Correlating Sampling and Intensity Statistics in Nanoparticle Diffraction Experiments

Authors

Hande Öztürk[a], Hanfei Yan[b], John, P. Hill[bc] and I. Cevdet Noyan[d]*

[a]Applied Physics and Applied Mathematics, Columbia University, 116th st. and Broadway, New York, New York, 10027, USA

[b]National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, New York, 11973, USA

[c]CMPMSD, Brookhaven National Laboratory, Upton, New York, 11973, USA

[d]Applied Physics and Applied Mathematics, Columbia University, 116th st and Broadway, New York, New York, 10027, USA

Correspondence email: icn2@columbia.edu

Synopsis Individual crystallites contribute multiple beams to diffraction spectra from nanocrystalline powders. This complicates correlating the number of diffracting grains with diffracted intensities.

Abstract In a previous article (Ozturk *et al.*, 2014) we showed that the sampling statistics of diffracting particle populations within a polycrystalline ensemble depended on the size of the constituent crystallites: broad X-ray peak breadths enabled some nano-sized particles to contribute more than one diffraction spot to Debye-Scherrer rings. Here we show that the equations proposed by Alexander, Klug and Kummer (AKK) (Alexander *et al.*, 1948) to link diffracting particle and diffracted intensity statistics are not applicable if the constituent crystallites of the powder are below 10 nm. In this size range, (1) the one-to-one correspondence between diffracting particles and Laue spots assumed in the AKK analysis is not satisfied, and (2) the crystallographic correlation between Laue spots originating from the same grain invalidates the assumption that all diffracting plane normals are randomly oriented and uncorrelated. Such correlation produces unexpected results in the selection of diffracting grains: For example, three or more Laue spots from a given grain for a particular reflection can only be observed at certain wavelengths. In addition, correcting the diffracted intensity values by the traditional Lorenz term, $\frac{1}{cos\theta}$, to compensate for the variation of particles sampled within a reflection band does not maintain fidelity to the number of Laue spots contributing to the diffracted signal. A new term, $\frac{cos\theta_B}{cos\theta}$, corrects this problem .

Keywords: Diffraction, nanoparticle, Lorenz factor, X-rays, statistics.

Nomenclature

$\vec{k}_i$	Incoming X-ray wavevector
$\vec{k}_d$	Diffracted X-ray wavevector
λ	X-ray wavelength
$\vec{q}$	Momentum transfer vector
$\rho(\vec{r})$ $\rho_\infty(\vec{r})$	Electron density functions (Local density above and 3D periodic density below)
$y(\vec{r})$	Shape function of a diffracting particle
F_{hkl}	Structure factor of the unit cell
H_{hkl}	Pole enhancement factor
D	Particle diameter
N_c	Number of unit cells in a particle in the direction of the momentum transfer vector
β	Full width at half maximum of an intensity peak
C	Scherrer shape factor
Γ	The complementary angle to the angle between the pole vector with the incoming wavevector
s	Range multiplier (integer)
m_{hkl}	Multiplicity of the {hkl} planes
P^*_{hkl}	Activated hkl pole
α_{hkl}	Reflection band-width
N_G	Number of irradiated grains in the powder
$N_{P,hkl}$	Number of poles belonging to the {hkl} planes
$N_{G^*,hkl}$	Number of grains that are diffracting at the hkl reflection
$p_{P^*,hkl}$	Probability of having an activated pole within the hkl reflection band
$N_{P^*,hkl}$	Number of activated poles within the hkl reflection band
P^*_1	An activated singlet pole
P^*_n	An activated n-let pole
$\vec{C}_i$	The vector pointing from the origin of the reference sphere to the i'th pole on the surface of the sphere
$\vec{C}_A$	Grain orientation vector
P^*_{C3}	An activated pole corresponding to the $\mathbf{C_3}$ pole vector

p_{P*2}	Probability of having two poles activated simultaneously
$p_{P*2,1}$	Probability of the first member of the doublet poles to be activated
γ_{ij}	Interplanar angle between 'i' and 'j' planes
d_{hkl}	Interplanar distance between {hkl} planes
λ_{ES}	Enhanced selection wavelength (special wavelength where highly correlated poles appear)
N_{G*T}	The total number of all diffracting grains within a powder
N_{P*T}	The total number of all activated poles from a powder sample
$\bar{I}$	Integrated intensity
CT_1, CT_2	Correction factors for the activated pole distributions
$\tilde{u}_X$	Relative uncertainty of quantity 'x' based on a given set of measurements

1. Introduction

Structural parameters obtained from analysis of powder diffraction data, such as crystallite size, atomic site occupation parameters, defect concentration, texture, strain and phase fractions are complicated average values of the actual distributions existing within the sampling volume (Fewster, 2014). There are two independent factors which determine the uncertainty associated with these structural parameters. First, the total volume fraction of crystallites contributing diffracted intensity determines how representative refined structural parameters might be of the entire sample. The second factor is the quality of the measured intensity profiles: low peak intensity values, coupled with high background and overlapping diffraction peaks can lead to erroneous structural characterization of the sample.

Combining the uncertainties from both factors to develop quality metrics for diffraction profiles is a complex problem. For a non-textured powder sample estimating the number of grains which will contribute finite intensity to a particular reflection is traditionally carried out using the Lorenz formulation (Cullity, 1978). This formulation assumes that each crystallite contributes a single ray to the diffraction pattern (*single-pole per particle (SPP) assumption*). Consequently, any statistical variation in the number of grains[1] oriented favourably for diffraction will be directly proportional to the variation in the measured diffracted intensity. In other words, with the exception of the uncertainties in photon counting, the statistical variation of diffracted intensity can be computed directly from grain sampling statistics. The Alexander, Klug and Kummer analysis (henceforth referred to as the AKK analysis) utilizes this idea to construct equations which predict the expected uncertainty in diffracted intensities as functions of grain size, Bragg angle, X-ray beam divergence, etc. These equations were originally tested with experimental intensity data from quartz powder samples with particle sizes between 5 to 50 micrometers. For this size range, the measured and computed uncertainties showed reasonable agreement.

Neither the AKK formulation, nor the Lorenz analysis has any provisions for large scale failure of the single-pole assumption; namely, the presence of a large fraction of grains with multiple poles, satisfying the diffraction condition and resulting in many diffraction spots per crystallite. Such failure is possible: in a recent study (Ozturk *et al.*, 2014) it was shown through modelling that, for nano-sized particles with dimensions smaller than 10 nm, individual crystallites contribute multiple spots to the Debye-Scherrer ring of a given reflection, and the one-to-one correspondence between the number of diffracting particles

[1] In this article, the terms "(powder) particle", "crystallite" and "grain" are used interchangeably to describe a single, perfectly crystalline volume, similar to standard usage in diffraction literature, e.g. Smith, D. K. (2001). *Powder Diffr* **16**, 186-191, ibid.

and number of diffraction spots in the Debye-Scherrer ring did not hold. In that reference, the ratio of the diffraction spots to the number of diffracting grains was parameterized as the "pole enhancement factor" (H_{hkl}), which increased as the particle size decreased and multiplicity of a reflection increased; the H_{hkl} value computed for the 311 Debye-Scherrer ring of 1.6 nm gold particles illuminated by a planar monochromatic X-ray beam of 12 keV energy was ~1.25, indicating that there were 25% more spots than grains.

The H_{hkl} parameter described by Öztürk, et.al. does not provide the link from particle sampling statistics caused by variations in the orientation distribution of sample crystallites to the diffracted intensity statistics. In addition, H_{hkl} does not differentiate between grain subsets which contribute two, three or more spots to the ring. Consequently it cannot be used to investigate the correlation between the statistical variations in diffracted intensity with the statistical variations in the number of diffracting grains. Experimental determination of such correlation through experiments is also a nontrivial undertaking. In the current manuscript these issues are addressed through modelling and geometric analysis.

2. Theory

2.1. Basic Formalism

In this section a simple formalism for modelling the powder diffraction process from a random polycrystalline sample is presented. In contrast to the venerable Debye equation (Hall, 2000), this formalism permits the sampling statistics to be unambiguously linked to diffracted intensity statistics: each diffraction peak is modelled by identifying all crystallites which satisfy the diffraction condition for the particular reflection, and then summing up their (angle-resolved) intensity contributions. To facilitate interpretation of the results by removing all uncertainty sources unrelated to sampling, the following ideal conditions will be assumed:

2.1.1. Basic Assumptions

1. All crystallites in the powder sample are exactly spherical gold particles and have the same diameter, D. Their orientations in the laboratory space are completely uncorrelated, resulting in a perfectly "random" powder specimen.
2. All atoms of a crystallite belong to complete unit cells of a single space group, described by a the single structure factor, F_{hkl}, corresponding to the particular hkl reflection of gold.

3. All atoms are static, with no thermal vibrations. This corresponds to a temperature of 0 K in classical mechanics.
4. The incident beam is a perfect, monochromatic plane wave of exact wavelength, λ. There is no angular divergence in either incident or diffracted beams.
5. There is no photoelectric absorption, fluorescence or inelastic scattering.
6. The diffraction process operates at the kinematical limit.
7. An ideal two-dimensional detector with infinite resolution and dynamical range, infinitesimal pixel size, zero dark current and no background noise is used.
8. There are no alignment errors of any kind.

Assumptions 1 through 3 define an "ideal powder sample"; assumptions 4 through 8 define an ideal experimental set-up. Combined, these assumptions define an ideal experiment which should yield the lowest possible uncertainty values associated with (modelled) sampling and intensity parameters. In other words, one cannot achieve higher accuracy in their diffraction data than the limits investigated in this article.

2.2. Computation of diffracted intensity from a single crystallite

Assumptions 1 to 8 permit the use of the Patterson formalism to exactly determine the rocking curve of a spherical crystallite (Patterson, 1939a). Accordingly, the angle dependent intensity, I(Δ2θ), can be written as

$$I(\Delta 2\theta) = \left|A(\Delta 2\theta)\right|^2 = N_C^2 \left|F_{hkl}\right|^2 [Y(\vec{q})]^2 \qquad \text{(1-a)}$$

where $A(\Delta 2\theta)$ is the complex amplitude, N_C is the number unit cells in the particle, F_{hkl} is the structure factor and $Y(\vec{q})$ is the Fourier transform of the shape function associated with the momentum transfer vector $\vec{q}$ given by

$$Y(\vec{q}) = \left[3\frac{-\varsigma\cos\varsigma + \sin\varsigma}{\varsigma^3}\right] \qquad \text{(1-b)}$$

Here $\varsigma = \frac{\pi}{\lambda} D \cos\theta_B \Delta 2\theta$; where D is the diameter of the perfectly spherical crystalline particle[2], θ_B is the Bragg angle of the hkl reflection being measured and $\Delta 2\theta$ is the deviation of the Laue spot from twice the exact Bragg angle, $2\theta_B$, of this reflection. Numerical analysis shows that, for finite D, $I(\Delta 2\theta)$ exhibits a strong maximum at the exact Bragg condition, $\Delta 2\theta = 0$, with subsidiary maxima or "thickness fringes" bracketing this primary peak; the thickness fringe intensities tend to zero at $\Delta 2\theta = \pm\infty$. As D increases, $[Y(\vec{q})]^2$ decays quickly and the peak function becomes narrower, converging to a Dirac delta in the limit of infinite sphere diameter. Since $[Y(\vec{q})]^2$ tends to unity at $\Delta 2\theta = 0$ for all D, the maximum intensity of the primary peak of the rocking curve for any sphere with "N_c" unit cells is equal to $N_C^2 |F_{hkl}|^2$.

In Figure 1 normalized diffracted intensity plots, $I(\Delta 2\theta) / I(\Delta 2\theta = 0)$, calculated by using Eq. (1-a) for the 200 reflection of spherical gold particles, with diameters of 5, 10 and 50 nm respectively, are shown. For these computations, a planar X-ray beam of 12.4 keV energy (λ=1 Å) was used, resulting in 200 diffraction peaks from the gold lattice at the Bragg angle $2\theta_B = 28.36^o$. Within the angular range covered in this figure, the peak function of the largest (50 nm) particle decays five orders of magnitude before the peak function of the smallest (5 nm) particle reaches to 50% of its peak value.

The full-width-at half-maximum (FWHM), β , of the primary peak of the rocking curve or the period of the minima (zeroes), τ , of the thickness fringes predicted by Eq.(1-a) can be obtained from:

$$\beta = \frac{1.17\lambda}{D\cos\theta_B} \qquad \text{(2-a)}$$

$$\tau = \frac{\lambda}{D\cos\theta_B} \qquad \text{(2-b)}$$

Eq. (2-a) is a particular form of the traditional Scherrer equation[3] (Patterson, 1939b):

[2]Due to the requirements of the crystalline lattice, not all atoms on the surface of the sphere can be exactly at the geometric sphere boundary. This approximation becomes weaker as the size of the sphere decreases and should fail when the crystallite size is comparable to the unit cell size.(Palosz, B., Grzanka, E., Gierlotka, S. & Stelmakh, S. (2010). *Z Kristallogr* **225**, 588-598.)

[3] This equation is obtained from Eq. (1-a) by setting the left-hand side to half of the maximum intensity, $I(\Delta 2\theta = 0)/2$, and numerically solving for the corresponding $\Delta 2\theta$ values. Similarly, Eq. (2-b) is obtained by determining the position of all minima in the thickness fringes, where $I(\Delta 2\theta) = 0$. (Yan, H. (2006). PhD thesis, Columbia University, Ying, A. (2010). PhD thesis, Columbia University, Ying, A. J., Murray, C. E. & Noyan, I. C. (2009). *J Appl Crystallogr* **42**, 401-410.The thesis works can be downloaded from the following links, respectively:

$$\beta = \frac{C\lambda}{D\cos\theta_B} \qquad \text{(2-c)}$$

where "*C*" is the Scherrer constant. For spherical crystallites, C is equal to 1.17.

2.2.1. Basic diffraction geometry

Figure 2 shows a schematic of the diffraction geometry when a single particle is irradiated with a parallel plane X-ray beam of wave-vector $\overrightarrow{\boldsymbol{k_i}}$ and wavelength λ. The wave-vector of the diffracted beam is $\overrightarrow{\boldsymbol{k_d}}$, [hkl] is the normal vector to the set of diffracting planes, (hkl), and $\boldsymbol{\theta_{B,hkl}}$ is the (exact) Bragg angle for the hkl reflection. The [hkl] vector forms the polar angle $\boldsymbol{\Gamma_B} = \frac{\boldsymbol{\pi - 2\theta_{B,hkl}}}{\mathbf{2}}$ with $\overrightarrow{\boldsymbol{k_i}}$. The intersection of the [hkl] vector with the surface of the unit sphere is termed an activated hkl pole, $\boldsymbol{P^*_{hkl}}$. This pole corresponds directly to a particular diffracted spot on the reference sphere surface, and to its projection on a 2-D detector. This (projected) intensity spot on the detector is termed a "Laue spot". In this construct, (activated) poles are elements belonging to the (powder) sample space whereas the diffracted spots or Laue spots are elements of the detector space.

If an ideal powder specimen is placed at the unit sphere origin in Figure 2, the locus of activated poles from the (randomly oriented) grains diffracting at the exact Bragg angle, $\boldsymbol{\theta_{B,hkl}}$, will be the hkl reflection circle (the dashed line in Fig. 2). Crystallites whose orientations deviate from the Bragg condition while still remaining within the rocking curve will broaden the reflection circle into a reflection band of angular width, $\boldsymbol{\alpha_{hkl}}$. The corresponding loci of all diffracted beams on the reference sphere will be the Debye-Scherrer halo. The projection of this halo on the 2-D detector will yield the Debye-Scherrer ring which contains all Laue spots of finite intensity. The "range factor", *s*, defines the angular range, in terms of a multiple of the FWHM, $\boldsymbol{\beta}$, within which the activated poles contribute finite intensity to the hkl reflection. This term depends on the angular coverage desired in the experiment. For a given range factor, s, the relationship between the angular width, $\boldsymbol{s\beta_{hkl}}$, of the Debye-Scherrer ring and the breadth of the reflection band, $\boldsymbol{\alpha_{hkl}}$, which contains the corresponding (activated) poles is $\boldsymbol{\alpha_{hkl} = s\beta_{hkl}/2}$ (Ozturk *et al.*, 2014).

https://www.researchgate.net/publication/33743142_Modeling_of_strain_fields_in_semiconductor_single-crystals_using_dynamical_diffraction_theory_
http://gradworks.umi.com/34/42/3442837.html)

2.2.2. Computation of the diffracted intensity from the ideal powder sample

The intensity diffracted by the ideal powder sample at a particular Δ2θ angle around the Bragg peak will be due to the contributions from all crystallites oriented for diffraction at this particular angle. Since nanosized crystallites can contribute more than one Laue spot to the Debye Scherrer ring, the contributions from all Laue spots of finite intensity, rather than crystallites, must be summed together. In that case the intensity expression becomes:

$$I_{tot}\left(\Delta 2\theta_j\right) = \sum_{i=1}^{N_{S_j}} I_{j,i} = N_c^2 \left|F_{hkl}\left(\Delta 2\theta_j\right)\right|^2 \sum_{i=1}^{N_{S_j}} \left[3\frac{-\varsigma_{j,i}\cos\varsigma_{j,i} + \sin\varsigma_{j,i}}{\varsigma_{j,i}^{3}}\right]^2 \qquad (3)$$

Here $I_{j,i}$ is the intensity contribution due to spot number 'i' at the "jth" diffraction angle; $2\theta_j = 2\theta_B \pm \Delta 2\theta_j$, and N_{S_j} is the number of spots properly oriented to scatter at this angle. Equation (3) shows that, for the ideal powder sample under assumptions 1-8, the diffracted intensity at a given angle has two components: **1)** A deterministic value, given by Eq. (3), which depends only on the symmetry, size and composition of the crystallite and **2)** the number of spots oriented to scatter at this angle, N_{S_j}, which is a stochastic quantity. This term will have different values for every member of a set of identical ideal powder samples, each with the same number of grains. The variation among N_{S_j} is due to the random distribution of crystallite orientations. Thus, neglecting counting statistics, the statistical variation in the measured intensities will be caused by the statistical variations in N_{S_j} for all angles, $\Delta 2\theta_j$.

Determination of N_{Sj} is not straightforward since the relationship between this term and the number of diffracting grains within the powder sample depends on the particle size (Ozturk *et al.*, 2014). For a set of identical powders, the actual number of grains oriented at a particular solid angle should be distributed normally around a mean value. If each grain had only one pole associated with it (SPP assumption), the distribution of poles on the surface of the unit sphere would also be random, and the mean and standard deviation of the poles would be equal to the corresponding grain values. When multiple poles are associated with a single grain, the assumption that all poles are randomly distributed on the unit sphere surface becomes weaker since the orientations of diffracting poles from a single grain are linked through crystal symmetry. To understand how such linkage determines the functional dependency between grain and pole populations, we used a simple geometric analysis, which is an extension of the classical Lorenz formulation. This is described in the following.

2.3. Diffracting Grain vs. Activated Pole Fractions in an Ideal Powder Sample

2.3.1. Fraction of Total Activated Poles in a Given Reflection Band

If the single crystallite in Figure 2 is replaced by an ideal powder sample which contains N_G grains, the total number of poles, $N_{P,hkl}$, on the surface of the sphere corresponding to an hkl reflection of multiplicity, m_{hkl}, will be $N_{P,hkl} = m_{hkl} N_G$. Among these poles, only those which fall within the reflection band will become activated and have corresponding Laue spots in the Debye-Scherrer ring. If all poles were uncorrelated, the probability, $p_{P^*,hkl}$, of any one hkl pole falling within the reflection band would be equal to the fractional area of this band with respect to the surface area of the reference sphere:

$$p_{P^*,hkl} = \frac{A_{band}}{A_{sph}} \cong \frac{\int_{\Gamma=\frac{\pi}{2}-\theta_B-\frac{s\beta}{4}}^{\Gamma=\frac{\pi}{2}-\theta_B+\frac{s\beta}{4}} \int_{\phi=0}^{2\pi} r d\Gamma \, r \sin\Gamma \, d\phi}{4\pi r^2} = \cos\theta_B \sin\left(\frac{s\beta}{4}\right) \qquad \text{(4-a)}$$

Substituting for β from Eq. (2-c), and assuming small $\left(\frac{s\beta}{4}\right)$ values, one obtains the modified Lorenz equation (Ozturk *et al.*, 2014):

$$p_{P^*,hkl} \approx s\frac{C\lambda}{4D} \qquad \text{(4-b)}$$

The number of activated poles expected in the reflection band is, then, given by:

$$N_{P^*,hkl} \approx N_G \; m_{hkl} \; p_{P^*,hkl} = N_G \; m_{hkl} \; s\frac{C\lambda}{4D} \qquad \text{(4-c)}$$

Eq. (4-b) predicts that, for typically used X-ray wavelengths and large crystallite sizes, D>500 nm, the probability of any one pole falling within the reflection band, and thus contributing finite intensity to the rocking curve is quite small (on the order of 10^{-4}). Consequently, each diffracting crystallite can be assumed to possess only one, single, activated pole. For crystallite diameters around 10 nm, $p_{P^*,hkl}$ values are on the order of 10^{-2}. Below this size the angular breadth of the reflection band is of the order of several degrees and, thus, the probability of multiple poles from a single crystallite falling within this band is finite. (Ozturk *et al.*, 2014). To investigate the differences between distributions of activated poles and diffracting grains in detail, two types of activated poles are defined:

1) **Uncorrelated (Independent) Poles**: These "singlet" poles, P_1^*, have no other members of the <*hkl*> family belonging to the same grain within the reflection band of width $s\beta_{hkl}$.

2) **n-fold Correlated Poles:** These poles belong to "n-let" groups, such as doublets, triplets, etc. with each group having 'n' members $(n \geq 2)$ of the hkl poles belonging to one particular grain. The i-th activated pole belonging to a particular n-let is denoted by $P_{n,i}^*$. The orientations of $P_{n,i}^*$ with respect to each other are fixed by the crystal symmetry.

The total number of (activated) poles within the reflection band is the sum of poles for all groups:

$$N_{P^*,hkl} = N_{P^*,1} + \sum_{n=2}^{r} nN_{P^*,n} \qquad \text{(5-a)}$$

Here $N_{P*,n}$ is the population of n-fold correlated poles, and "*r*" is the largest n-let dimension that can be observed for the particular crystal in the particular diffraction geometry.

From equations (5-a) and (4-b) we obtain:

$$p_{P^*,hkl} = \frac{N_{P^*,hkl}}{N_G \cdot m_{hkl}} = \sum_{n=1}^{r} p_{P^*n} \approx s\frac{C\lambda}{4D} \qquad \text{(5-b)}$$

Eq. (5-b) indicates that the probability of finding an activated pole belonging to a given <hkl> family should be approximately equal to the probability predicted by the modified Lorenz equation. Any deviations will be due to the correlation among n-lets.

Equation (5-b) cannot be factored to yield the probabilities of the activated pole subgroups (singlets, doublets, triplets, etc.) existing in the reflection band. We now investigate if these subgroup probabilities can be obtained from geometric considerations.

2.3.2. Expected Fraction of Doublets in the Reflection Band

Consider Figure 3, where the intersection of the reference sphere surface with vector $\vec{C}_3$, belonging to the <h00> family[4], forms an activated pole, $P_{C_3}^*$, on the reflection circle. While keeping the angular position of $P_{C_3}^*$ invariant, the pole of the $\vec{C}_2$ vector, P_{C_2}, can be brought within the reflection band through a

[4] In this formulation the <h00> family is used for convenience; this family has the lowest multiplicity and all family members are orthogonal.

rotation ξ around $\vec{C}_3$. The locus of angular positions within the reflection band where P_{C_2} will be activated simultaneously with $P^*_{C_3}$ is the general quadrilateral within the reflection band (shaded area in Figure 3) and its complement on the other side of the half-sphere. These quadrilaterals are formed by the intersection of the two great circles which form the loci of the h00 poles when $\vec{C}_2$ is rotated around $\vec{C}_3$ at the two limiting polar angles, $\Gamma \pm s\beta / 4$, of the reflection band. Thus, to a first approximation the probability of having two simultaneously activated correlated poles, p_{P*2}, will be given by the product of: **1)** the probability, $p_{P*2,1}$, of the first pole, P_{C_3}, being within the reflection band and, **2)** the probability, $p_{P*2,2}$, of the second pole, P_{C_2}, falling within either of the general quadrilaterals on the reflection band. The first probability term, $p_{P*2,1}$, is given by Eq. (5-b). The activation probability of the second pole, $p_{P*2,2}$, is proportional to the ratio of the quadrilateral area to the area of the spherical lune:

$$p_{P*2,2} \propto \frac{A_{quad}}{A_{lune}} = \frac{2 \times \int_{\Gamma=\frac{\pi}{2}-\theta_B-\frac{s\beta}{4}}^{\Gamma=\frac{\pi}{2}-\theta_B+\frac{s\beta}{4}} \int_{\phi=-\frac{s\beta}{4}}^{\phi=+\frac{s\beta}{4}} r d\Gamma \, r \sin\Gamma \, d\phi}{2 \times \frac{s\beta/2}{2\pi} 4\pi r^2} = \frac{r^2 s\beta \cos\theta_B \sin\left(\frac{s\beta}{4}\right)}{2r^2(s\beta)} = \frac{1}{2}\cos\theta_B \sin\left(\frac{s\beta}{4}\right) \qquad (6)$$

The probability of simultaneously activating two poles belonging to the <h00> family will be:

$$p_{P^*2} \propto p_{P*2,1} . p_{P*2,2} = \left[\cos\theta_B \sin\left(s\beta/4\right)\right]^2 \qquad (7\text{-a})$$

Substituting for β_{hkl} from Eq. (4-c), and using the small angle approximation we obtain:

$$p_{P^*2} \propto \left[\frac{sC\lambda}{4D}\right]^2 \qquad (7\text{-b})$$

For a fixed crystallite shape and constant wavelength, the probability of activating two h00 poles from a given crystallite is predicted to be proportional to $(s/D)^2$.

For the <h00> family of directions, Equation (7-b) describes the probability of a given pair of h00 poles being activated together; it does not preclude the probability of a third pole (if exists) also being simultaneously activated. Since the position of this third pole is dictated by crystal symmetry, this third

probability cannot be computed from simple area fractions. For general hkl type reflections, where the plane normal vectors are not mutually orthogonal, the integrals in Eq. (6) must be modified.

2.3.3. Expected Fraction of Triplets in the Reflection Band

Consider the case where three poles, $P^*{}_{C_1}, P^*{}_{C_2}, P^*{}_{C_3}$ belonging to three members of the general <hkl> family, $\vec{C}_i = [h_i, k_i, l_i], i = 1,3$, are exactly on the reflection circle (Fig.4-a). In this case a crystal orientation vector, $\vec{C}_A = [m, n, p]$, which is anti-parallel with the incident wave vector, $\vec{k}_i$, must simultaneously satisfy the angular relationship:

$$\vec{C}_A \cdot \vec{C}_i = |C_A||C_i| \cos\Gamma \quad \Rightarrow \quad \Gamma = \cos^{-1}\left[\frac{mh_i + nk_i + pl_i}{\sqrt{m^2 + n^2 + p^2}\sqrt{h_i^2 + k_i^2 + l_i^2}}\right] \qquad \text{(8-a)}$$

for all $\vec{C}_i$. Here h_i, k_i, l_i are the indices of the particular vector, $\vec{C}_i$, of the <hkl> family.

The crystallography of the unit cell imposes three additional constraints on the angular relationships between the vectors $\vec{C}_i$:

$$\vec{C}_i \cdot \vec{C}_j = |\vec{C}_i||\vec{C}_j| \cos\gamma_{ij} \quad \Rightarrow \quad \gamma_{ij} = \cos^{-1}\left(\frac{h_i h_j + k_i k_j + l_i l_j}{h^2 + k^2 + l^2}\right) \qquad \text{(8-b)}$$

Here $\gamma_{ij,}$ are the interplanar angles between the planes normal to $\vec{C}_i, \vec{C}_j$; these angles can be calculated, or obtained from published tables (Kelly & Knowles, 2012; Won-Sik Han, 2007). The angle, γ_{ij} , between vectors, $\vec{C}_i, \vec{C}_j$, defines the angular separation of the corresponding activated poles in polar coordinates.

Eq. (8-a) defines the diffraction condition, where a particular orientation vector $\vec{C}_A$ makes the complementary angle $\Gamma = \frac{\pi}{2} - \theta_B$ with all members of a select triplet of the poles belonging to the <hkl> family. This defines the condition where the planes normal to the three crystal vectors, $\vec{C}_i$, exactly satisfy the Bragg condition ($\Delta 2\theta = 0$) and, thus, diffract waves at maximum intensity (Eq. (1-a)). In this case, the triplet of the $\vec{C}_i$ vectors has rotational symmetry around the incident beam vector; rotating the crystal

around this axis will not change the angular positions of the poles, $P^*{}_{C_i}$, with respect to $\vec{k}_i$, or the diffracted intensities of the corresponding Laue spots.

Deviation of $\vec{C}_A$ and $\vec{k}_i$ vectors from the exact antiparallel condition will move the $\vec{C}_i$ vectors away from the Bragg condition, decreasing the intensity diffracted by the corresponding {hkl} planes. As long as the misorientation between $\vec{k}_i$ and $\vec{C}_A$ is small enough to retain the poles of the $\vec{C}_i$ vectors within the reflection band area, all three poles will remain activated; the intensities of the corresponding Laue spots will be proportional to the relative elevation angle that each pole is displaced from the reflection circle. Then, as long as the pole P_{C_A} belonging to the orientation vector $\vec{C}_A$ is within a small area on the sphere surface centred around $\vec{k}_i$, then the triplet of poles, $P^*{}_{C_1}, P^*{}_{C_2}, P^*{}_{C_3}$, will remain activated. Thus, the probability of activating a triplet of poles will be proportional to the ratio of this area to that of the sphere surface.

In Figure (4-b), the areas swept by the pole, P_{C_A}, while keeping the normal vectors, $\vec{C}_2, \vec{C}_3$, within reflection bands of breadths, $\alpha_{h00} \approx \beta_{h00}$ and $\alpha_{h00} \approx 2\beta_{h00}$ are shown. Using numerical analysis the area ratios of these bands to the sphere surface were found to be approximately proportional to $(\cos\theta_B)^2 (s\beta/4)^2 = (sC\lambda/4D)^2$.

2.3.4. Dependency of Triplet Activation on the Illumination Wavelength

For low multiplicity reflections, the set of equations given in Eq. (8-a) uniquely defines the diffraction condition for the activation of a triplet of correlated poles. In the limiting case, for h00 reflections from cubic crystals, where $m_{hkl} = 6$, only a unique crystal vector $\vec{C}_A$, anti-parallel with the incident wave vector, $\vec{k}_i$ at one single wavelength, λ, can satisfy the diffraction condition for triplets. This condition can be observed by combining Eq. (8-a) with Bragg's law, $\lambda = 2d_{hkl}\sin\theta_B$, and substituting $\Gamma = \frac{\pi}{2} - \theta_B$, to yield:

$$\vec{C}_A \cdot \vec{C}_i = |C_A||C_i| \sin\left[\arcsin\left(\frac{\lambda}{2d_{hkl}}\right)\right] \qquad \text{(9-a)}$$

which, upon re-arranging, becomes:

$$\sin\left[\arcsin\left(\frac{\lambda}{2d_{hkl}}\right)\right] = \frac{\vec{C}_A \cdot \vec{C}_i}{|C_A||C_i|} \qquad \text{(9-b)}$$

For an h00 reflection from a cubic crystallite, the $\vec{C}_A$ axis must have three-fold symmetry; this condition is satisfied only by one of the cube diagonals, <111>. Consequently, Eq. (9-b) indicates that triplets of activated h00 poles at the exact Bragg condition can be observed only for a particular enhanced-selection wavelength, λ_{ES}:

$$\lambda_{ES} = \frac{2d_{h00}}{\sqrt{3}} \qquad \text{(9-c)}$$

In the case of a gold crystallite with lattice parameter 4.08Å, 200 type triplets with all poles at the exact Bragg condition would be observed for λ_{ES}=2.36 Å. The finite breadth of the reflection band, $s\beta_{hkl}/2$, within which pole activation is possible permits the observation of these triplets over a range $\lambda_{ES} + \Delta\lambda$. This range can be determined by substituting $\theta = \frac{\pi}{2} - \Gamma \mp \frac{s\beta}{4}$ into Bragg's law.

2.3.5. Expected Fraction of Singlets in the Reflection Band

For a single crystallite at the origin of the reference sphere, the number of poles in the half-sphere containing the reflection band is half of the multiplicity of the hkl reflection, $m_{hkl}/2$. To treat the case where only one of these correlated poles can be within the reflection band, the probability formalism must be modified to include a term describing the conjunctive probability of having none of the remaining $(\frac{m_{hkl}}{2} - 1)$ poles to be within the reflection band. Since the poles that are associated with a single diffracting grain are perfectly correlated through crystal symmetry, the probability p_{P*1} cannot be evaluated from $\prod_{j=2}^{g}(1 - p_{P*j})$ where $g = \frac{m_{hkl}}{2}$ and $p_{P*,hkl}$ is given by Eq. (5-b). However, using Eq. (5-b), it is possible to write:

$$p_{P*1} \cong p_{P*,hkl} - \sum_{n=2}^{r} p_{P*n} \qquad \text{(10-a)}$$

Here the summation is over all possible "n-lets" (n≤g) which can occur in the reflection band. For the case of the h00 reflection Eq. (10-a) becomes:

$$p_{P*1} \cong p_{P*,hkl} - (p_{P*2} + p_{P*3}) \quad \text{(10-b)}$$

$$p_{P*1} \cong \frac{sC\lambda}{4D} - (p_{P*2} + p_{P*3}) \quad \text{(10-c)}$$

This equation predicts the following:

1) Since both p_{P*2} and p_{P*3} are proportional to $(s/D)^2$, p_{P*1} values should vary nonlinearly with s/D.
2) For the hypothetical case where the diameter "D" is small enough to broaden the reflection band over the entire half-sphere, the probability of having one and only one independent pole per grain to be in the diffraction condition will be identically equal to zero since all poles will have correlated siblings.
3) For large *D*, Eq. (10-c) tends to the classical probability (of any active pole) given by Eq. (4-b) since as both p_{P*2} and p_{P*3} rapidly approach zero.
4) For intermediate *D* values the probability, p_{P*1}, will be smaller than the probability values computed using the modified Lorenz approximation, i.e., Eq.(4-b).

2.3.6. Fractions of higher multiplicity reflections

Equations (8)-(10) can be extended to reflections of higher multiplicity through geometric analysis. In the general case, many activated triplets within the general <hkl> family, corresponding to arrays of orientation vectors $\vec{C}_A$ and enhanced selection wavelengths, λ_{ES}, will exist. Similar considerations apply to having simultaneously activated quadruplets, quintuplets or higher order n-lets of poles for higher multiplicity reflections. In such cases, the governing equations are also quite unwieldy; therefore, they will be investigated through modelling.

In conclusion, for those cases where the conditions for multiple pole activation are satisfied, the diffracting grain population can no longer be considered equivalent to the diffracted Laue spot population. Moreover, although the orientation distribution of the crystallites making up the ideal powder sample must be random by definition, the distribution of the diffracted spots within a given Debye-Scherrer ring cannot be completely random: the multiple spots that are diffracted from one crystallite will have definite orientation relationships imposed by the crystallographic structure of the crystallite. To understand how

these issues influence the diffracted intensity, the true intensity profiles from ideal powder samples must be computed using the angular distribution of all activated poles. This task is not analytically tractable. We used rigorous numerical modelling for this purpose.

3. Modelling Procedure and Results

In this section, the process of modelling the diffraction signal expected from an ideal, untextured, powder ensemble will be described in detail and the results obtained from these computations will be compared to the predictions from the equations derived in Section 2 where appropriate. Our modelling algorithm consists of the following steps:

1-A perfectly spherical crystalline gold particle of a certain diameter is defined.
2-We use a true randomizing algorithm (Kuffner, 2004) to assign a random orientation to this crystallite in the laboratory coordinate system in which diffraction will be modelled.
3-We place this particle at the centre of a hypothetical diffractometer equipped with a 2-D detector, and a monochromatic plane wave incident beam.
4-For a given family of planes, {hkl}, we check the orientation of each set of planes (hkl) to see if the normal to this set of planes, the vector [hkl], falls within the corresponding reflection band.
5-In the laboratory coordinate system, the angular coordinates of those plane normal, [hkl], which intersect the reflection band, and thus form activated poles, are noted. A typical distribution of activated poles within the reflection band area is depicted in Figure 5-a.

6-Each coordinate set is used to determine the position of the Laue spot within the Debye-Scherrer ring corresponding to a particular set of planes (hkl), and the Patterson function for the particular shape is used to compute the diffracted intensity for each Laue spot belonging to the particular crystallite.
7-This procedure is repeated for all crystallites in the powder ensemble. This yields a “spotty” Debye-Scherrer ring (Figure 5-b) for the hkl reflection in which each particular Laue spot can be referred to the crystallite from which it was diffracted from. In addition, the “siblings” of any Laue spot (those originating from the same grain) can be identified. This information is tabulated.
8-The Debye-Scherrer rings are azimuthally compressed (caked) (Sulyanov *et al.*, 1994) to yield the rocking curves or the radial ($2\theta / \theta$) scans (Figure 5-c) of the corresponding reflections. These plots are analysed to obtain parameters such as the integrated and maximum peak intensities.

9- Peak intensity values are correlated with the number of diffracting grains and activated poles, and used to test the applicability of the previously published statistical equations.

Simulations were carried over a range of wavelengths, λ, and particle sizes, *D*. Several reflections, corresponding to a broad range of multiplicities were modelled. The simulation parameters are summarized in Table 1.

Particle Shape	Spherical; *C*=1.17
Particle Diameter (nm)	2.4, 2.86, 4.0, 5.0, 8.0, 10.0, 25.0, 30.0, 80.0, 200, 250, 300, 700
Wavelengths, λ, (Å)	0.7, 1.0, 1.5, 2.0, 2.29, 3.0
Reflections used (multiplicities)	111 (8), 200 (6), 220(12), 311(24)
range multiplier (*s*)	2, 3, 4
# of irradiated grains, N_G	10^4, $5x10^4$,10^5

Table 1 Simulation parameters used in the modelling procedure.

3.1. Grain and Pole Sampling Statistics

In Figure 6 the total number of diffracting grains, N_{G^*T} the total number of activated poles, N_{P^*T} , and the populations of all correlated n-lets, N_{P^*n} for the 200 reflection from 2.86 nm diameter spherical gold nanoparticle ensembles are shown as a function of wavelength for sampling interval, s, equal to 4. In this range, N_{P^*1} and N_{P^*2} obtained from the simulation program are finite, continuous, non-monotonic and nonlinear over the tested wavelength range. The number of activated poles belonging to triplets, on the other hand, changes discontinuously: N_{P^*3} is finite only at λ=2.29Å, which is within the selection range defined by Eq. (9-c); N_{P^*3} is zero for all other wavelengths used in the simulation. The number of activated poles belonging to doublets, N_{P^*2}, drop steeply when triplets are activated and shoot up when triplets are supressed again, surpassing the number of singlets, N_{P^*1} , for λ=3Å. At this wavelength, the reflection band is quite wide and doubly correlated poles are more probable, as discussed previously.

In contrast to the behaviour observed for N_{P^*1}, N_{P^*2} and N_{P^*3}, the variation of the total number of activated poles, N_{P^*T} , with respect to λ is continuous and linear. In addition, N_{P^*T} obtained from the

simulation program is equal, within statistical error, to the value computed from Eq. (4-c), indicating that the area fraction of the reflection band is an acceptable approximation for the total pole probability, p_{P*}. Similar behaviour was observed for all modelled reflections: for a fixed selection interval, s, the variation of the total number of activated poles, N_{P^*T}, was continuous and linear over the tested wavelength range, and agreed within statistical error with the N_{P^*T} values computed from Eq. (4-c). On the other hand, the variation of the number of diffracting grains, N_{G^*T} with λ was nonlinear for all cases.

Eq. (4-c) also intimates that the variation of N_{P^*T} should be linear with the parameter, s/D, for fixed wavelength, λ, for all reflections. Our simulation results for λ=2.29Å showed that this prediction was largely obeyed (see Figures 7-a and 7-c for the 200 and 311 reflections of the spherical gold nanoparticle ensembles). For the 200 reflection, the variations of N_{P^*1}, N_{P^*2} and N_{P^*3} show a quasi-parabolic dependence on s/D, as anticipated through Equations (7-b) to (10-c). For the 311 reflection, all $N_{P*,311-n}$, with the exception of $N_{P*T,311}$, exhibit non-monotonic and nonlinear behaviour over the s/D range used in the simulations. The 111 and 220 reflections showed similar behaviour.

The variation of the total number of diffracting grains, $N_{G^*T,hkl}$, and the populations belonging to their respective subsets with n-let activated poles are depicted in Figures 7-b and 7-d for these two reflections. All of these populations exhibit nonlinear behaviour over the s/D range used in the simulations.

Figures 7-a and 7-c show that the variations of $N_{P*T,hkl}$ are linear, and approximate the predictions of Eq. (4-c), over the s/D range used in our simulations. In contrast, the total number of diffracting grains (Figures 7-b and 7-d) approach linear behaviour only for $\frac{s}{D} < 0.02$. For s=4, this corresponds to a particle diameter of 200 nm. Below this size $N_{G*T,hkl}$ cannot be obtained from the Lorenz formulation.

A closer examination of Figure 7-d reveals that $N_{P*T,311}$ obtained from the simulation diverge slightly from the predictions of Eq. (4-c), at the larger s/D values. This is shown more clearly in Figure 8-a, where $\Delta N_{P^*T,hkl} = (N_{P^*T,hkl})_{sim} - (N_{P^*T,hkl})_{Eq.(4-c)}$ is plotted as a function of s/D.

Figure 8-a shows that $\Delta N_{P^*T,hkl}$ systematically increases with the multiplicity of the reflection and also with increasing s/D; the dashed line in Figure 8-a depicts a 3rd order polynomial fitted to the $\Delta N_{P^*T,311}$ values. This deviation is due to the assumption, implicit in the derivation of Eq. (4-c), that the poles

existing within the reflection band are completely uncorrelated. As expected, this assumption becomes weaker with increasing multiplicity and s/D, as shown in Figures 7-a and 7-c.

Figure 8-b shows $\Delta N_{P^*T,hkl}$ as % fraction of the total number of activated poles. It is seen that, on average a 1% systematic difference is expected for the higher multiplicity reflections at large s/D values. At the very low values, the large error bars are due to the large propagated statistical error: the number of grains selected for diffraction is much smaller at these large sizes for all reflections.

The close correspondence of the total number of poles obtained from simulation, $\left(N_{P^*T,hkl}\right)_{sim}$, with the values computed from Eq. (4-c) also intimates that the ratios of $\left(N_{P^*T,hkl}\right)_{sim}$ from various reflections, $R_{P*} = N_{P*T,hkl} / N_{P*T,h'k'l'}$, should yield values approximately equal to the ratios of the multiplicities from these reflections. These ratios are plotted in Figure 9-a for various reflection pairs; the corresponding multiplicity ratios are depicted by dashed lines. The results are quite close. In contrast, the ratios of the total number of grains from various reflections (Fig. 9-b) approach the multiplicity ratios only for small s/D, where the probability of diffracting multiple Laue spots per grain is much lower.

3.1.1. Influence of the Range Multiplier "*s*" on Sampling Statistics

The probabilities of activated poles of all flavours, singlets through n-lets, as well as the total activated pole probability for any reflection, computed from simple area ratios (Eq. (4), Eq. (7)), depend, either linearly or through a power law, on the normalized angular width, $s\beta/2$, which is expressed in radians. After substituting for β from the Scherrer Equation (Eq. (4-c)), this width becomes the dimensionless quantity $sC\lambda/4D$. This expression implies that all of these probabilities should increase continuously as the parameters s, λ and $1/D$ increase. In the case of the wavelength, λ, this intimation is wrong: as discussed in section 2.3.4 and shown in Figure 6, simple area ratios do not capture the constraints imposed on diffracting pole selection by crystal symmetry for n-lets with dimensions equal or higher than three; such n-lets are only activated at special wavelengths, λ_{ES}, given by Eq.(9-c). Consequently, the probabilities (and populations) of all flavours of activated poles should change continuously with the size-normalized selection interval, s/D only for a fixed wavelength. This prediction was verified by our simulation results[5]. We note that, while the inverse sample diameter, $1/D$, can exist over a (mostly)

[5] If Figures 7-a and 7-c are replotted with simulation results from multiple wavelengths, with the abscissa based on parameter $s\lambda/D$, some of the traces are no longer smooth. These plots are not included for brevity.

continuous range, the range multiplier term "s" is meaningful only for those values of $s\beta$ within which Laue spots have finite intensities. It is possible to determine this range independently from particle diameter, D, by expressing the angular variable of the abscissa of Figure 1 in integral multiples of β. The results are shown for three particle sizes in Figure 10-a. Almost perfect agreement is observed between the respective normalized intensities. Figure 10-a also indicates that most of the integrated peak intensity, $\bar{I} = \int_{-\infty}^{\infty} I(\Delta 2\theta) d\Delta 2\theta$, is captured for $|\Delta 2\theta| \leq 4\beta$ or $s = 4$. In Figure 10-b, the fractional integrated intensities, $\%\bar{I}_f = \int_{-s\beta}^{s\beta} I(\Delta 2\theta) d\Delta 2\theta / \int_{-\infty}^{\infty} I(\Delta 2\theta) d\Delta 2\theta$, are plotted over the range of $0 < s \leq 8$. The fractional (percent) intensity, $\%\bar{I}_f$, tends asymptotically to 100% following the empirical equation: $\%\bar{I}_f = 100\,(0.8) - 174(6) * [0.105(0.01)]^s$. Here the values in the parentheses are the errors associated with the fitted values. It can be seen from this equation and from Figure 10-b that regions with s greater than 4 contribute negligibly to the integrated peak intensity. Thus, reflection bands of width 2β are sufficiently wide to capture all activated poles for any wavelength and crystallite diameter. Consequently, setting s equal to 4 and varying only particle size, D, should be sufficient to compare sampling and intensity statistics for each wavelength. This approach will be utilized in the following sections where the intensity profiles corresponding to the activated pole distributions will be presented.

3.2. Distribution of the Activated Pole Populations and the Resulting Intensity Profiles in the Vicinity of the Bragg Angle

Figure 11-a shows the activated pole distributions contributing to the 200 reflection of an ideal gold nanoparticle ensemble irradiated with 1, 1.54 and 2.29 Å wavelength X-rays, respectively. In these computations 10^4 randomly oriented, spherical particles, 28.56 Å in diameter, with lattice parameter, $a_0 = 4.08$Å, were modelled. Per Section 3.1.1, identification of activated poles were limited to an angular band of width $4\beta_{200}$, centred at the Bragg angle, $2\theta_B$, of the 200 reflection. To generate statistically representative data, 10 independent runs were executed. Due to the random orientations of the constituent crystallites, statistically equal numbers of activated poles at each angular step, Δ2θ, were expected. However, this was not the case: for all three wavelengths $N_{P*T,200}$ changed linearly with Δ2θ, with fewer poles being activated on the high angle side of the Bragg peak, indicating the presence of a (pole) sampling bias. The least-squares lines fitted to the $N_{P*T,200}$ vs. Δ2θ data were steeper with increasing

wavelength; the slopes of the corresponding regression lines were -0.96 (±0.29), -1.36 (±0.36) and -5.47 (±0.50).

Figure 11-b shows the polycrystalline diffraction profiles obtained from the product of these activated pole distributions with the rocking curve scans of properly oriented single crystallites (Eq. (3)). Due to the negative slopes of the activated pole profiles, apparent peak shifts to lower angles are expected. This is shown in Table 2 where the Bragg angle, $2\theta_B$, FWHM (β) and the peak shift due to the sampling bias (in micro strain), corresponding to the profiles in Figure 11-b are tabulated.

Wavelength (Å)	θ_B (expected, deg)	β (expected deg-2θ,)	θ_B (fit, deg)	β_{fit} (fit, deg-2θ)	Peak shift $\frac{a_{sim}-a_0}{a_0} * 10^6$ (με)
1	14.188	2.42	14.19 (0.01)	2.50 (0.06)	207
1.54	22.176	3.90	22.17 (0.01)	3.90 (0.06)	214
2.29	34.144	6.50	34.10 (0.02)	6.48 (0.09)	1082

Table 2 Expected and fitted peak parameters and the peak shift error (in micro strain) for the profiles shown in Figure 11-b. The first two columns were computed from Bragg's law and Scherrer equation, respectively, for the listed wavelengths. The corresponding simulation values were obtained from fitting the intensity profiles using Gaussian functions. Fit errors are shown in parentheses.

This sampling bias was first predicted by Lorenz (This was mentioned in an article by P. Debye and P. Scherrer, based on a simple geometric analysis. See page 197 of the compilation of early papers on X-ray diffraction analysis (Bijvoet *et al.*, 1969).); it is caused by the variation of the $\cos\theta$ term (Eq. (4-a)) over the width of the reflection band. Thus, for a given reflection, hkl, the sampling bias would be more

pronounced at larger wavelengths where the Bragg angle of the reflection shifts to higher values and the value of β increases. In a similar fashion, for a given wavelength, λ, the Lorenz sampling bias is expected to increase with decreasing interplanar spacing. This is shown in Figure 12, where the variation of activated poles for the 111, 200, 220 and 311 reflections with Δθ are plotted for the powder sample discussed in Figure 11. The slopes of least-squares lines fitted to these data are -5.08 (±0.46), -5.47 (±0.50), -23.21 (±0.55) and -149.11 (±0.88), respectively. The highest peak shift will, thus, occur for the 311 reflection.

To avoid the spurious shifts in Bragg peak positions due to the systematic variation of $N_{P^*T,hkl}$ within the reflection band, a geometry-based correction must be used (Reynolds & Reynolds, 1986; Yinghua, 1987; Fitzsimmons *et al.*, 1991; Bijvoet *et al.*, 1969; Buerger, 1940). An examination of Eq. (4-a) shows that the dependency of $N_{P^*T,hkl}$ on Δθ can be remedied by multiplying $N_{P^*T,hkl}(\Delta\theta)$ either by $\boldsymbol{CT_1} = 1/\cos\theta$, or with $\boldsymbol{CT_2} = \cos\theta_B / \cos\theta$; this second correction term is obtained by removing the (implicit) assumption used in the derivation of Eq. (4-c): that $\cos\theta$ is approximately equal to $\cos\theta_B$ within the reflection band. In Figure 13 the application of these corrections to $N_{P^*T,311}$ vs. Δθ data are shown. Both corrections significantly reduced the sampling bias: the slopes of regression lines fitted to the "corrected" pole distributions are -4.43 (±2.54) or -1.62 (±0.93) for corrected data with $\boldsymbol{CT_1}$ or $\boldsymbol{CT_2}$, respectively; these values correspond, approximately, to 4% and 1% of the original slope. Diffraction peak profiles computed for the corrected pole distributions yielded peak shifts corresponding to an apparent strain of approximately 80 με, which is an acceptable uncertainty.

Figure 13 also shows that, while the profile corrected by the $\boldsymbol{CT_2}$ term maintains fidelity with the total number of activated poles, $N_{P^*T,311}$, within the reflection band (22251±149 for the corrected profile vs. 22308±162 for the uncorrected one), the $\boldsymbol{CT_1}$ correction increases this parameter by approximately three times (61002±444). Thus correction by the $\boldsymbol{CT_2}$ term appears to be preferable for single peak analysis. For full-profile fitting (such as Rietveld analysis) this term cannot be used directly: multiple peaks with their individual pole populations must be considered simultaneously and separate correction term are needed for each reflection. We are using further modelling to investigate how fidelity to the diffracting pole populations can be maintained for this case. These results will be reported in a later article.

We note that, in the literature, the correction for sampling bias is usually combined with other diffraction-geometry dependent terms to derive specific intensity correction formulations. These are termed, in aggregate, as the Lorentz factor, L(θ). For example, for synchrotron-radiation diffraction analysis of

randomly oriented nanometer-sized crystalline Pd, the Lorentz factor was given as $L(\theta) = \frac{1}{\sin\theta \sin 2\theta}$ (Fitzsimmons *et al.*, 1991) . Such intensity corrections cannot be used in the sampling domain and are not directly applicable to pole distributions, such as Figures (11-a) or (12).

3.3. Comparison of Sampling vs. Intensity Statistics

At this point the simulation process can be used to test the statistical analysis, published by Alexander, Klug and Kummer in 1948 (Alexander *et al.*, 1948), where measured intensity values were related to the number of diffracting crystallites through Laplace's generalized probability formula. This analysis is based on the *SPP* assumption, and postulates that the intensity of the diffracted ray will be directly proportional to the number of diffracting particles. This postulate can be tested by comparing the diffracting grain populations obtained from the simulations to the (computed) diffracted intensities. Table 3 summarizes the results of such a test where intensity and sampling data for the 200 and 311 reflections were computed for 10^4 , $5x10^4$ and 10^5 particle polycrystalline gold samples. For each reflection, the 10x increase in the total irradiated particle population results in an equal increase in the peak and integrated intensities, as well as the total number of activated poles and diffracting particles. The ratios of the maximum and integrated intensities I_{311}/I_{200}, and the total number of activated poles, $\frac{N_{P^*T,311}}{N_{P^*T,200}}$ scale, within statistical error, with the ratio of the corresponding multiplicities, $(m_{311}/m_{200}) = 4$. On the other hand, the ratio of the number of diffracting particles, $\frac{N_{G^*T,311}}{N_{G^*T,200}}$, is approximately 2.1. This ratio correlates neither with the increase in the irradiated grains, nor with the multiplicity ratio of the corresponding reflections and indicates the breakdown of the *SPP* assumption in the intensity domain. (These results complement those presented in Figures 9-a and 9-b which show that the ratios of the activated pole populations are equal to the ratios of the multiplicities of any two reflections over the entire normalized selection range, s/D. The diffracting particle populations are equal to this ratio only when the SPP assumption is valid.)

N_G	I_{max}	I_{int}	N_{G*T}	N_{P*T}	I_{max}	I_{int}	N_{G*T}	N_{P*T}

(λ=2.29A)	(200)	(200)	(200)	(200)	(311)	(311)	(311)	(311)
10^4	57.93 (±1.2)	58.27 (±0.8)	4793 (±60)	5589 (±65)	234 (±2)	231 (±2.6)	9864 (±7)	22273 (±113)
$5x10^4$	287.22 (±2.9)	289.53 (±2.4)	23761 (±120)	28038 (±129)	1158 (±6)	1155 (±5)	49331 (±25)	111370 (±170)
10^5	577 (±7)	579 (±5)	47595 (±113)	56175 (±222)	2325 (±11)	2316 (±6)	98677 (±25)	222770 (±297)

Table 3 Sampling and intensity parameters of 200 and 311 reflections from 2.86 nm diameter gold nanoparticle ensembles with 10^4, $5x10^4$ and 10^5 particles, respectively, illuminated with 2.29 Å X-rays. The errors in parentheses show the standard deviation from the mean of 10 independent simulation runs.

Consequently, for any *s/D* value where the SPP assumption is unjustified, changes in the diffracted intensities caused by changes in the (angular) width of the reflection band cannot be directly linked to the changes in the number of diffracting particles. An example is shown in Table 4, where the maximum and integrated peak intensities of the 200 intensity profiles corresponding to the 10^4 particle ensemble are tabulated for illumination with 1, 1.54 and 2.29 Å wavelengths, respectively, along with the numbers of the corresponding diffracting particles and activated poles[8]. The relative increases in the intensity parameters with increasing wavelength track the corresponding changes in the activated pole counts, while the relative increases in the numbers of diffracting particle counts are lagging.

λ (Angstrom)	**Intensity Max**	**Integrated Intensity**	N_{G*T}	N_{P*T}
1	25.21 (±1.56)	25.87 (±0.75)	2321 (±41)	2457 (±42)
1.54	38.60 (±1.48)	39.50 (±1.01)	3425 (±34)	3787 (± 38)
2.29	57.93 (±1.2)	58.27 (±0.82)	4739 (±60)	5589 (±65)

[8] These data correspond to the profiles shown in Figures 11-a and 11-b.

Table 4 Peak and integrated intensities of the 200 reflections shown in Figure 11-b, and the grain and activated pole populations contributing to these reflections. The errors in parentheses are based on the standard deviation from the mean values for 10 independent simulation runs each with 10^4 particles. The individual particle diameter is 2.86 nm.

In conclusion, our analysis shows that when the *SPP* assumption fails, the variation of the diffracting grain populations do not maintain fidelity with the variation of the intensity parameters. Thus, diffracted intensity ratios cannot be used to compare the corresponding diffraction volumes even when all other parameters, such as the structure factor, absorption, polarization, etc., are taken into account. As a corollary, in such cases the uncertainty due to statistical sampling of diffracting particles from an ideal powder sample might not be linked directly to the uncertainty in the resulting intensity data or vice-versa. This problem is examined next.

3.4. Comparison of Statistical Sampling Uncertainty and Intensity Uncertainty

In the AKK formulation, the relative uncertainty, $\tilde{u}_{G^*,hkl}$, in the number of diffracting grains, N_{G^*}, within a powder sample with N_G illuminated grains is assumed to be equal to the uncertainty in intensity data, $\tilde{u}_{I,hkl}$, given by the Laplacian probability equation:

$$\tilde{u}_{I,hkl} = \tilde{u}_{N_{G^*,hkl}} = \sqrt{\frac{1-p_{G^*,hkl}}{N_G p_{G^*,hkl}}} \qquad (11)$$

The relative uncertainty, $\tilde{u}_x$, is defined as ratio of the standard deviation of the relevant quantity 'x' to its mean value, $\bar{x}$, and $p_{G^*,hkl}$ is the probability that a given grain is in the diffraction condition for the hkl reflection. In terms of the FWHM through the Scherrer equation, this term becomes:

$$p_{G^*,hkl} = \frac{m_{hkl} s C \lambda}{4D} \qquad (12)$$

Relative uncertainties for the diffracting grain and activated pole populations, $\tilde{u}_{G^*,hkl}$, $\tilde{u}_{P^*,hkl}$, as well as for the maximum and integrated peak intensities, $\tilde{u}_{I-max,hkl}$, $\tilde{u}_{I-int,hkl}$ were computed from the data sets shown in Table 3; these mean values and standard deviations were obtained from powder samples with 10^4, $5\text{x}10^4$ and 10^5 particles (N_G), respectively. For each ensemble size, 10 independent simulations were

performed. The results are presented in Table 5, where the predictions of the AKK formulation (Eq. (11)) are also tabulated.

N_{G*}	$\tilde{u}_{I-\max,200}$	$\tilde{u}_{I-\mathrm{int},200}$	$\tilde{u}_{N_{G*,200}}$	$\tilde{u}_{N_{P*,200}}$	$\tilde{u}_{I,hkl} = \tilde{u}_{N_{G*}}$ Eq.(11)
10^4	0.0207	0.0137	0.0125	0.0116	0.0088
$5x10^4$	0.0102	0.0083	0.0051	0.0046	0.0039
10^5	0.0121	0.0085	0.0024	0.0040	0.0028

Table 5 Fractional relative uncertainty values for the intensity parameters, diffracting particle and activated pole populations computed from the dataset shown in Table 3 for the 200 reflection. The simulations used a particle size of 2.86 nm, wavelength of 2.29 Å and s=4. 10 independent runs were used to obtain the mean and standard deviation of the relevant quantities for each. The intensities were computed after pole distributions were corrected by the CT_2 term. The last column lists the uncertainties predicted by Eq. (11).

Table 5 shows that, as the total irradiated particle population increases, the uncertainties in both intensity and sampling parameters decrease in a manner consistent with the Laplacian probability theory utilized the AKK formulation, which predicts the decay rates for the uncertainties to be proportional to the square root of the relative increase in the sample size. However, in contrast to the AKK formulation, the uncertainty values associated with the intensity and sampling parameters are not equal. The relative uncertainty values in the peak and integrated intensity parameters are much higher than the uncertainty values associated with diffracting grains or activated poles. Consequently, for those particle sizes where the *SPP* assumption fails, the uncertainties in the diffracting particle populations cannot be used to compute the expected relative uncertainties associated with the intensity parameters.

4. Summary and Conclusions

In this article, a diffraction experiment on an ideal nanocrystalline powder sample was reconstructed using rigorous modelling. The simulation started by defining an aggregate of nanocrystallites where the orientation of each was defined in the diffraction system coordinates. Based on this information, those

grains within the powder aggregate which satisfied the Bragg condition for one or more activated poles (those poles which fall within the reflection circle and correspond to Laue spots on the Debye-Scherrer ring) were identified. The (kinematic) diffracted intensity for each grain was computed by summing the diffracted intensity corresponding to each of its activated poles. The diffracted intensity from the aggregate for a given reflection was computed by summing the intensities from all particular diffracting grains. This procedure was repeated for several reflections spanning a range of multiplicities. In contrast to an actual diffraction experiment, this procedure provided full sampling information: the number and exact orientation of all diffracting grains and their corresponding activated poles were available in addition to the diffracted intensities. Thus, the sampling parameters contributing to the "diffraction inverse problem" were accessible.

Our analysis showed that the classical Lorenz formulation for computing the expected fraction of grains in the Bragg condition in a randomly oriented aggregate of crystallites was valid only for the large particle regime where *SPP* assumption holds. The failure of this assumption also invalidated the intensity statistics developed by Alexander, Klug and Kummer, which linked the uncertainty in integrated and maximum intensity values to the statistical variations in the number of properly oriented particles. The relative standard deviations in intensity values obtained from the simulation were much higher (~4x) than the relative standard deviation in the number of diffracting grains. This finding precludes using the AKK formulations for estimating the number of particles needed in an irradiated volume interacting with the incoming X-ray beam for obtaining a specified relative standard deviation in peak intensity.

Analysis of the angular distribution of the activated poles within the reflection band showed that their fraction decreased with increasing θ. This sampling bias was first predicted by Lorenz (Bijvoet *et al.*, 1969); it is caused by the variation of the $\cos\theta$ term (Eq. (4-a)) over the width of the reflection band, and causes the Bragg peak position to shift to lower diffraction angles, corresponding to spurious "strains". This sampling bias was eliminated by multiplying the activated pole counts within the reflection band by the factor, $\boldsymbol{CT_2} = \cos\theta_B / \cos\theta$. This term removed the sampling bias while maintaining fidelity to the number of activated poles. Employing the traditionally recommended term, $\boldsymbol{CT_1} = 1/\cos\theta$, also mitigated most of the bias. However, this correction did not preserve the number of poles within the reflection band; for the 200 reflection of our ideal gold nanopowder sample the overall number of poles were increased by almost 4 fold.

A simple numerical computation showed that, for an ideal random powder with identical, monosized, spherical crystallites, almost all of the activated poles which contributed finite intensity to the diffraction

peak of a given reflection were within a corresponding reflection band of width 2β. The total number of poles in this band could be predicted within 1% of their true value by the "Modified Lorenz Equation" (Eq. (4-c)), which is obtained from the band's area fraction (Figure 2). This 1% error is due to the crystallographic correlation between the poles from individual grains and cannot be eliminated.

The activated poles within the reflection band are divided into two subgroups. The "independent poles" or singlets have no other member of the <hkl> family within the band. "*n-lets*", on the other hand have "*n*" perfectly correlated members, all referred to a particular crystallite. The largest "*n-let*" dimension, "*n*" is half of the multiplicity of the particular reflection, $m_{hkl}/2$. The grain fractions which contribute "n" poles to the reflection band cannot be predicted from simple area fractions due to the crystallographic correlations of the probabilities. It was interesting to note that, for n>2, crystallographic orientation relationships can limit the formation of *n-lets* to specific wavelengths, which we termed "enhanced-selection wavelength" λ_{ES}. For example, for the h00 reflection from a cubic particle, triplets could only be observed for $\lambda_{ES} = \frac{2d_{h00}}{\sqrt{3}}$. The presence of higher dimension "n-lets" can complicate the analysis of diffraction data; this is analogous to some grains voting more than once in the diffraction process.

In summary, rigorous grain-by-grain forward modelling of the diffraction process provides useful checkpoints for the correct analysis of diffraction data from nanoparticle powder ensembles. Such modelling enables direct comparison of sampling statistics to intensity data and can be useful in elucidating the meaning of "diffraction average" terms as they relate to polydisperse samples. We note, however, that our analysis yields a lower uncertainty limit since we did not consider complications due to peak overlaps over the entire diffraction spectra. In addition, the use of Fourier transforms in rocking curve calculations, which assume exact particle shapes, can be overly simplistic when the particle size is very small or in cases where the particles have irregular shapes. In such cases, least squares refinement of nanoparticle diffraction data will yield larger uncertainties associated with the (extracted) average structural parameters. We are currently working on these issued and will report our findings in a future article.

Acknowledgements We would like to thank Chi-Chang Kao and James Misewich for encouraging us to work in this area and Mikhail Treger and Seung-Yub Lee for valuable discussions. We also thank Mr.

Art Ellis from IBM Research for his contributions to Figures 3 and 4-a and Ms. Connie Phung for her contribution to Figure 4-b. This work was supported by the US Department of Energy, Office of Science under contract number DE-AC02-98CH10886.

References

Alexander, L., Klug, H. P. & Kummer, E. (1948). *J Appl Phys* **19**, 742-753.
Bijvoet, J. M., Burgers, W. G., Hèagg, G. & International Union of Crystallography. (1969). *Early papers on diffraction of X-rays by crystals*. Utrecht: Published for the International Union of Crystallography by A. Oosthoek.
Buerger, M. J. (1940). *P Natl Acad Sci USA* **26**, 637-642.
Cullity, B. D. (1978). *Elements of x-ray diffraction*, 2d ed. Reading, Mass.: Addison-Wesley Pub. Co.
Fewster, P. F. (2014). *Acta Crystallogr A* **70**, 257-282.
Fitzsimmons, M. R., Eastman, J. A., Muller-Stach, M. & Wallner, G. (1991). *Physical Review B, Condensed Matter (USA)* **44**, 2452-2460.
Hall, B. D. (2000). *J Appl Phys* **87**, 1666-1675.
Kelly, A. & Knowles, K. M. (2012). *Crystallography and crystal defects*, 2nd ed. Chichester, West Sussex: Wiley,.
Kuffner, J. J. (2004). *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, pp. 3993-3998 Vol.3994.
Ozturk, H., Yan, H. F., Hill, J. P. & Noyan, I. C. (2014). *J Appl Crystallogr* **47**, 1016-1025.
Palosz, B., Grzanka, E., Gierlotka, S. & Stelmakh, S. (2010). *Z Kristallogr* **225**, 588-598.
Patterson, A. L. (1939a). *Phys Rev* **56**, 972-977.
Patterson, A. L. (1939b). *Phys Rev* **56**, 978-982.
Reynolds, R. C. & Reynolds, R. C. (1986). *Clays and clay minerals* **34**, 359-367.
Smith, D. K. (2001). *Powder Diffr* **16**, 186-191.
Sulyanov, S. N., Popov, A. N. & Kheiker, D. M. (1994). *J Appl Crystallogr* **27**, 934-942.
Won-Sik Han, S. O. K., Il-Hwan Suh (2007). *Korean Journal of Crystallography* **18**, 7-9.

Yan, H. (2006). PhD thesis, Columbia University.
Ying, A. (2010). PhD thesis, Columbia University.
Ying, A. J., Murray, C. E. & Noyan, I. C. (2009). *J Appl Crystallogr* **42**, 401-410.
Yinghua, W. (1987). *J Appl Crystallogr* **20**, 258-259.

LIST OF FIGURES

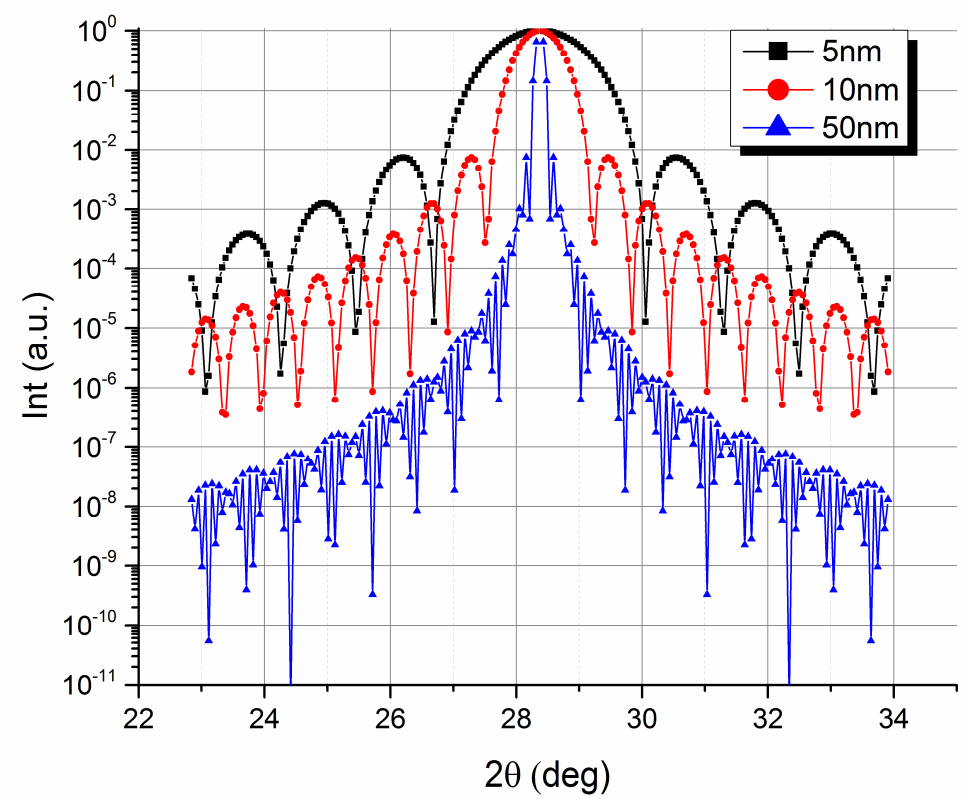


Figure 1 Rocking curve scans for a single, crystalline, gold nanoparticle computed around the 200 reflection. Each trace corresponds to a different particle size. The irradiation wavelength. λ, is 1 Angstrom, the Bragg angle is $2\theta_{200}$=28.36 degrees and the intensities are shown in log 10 scale.

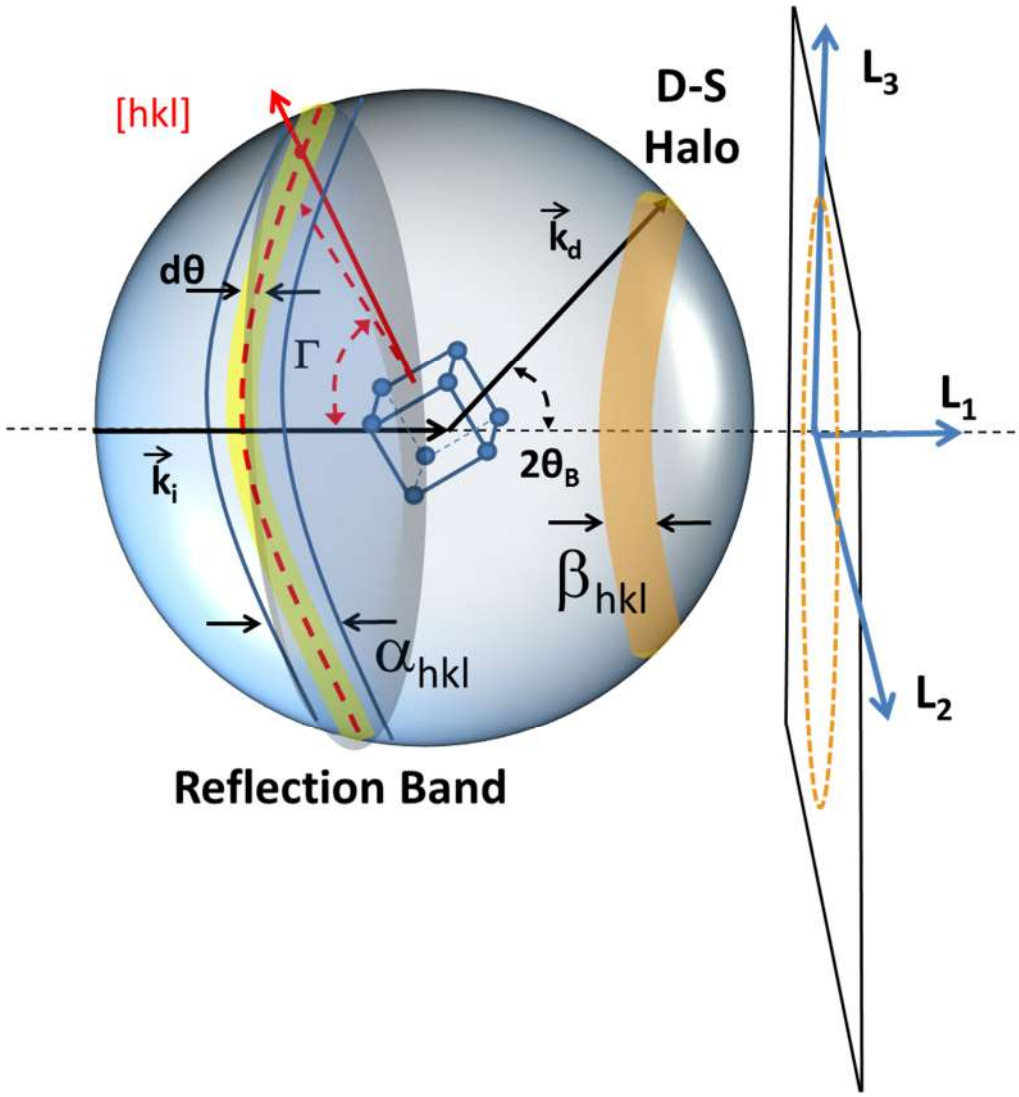


Figure 2 Schematic of the diffraction geometry for a single crystalline particle (shown by a simple cube). Here the reflection band contains the (activated) hkl poles which produce spots in the Debye-Scherrer band (D-S band). The projection of these spots onto the 2-dimensional detector forms the polycrystalline Debye ring. (This figure is adapted from our earlier work (Ozturk *et al.*, 2014).)

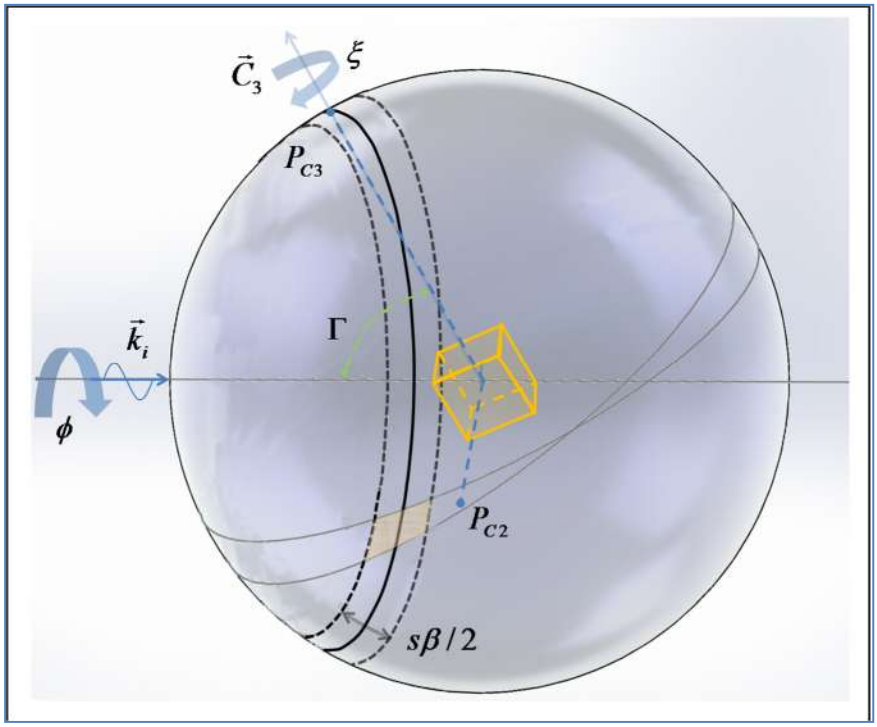


Figure 3 Diffraction geometry for forming an h00 doublet. For this condition both poles, P_{C_3} and P_{C_2} must fall within the reflection band of width $s\beta/2$. For the case shown, where P_{C_3} falls exactly on the reflection circle corresponding to the Bragg angle, a ξ rotation around the $\vec{C}_3$ vector is needed to rotate P_{C_2} into the shaded area defined by the intersection of the spherical lune with the reflection band. (This figure was created with the assistance of Mr. A. Ellis of IBM Research Division)

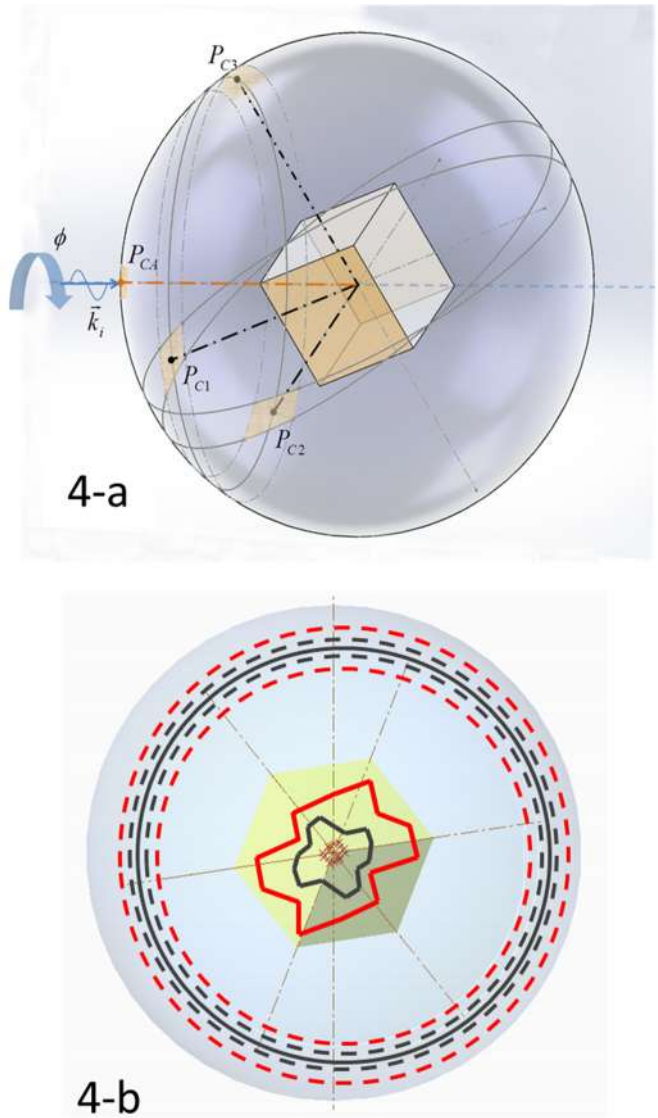


Figure 4 -a) Diffraction geometry when three h00 poles belonging to one crystallite with cubic symmetry (an h00 triplet) fall exactly on the reflection ring corresponding to the Bragg angle. This condition is automatically satisfied when a <111> body diagonal is antiparallel with the incident beam vector. **4-b)** Schematic of the areas swept by the pole of the $\vec{C}_A$ vector while keeping all three h00 poles within the reflection band; the red and black lines denote two reflection bands of 5 and 10 degrees angular width, respectively (not-to-scale). Here the viewer is assumed to be looking at the reference sphere along the incident beam vector, $\vec{k}_i$. (These figures were created with the assistance of Mr. A. Ellis of IBM Research Division and Ms. Connie Phung of the Department of Mechanical Engineering, Columbia University)

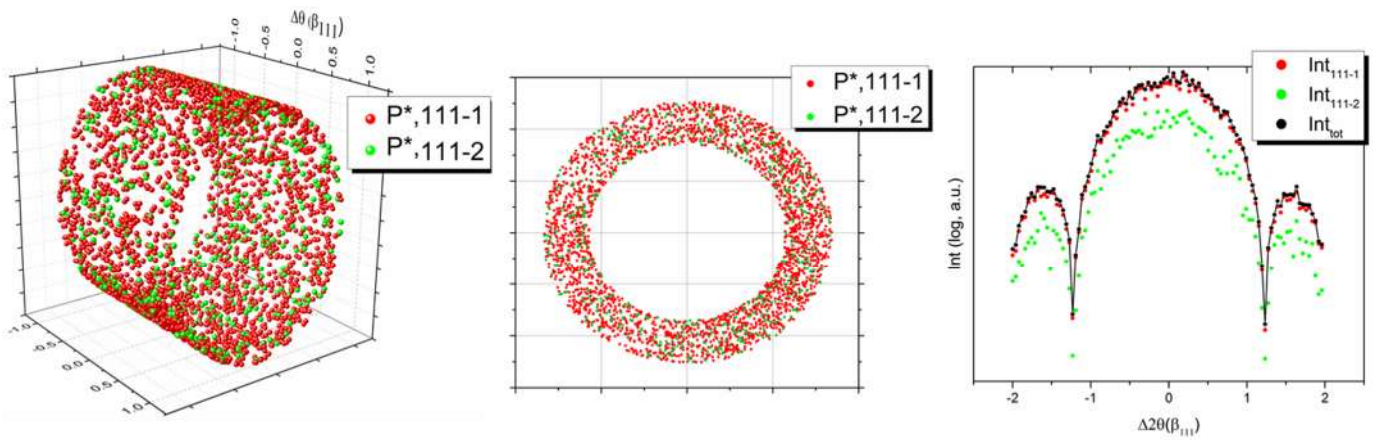


Figure 5 -a) The distribution of all singlet (red) and doublet (green) poles falling within the reflection band (s=4) for the gold 111 reflection, when 10^4 particles, each assumed to be a perfect 2.86 nm diameter sphere, are irradiated by X-rays of , λ=1Å. The reflection band is a conical segment with varying leading and trailing radii. Here the Δθ-axis corresponds to the acceptance angle of the crystallite ($s\beta/2$, s=4) which is the width of the reflection band. **5-b)** The corresponding 111 Laue spots on the detector, obtained from the intersection of the diffracted beams with the detector plane. A transmission pin-hole pattern with no texture is observed[9]. **5-c)** The "caked" intensity profile (in Log 10 scale) corresponding to the Laue spot distribution shown. The profiles represented by red and green dots correspond to the intensity contribution from uncorrelated and doubly-correlated 111 poles respectively. The profile shown with black dots is the total intensity from all activated 111 poles.

[9] The relative intensities of the Laue spots are not included for ease of viewing.

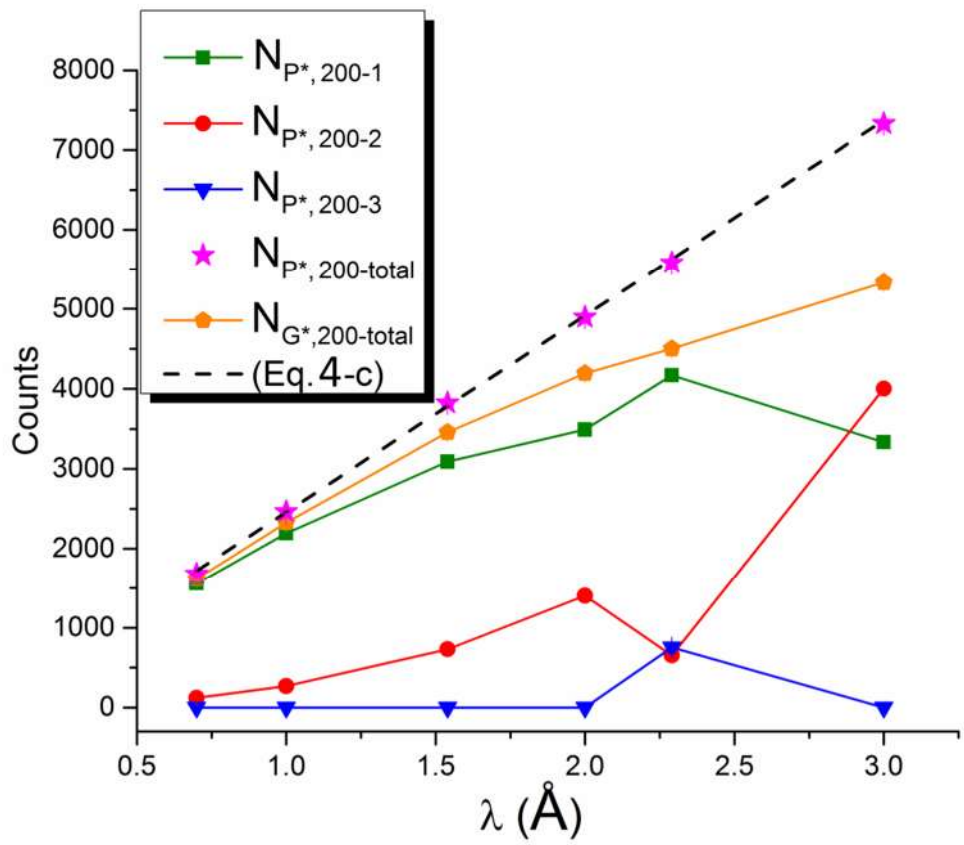


Figure 6 Dependency of the total population of diffracting poles, $N_{P*,hkl}$ and the populations of its subsets (n-lets) on irradiation wavelength for the 200 reflection. The data are obtained from 10^4 identical, gold nanoparticles with 2.86 nm diameter. The total population of the diffracting grains, $N_{G*,200}$ is also shown, along with the estimated values from Eq. (4-c). The simulation was repeated ten times, with independent populations. The statistical error bars are comparable to the size of the symbols.

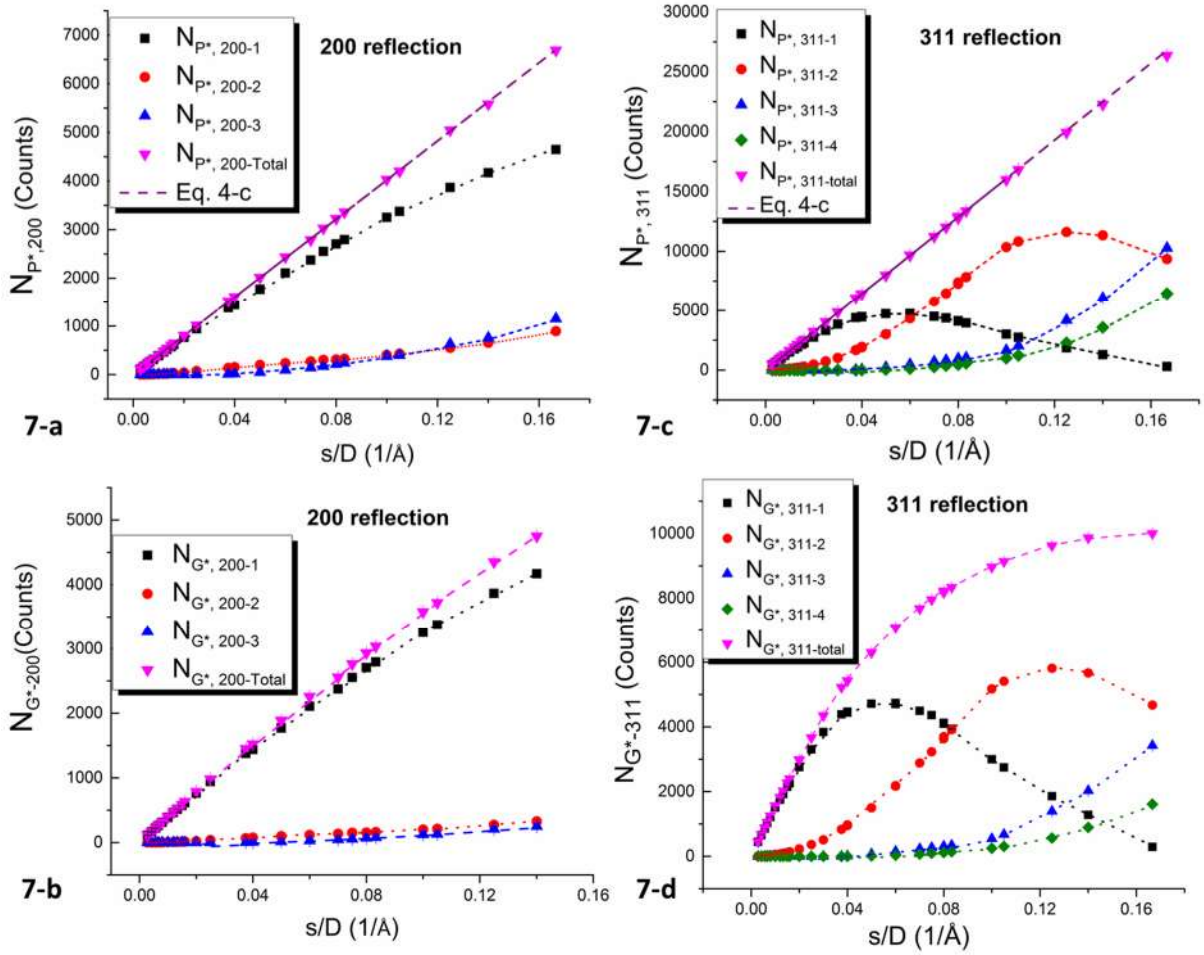
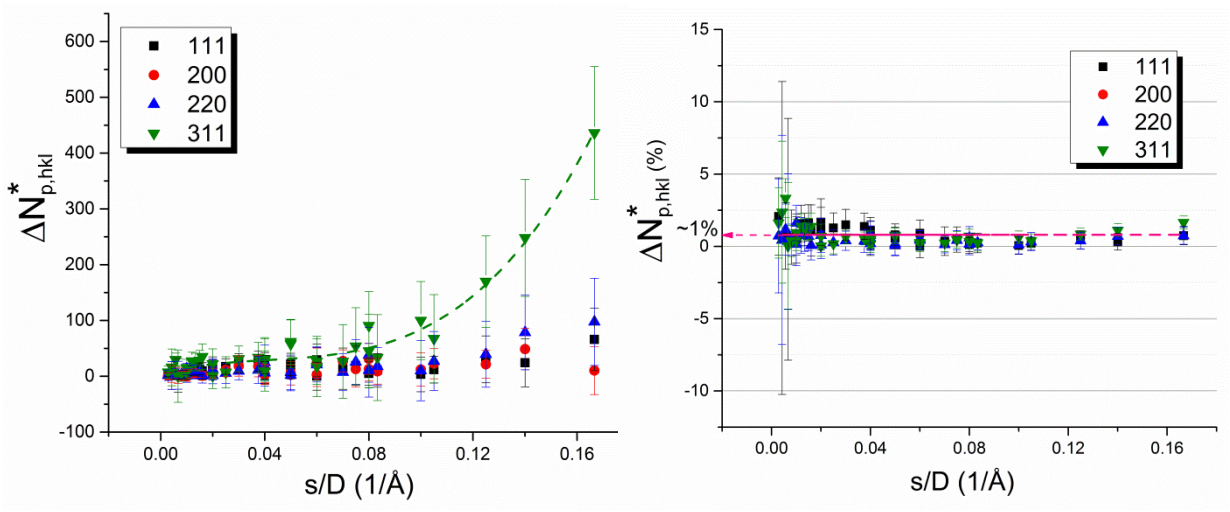


Figure 7 Variation of the diffracting pole populations, $N_{P^*,hkl}$ and their subsets with the normalized selection range, s/D , obtained for the simulation of 200 (Fig. **7-a)** and **7-c**) and 311 reflections (Fig. **7-b)** and **7-d**) when 10^4 gold spheres of various diameters were irradiated with 5.41 keV monochromatic X-rays. (The parameters considered are shown in Table 1.) In Figures **7-b)** and **7-d)**, the behaviours of the corresponding diffracting grain populations are shown.



8-a **8-b**

Figure 8 –**a**) The difference between the total number of poles obtained from simulation and Eq. (4-c), $\Delta N_{P^*T,hkl} = (N_{P^*T,hkl})_{sim} - (N_{P^*T,hkl})_{Eq.(4-c)}$, plotted as a function of the normalized selection range, s/D. In Figure **8-b**) $\Delta N_{P*,hkl}$ is expressed as the percent fraction of $N_{P*,hkl}$.

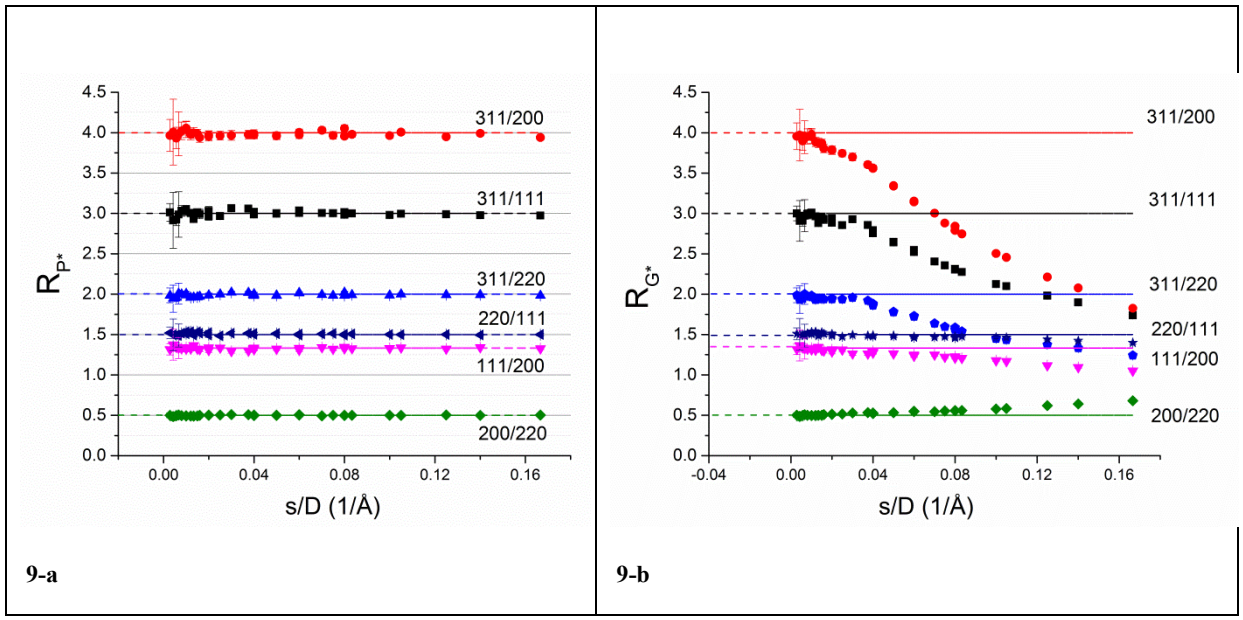
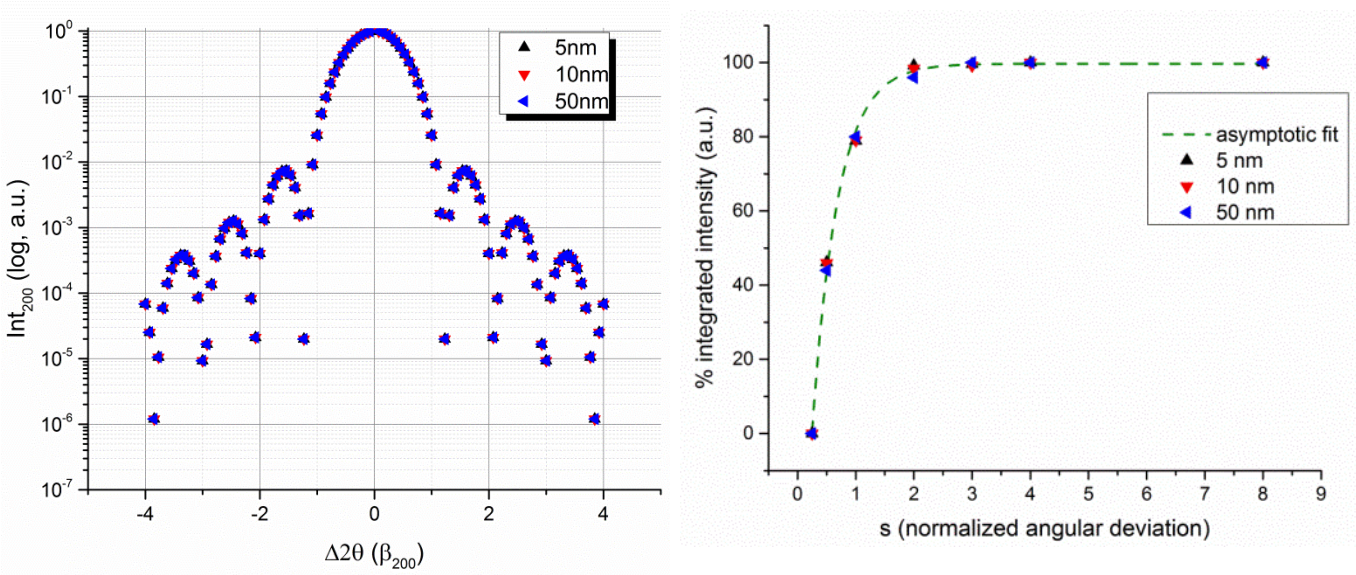
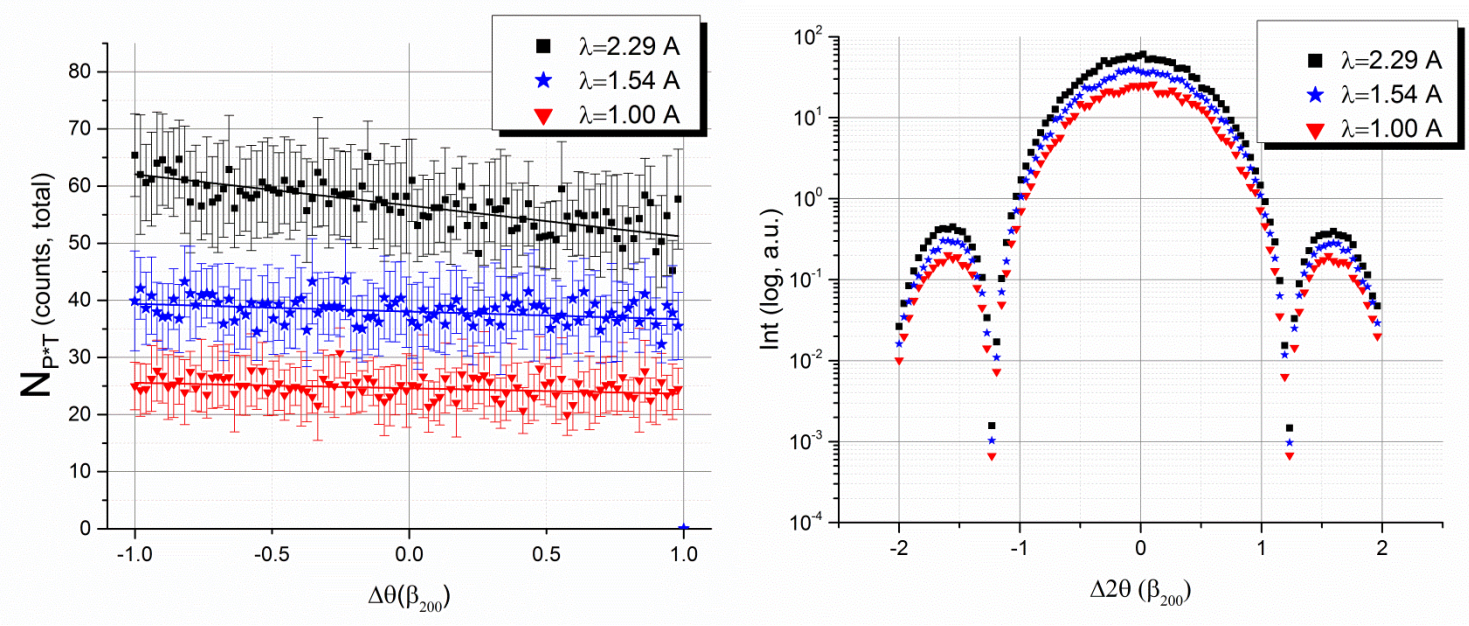


Figure 9 -a) Ratios of $\left(N_{P*,hkl}\right)_{sim}$ from various reflections: $R_{P*} = N_{P*,hkl} / N_{P*,h'k'l'}$ plotted as a function of the normalized selection range, s/D. The corresponding multiplicity ratios, $m_{hkl}/m_{h'k'l'}$ are shown as dashed lines. In Figure **9-b**, the corresponding "total diffracting grain" ratios are shown.



10-a **10-b**

Figure 10 **-a)** Rocking curve scans around 200 reflection for three spherical crystalline gold nanoparticles with 5, 10 and 50 nm diameters. The horizontal axis is scaled such that the deviation from the Bragg angle $\Delta 2\theta$ is given in multiples of the Scherrer breadth. (The data corresponding to 5, 10 and 50 nm particles coincide almost perfectly) - **b)** Integrated intensity vs the extent of the Bragg peak**.**



11-a **11-b**

Figure 11 **-a**) The distribution of the total activated pole populations around the 200 reflection for three irradiation wavelengths. Average values and their standard deviations of data from ten independent simulations, each with 10^4 identical spherical Au particles are shown. **11-b**) The corresponding intensity profiles of the 200 pole distributions on the left. This calculation includes only the mean pole counts and

excludes the atomic scattering factor, F_{hkl}, as well as the N_C^2 term in Eq. (3) to isolate and visualize the effect of the angle-dependent selection bias.

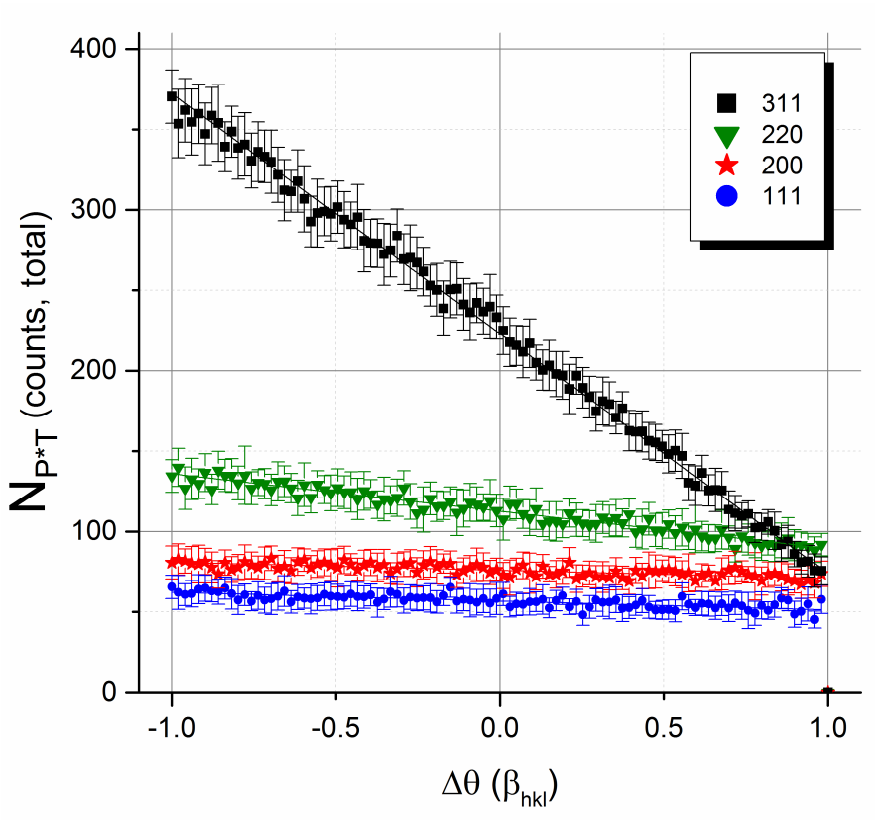


Figure 12 The distribution of the total activated pole populations around the first four Bragg reflections for a random gold powder sample of 10^4 particles irradiated with 2.29 A wavelength X-rays. The individual crystallite diameter is 2.86 nm. The data points correspond to mean counts of ten simulations whereas the error bars show the standard deviation from the mean.

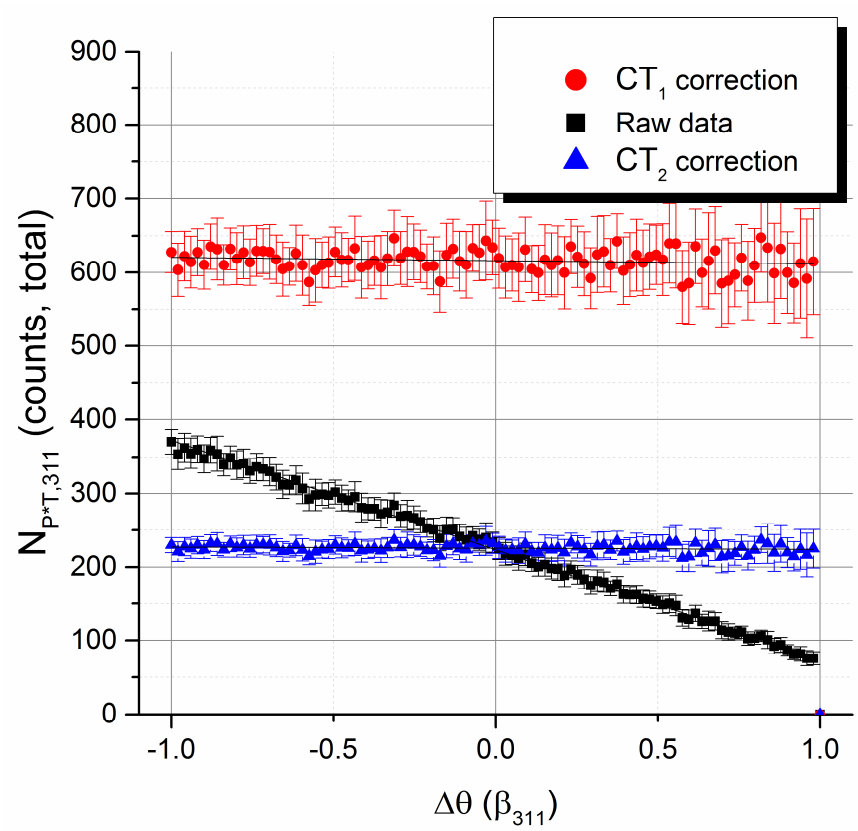


Figure 13 Correction of the sampling bias of the as-simulated 311 pole population (also shown in Figure 12, black trace) after multiplication with CT_1 (red trace) and CT_2 (blue trace) terms.