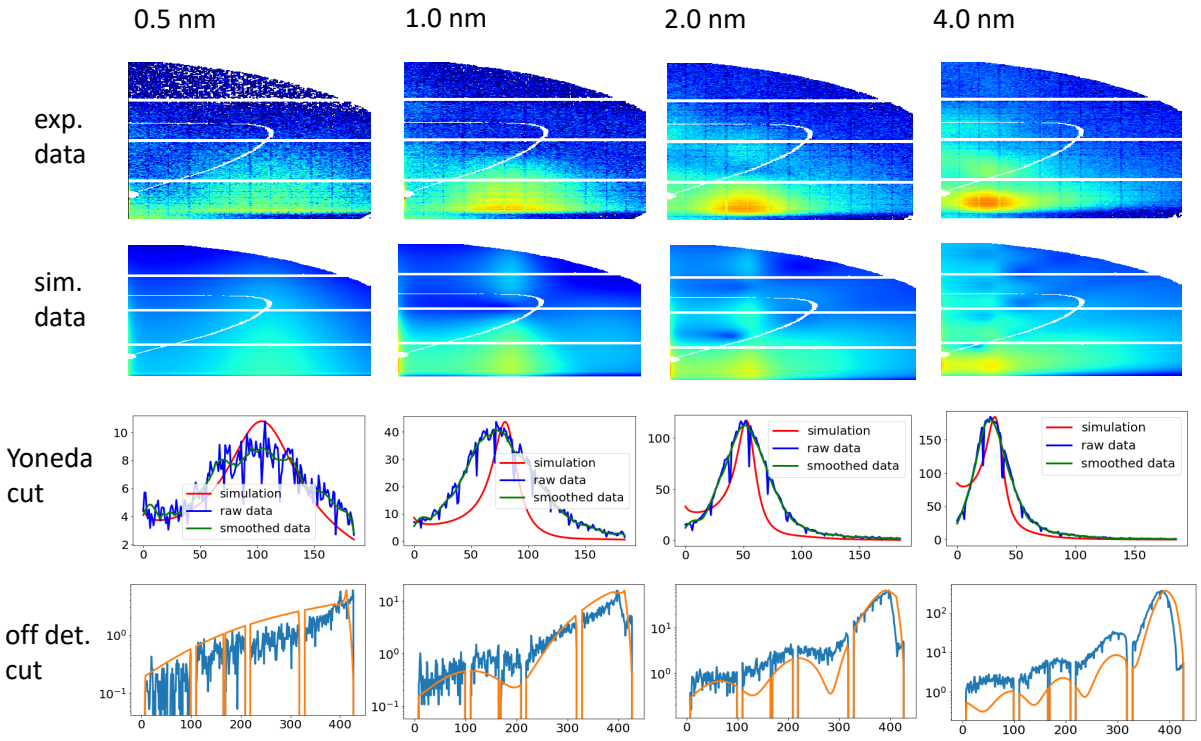

Figure 7: Measured (top) and simulated (middle) GISAXS pattern for 0.5 nm to 4 nm Au on Si. Also the corresponding Yoneda and off detector cuts are presented in the two lower panels.

Also measurements with thicknesses ranging from 2.0 nm to 4.0 nm, with a step size of 0.2 nm, were investigated. It turns out, that the parameter given for a thickness of 2.0 nm to 4.0 nm does match the corresponding values in tab. 4 pretty well. Also it shows that the movement through the parameter space is well reproducible, meaning that the maximum correlation is found again. By using the "probe nearest neighbours" searching mode after performing one comparison with much iterations for the first measurement, it is possible to find good correlations for the next comparison by only searching in the parameter neighbourhood.

5 Outlook and Conclusion

The work of this reports shows that a comparison of simulated and measured GISAXS data is suitable to obtain characteristic parameter describing the investigated system. By introducing a combination of "snake search" and "probe nearest neighbours" searching mode the number of iterations needed to bin the best match of simulation and measurement could be decreased. Calculating the Pearson coefficient as correlation quantity of the Yoneda cut and off detector cut, more reasonable values for R,D,σ and ω could be obtained.

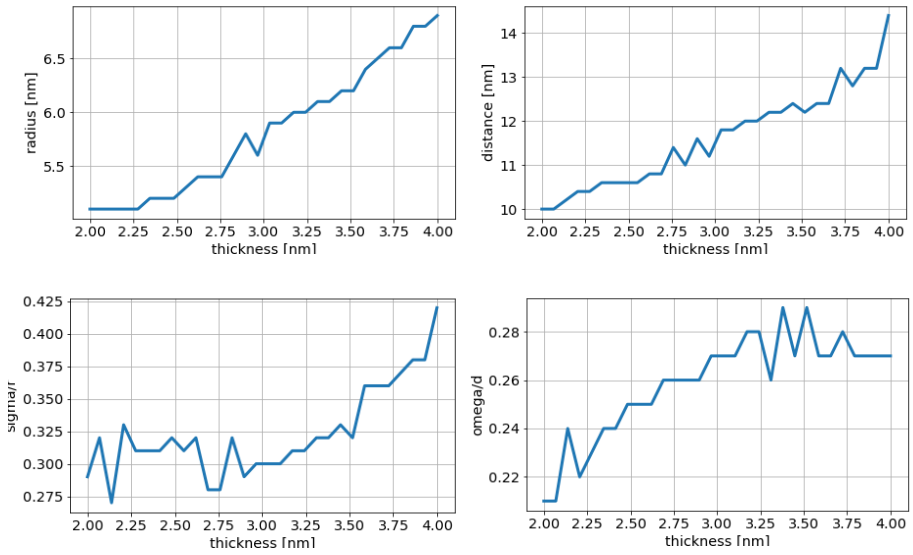

Figure 8: R, D, σ, ω for thicknesses ranging from 2 nm to 4 nm with a step size of 0.2 nm.

Never the less a more reliable concept should be developed to improve the comparison, since for now only a small amount of the present data is used to perform the comparison. Using deep learning methods could be a good approach to handle the complexity of GISAXS pattern. Also this algorithm should run on a super computer to test it with more data sets. Also the database should be extended for more particle layouts and a bigger range ob parameter.

6 Acknowledgments

I want to thank the whole staff of MiNaXs P03 group of PETRA III for a warm welcome and introducing me to the topic of GISAXS. Prof. Dr. Stephan V. Roth, Dr. Matthias Schwartzkopf, Dr. Pallavi Pandit, Dr Andrei Chumakov, Marc Gensch , Dipl. Ing. Jan Rubeck,Sven-Jannik Woehnert, Vivian Waclawek, Emanuel Schork, Calvin J. Brett .
My special acknowledgments go to my supervisor Dr. Matthias Schwartzkopf for helpful dicussion and supervison, as well as to Prof. Dr. Stephan V. Roth, Sven-Jannik Woehnert, Dr. Pallavi Pandit and Andre Rothkirch.

References

[1] Matthias Schwartzkopf, Adeline Buffet, Volker Korstgens, Ezzeldin Metwalli, Kai Schlage, Gunthard Benecke, Jan Perlich, Monika Rawolle, Andre Rothkirch, Berit Heidmann, Gerd Herzog, Peter Mueller-Buschbaum, Ralf Rohlsberger, Rainer Gehrke, Norbert Stribeck and Stephan V. Roth; From atoms to layers: in situ gold cluster growth kinetics during sputter deposition; Nanoscale, 2013, 5, 5053

[2] Matthias Schwartzkopf, Alexander Hinz, Oleksandr Polonskyi, Thomas Strunskus, Franziska C. Lohrer, Volker Koerstgens, Peter Mueller-Buschbaum, Franz Faupel and Stephan V. Roth; Role of Sputter Deposition Rate in Tailoring Nanogranular Gold Structures on Polymer Surfaces; ACS Appl. Mater. Interfaces 2017, 9, 56295637

[3] Adeline Buffet, Andre Rothkirch, Ralph Doehrmann, Volker Koerstgens, Mottakin M. Abul Kashem, Jan Perlich, Gerd Herzog, Matthias Schwartzkopf, Rainer Gehrke, Peter Mueller-Buschbaum and Stephan V. Roth; P03, the microfocus and nanofocus X-ray scattering (MiNaXS) beamline of the PETRA III storage ring: the microfocus endstation; J. Synchrotron Rad. (2012). 19, 647653

[4] Matthias Schwartzkopf, Alexander Hinz, Oleksandr Polonskyi, Thomas Strunskus,Franziska C. Loehrer, Volker Koerstgens, Peter Mueller-Buschbaum, Franz Faupel and Stephan V. Roth; Role of Sputter Deposition Rate in Tailoring Nanogranular Gold Structures on Polymer Surfaces; ACS Appl. Mater. Interfaces 2017, 9, 56295637

[5] Faupel, F.; Zaporojtchenko, V.; Strunskus, T.; Elbahri, M. Metal-Polymer Nanocomposites for Functional; Applications. Adv. Eng. Mater. 2010, 12, 11771190.

[6] Levine, J.R.; Cohen, J.B.; Chung, Y.W.; Georgopoulos, P. Grazing-incidence small-angle X-ray scattering: New tool for studying thin film growth. J. Appl. Crystallogr. 1989, 22, 528532.

[7] J. Burle, C. Durniak, J. M. Fisher, M. Ganeva, G. Pospelov, W. Van Herck, J. Wuttke, D. Yurov (2018), BornAgain - Software for simulating and fitting X-ray and neutron small-angle scattering at grazing incidence, http://www.bornagainproject.org

[8] R. Lazzari IsGISAXS: a program for grazing-incidence small-angle X-ray scattering analysis from supported islands J. Appl. Cryst. 2002, 35, 406-421. doi: 10.1107/S0021889802006088

[9] Thurn-Albrecht, T.; Schotter, J.; Kastle, G.A.; Emley, N.; Shibauchi, T.; Krusin-Elbaum, L.; Guarini, K.; Black, C.T.; Tuominen, M.T.; Russell, T.P. Ultrahigh-Density Nanowire Arrays Grown in Self-Assembled Diblock Copolymer Templates. Science 2000, 290, 21262129.