PDFgui – a small box modelling platform for nanoscale structure analysis

Emil S. Bozin
Institute of Physics Belgrade

January 12, 2026

A reflection on PDFgui usage: increasing popularity

PDFfit2 and PDFgui: computer programs for studying nanostructure in crystals

Farrow, CL; Juhas, P; (...); Billinge, SJL
3rd Workshop on Reverse Monte Carlo Methods

Aug 22 2007 | JOURNAL OF PHYSICS-CONDENSED MATTER 19 (33)

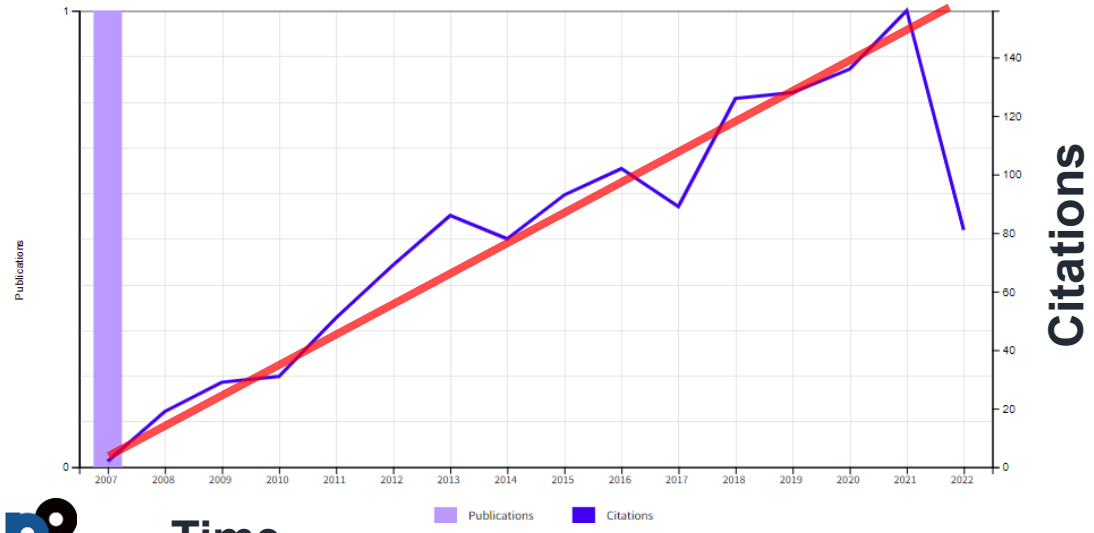
PDFfit2 is a program as well as a library for real-space refinement of crystal structures. It is capable of fitting a theoretical three-dimensional (3D) structure to atomic pair distribution function data and is ideal for nanoscale investigations. The fit system accounts for lattice constants, atomic p ... Show more

Full Text at Publisher

1,276
Citations

28
References

Clarivate
Web of Science™



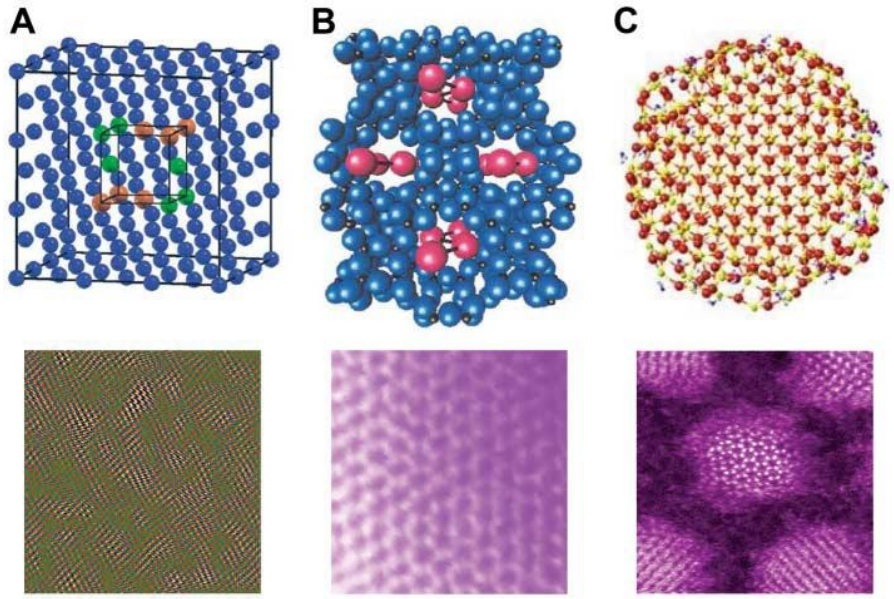
1 Publication	Sort by: Citations: highest first	1 of 1	Citations						
			< Previous year				Next year >	Average per year	Total
			2018	2019	2020	2021	2022		
	Total		126	128	136	156	81	79.75	1,276
⊖ 1	PDFfit2 and PDFgui: computer programs for studying nanostructure in crystals Farrow, CL; Juhas, P; (...); Billinge, SJL 3rd Workshop on Reverse Monte Carlo Methods Aug 22 2007 | JOURNAL OF PHYSICS-CONDENSED MATTER 19 (33)		126	128	136	156	81	79.69	1,275

Overview

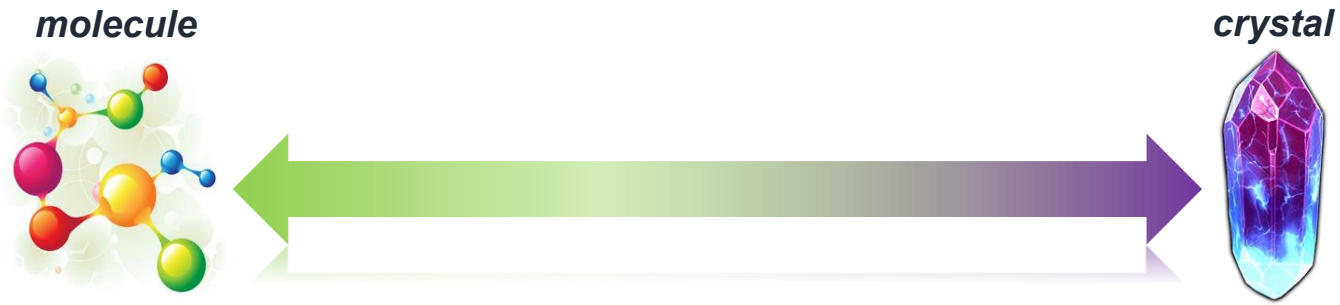
- Small box modelling and PDFgui
- Notes of relevance for PDF analysis
- PDFgui parameters & program structure
- Illustrative PDFgui-based science example

PDF approach

- Non crystalline materials
  (liquids, amorphous solids, polymers)
- Nanoscale materials
- Disordered crystalline systems with nanoscale heterogeneities



S.J.L. Billinge and I. Levin, **The Problem with Determining Atomic Structure at the Nanoscale**, *Science* **316**, 561 (2007).



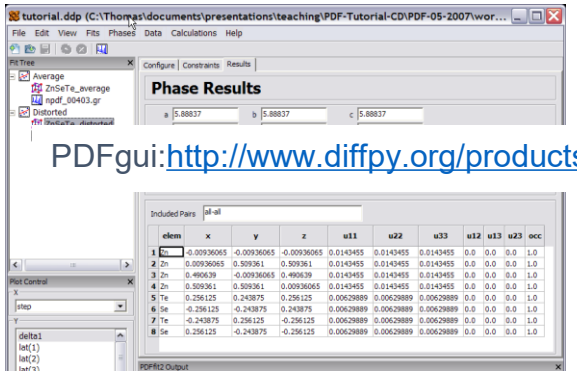
PDF data modeling

Small Models: Least Squares Refinement
Up to several hundreds of atoms
'Rietveld'-type parameters: *lattice parameters, atomic positions, displacement parameters, etc.*
Refinements as function of *r*-range

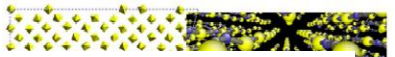
Large Models: Reverse Monte Carlo
20000 + atoms
Fit X-ray and neutron *F(Q), G(r)* , Bragg profile
Constraints utilized
Static 3-D model of the structure (a snap-shot)

Multi-level /Complex Modeling
Refine higher level parameters (not each atom)
Example nanoparticle: *diameter, layer spacing, stacking fault probability*
Choose minimization scheme

Emerging: *ab initio* and force-field based approaches
Density Functional Theory
Molecular Dynamics

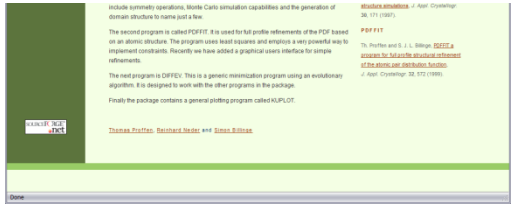


PDFgui:http://www.diffpy.org/products/pdfgui



RMCprofile:http://wwwisis2.isis.rl.ac.uk/rmc/
EPSR: www.facebook.com/disord.matt

DIFFEV and DISCUS: http://discus.sourceforge.net
DiffPy-CMI: http://www.diffpy.org/products/diffpycmi

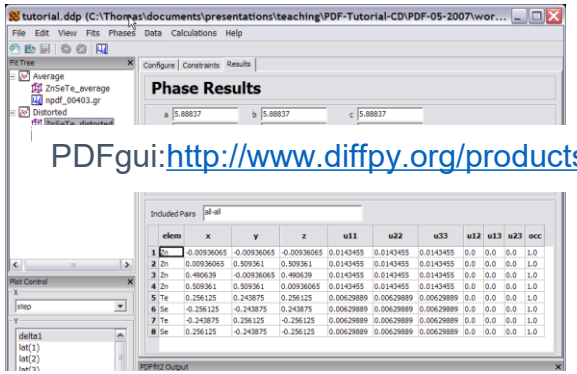


slide courtesy of Katharine Page

PDF data modeling

PDFgui

Small Models: Least Squares Refinement
Up to several hundreds of atoms
'Rietveld'-type parameters: *lattice parameters, atomic positions, displacement parameters, etc.*
Refinements as function of *r*-range

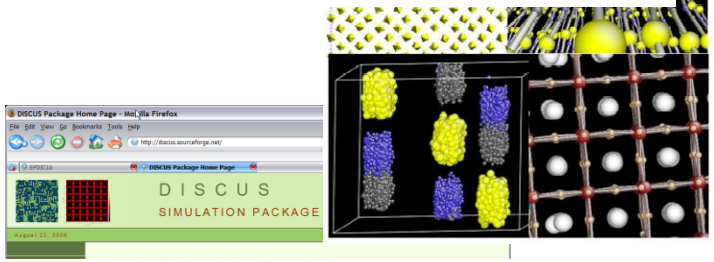


PDFgui:http://www.diffpy.org/products/pdfgui

Large Models: Reverse Monte Carlo
20000 + atoms
Fit X-ray and neutron *F(Q), G(r)* , Bragg profile
Constraints utilized
Static 3-D model of the structure (a snap-shot)

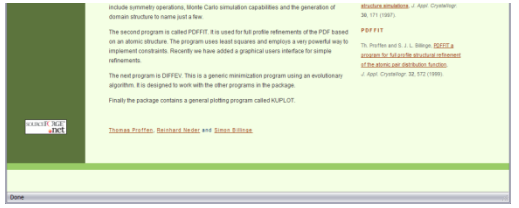
RMCprofile:http://wwwisis2.isis.rl.ac.uk/rmc/
EPSR: www.facebook.com/disord.matt

Multi-level /Complex Modeling
Refine higher level parameters (not each atom)
Example nanoparticle: *diameter, layer spacing, stacking fault probability*
Choose minimization scheme



DIFFEV and DISCUS: http://discus.sourceforge.net
DiffPy-CMI: http://www.diffpy.org/products/diffpycmi

Emerging: *ab initio* and force-field based approaches
Density Functional Theory
Molecular Dynamics



slide courtesy of Katharine Page

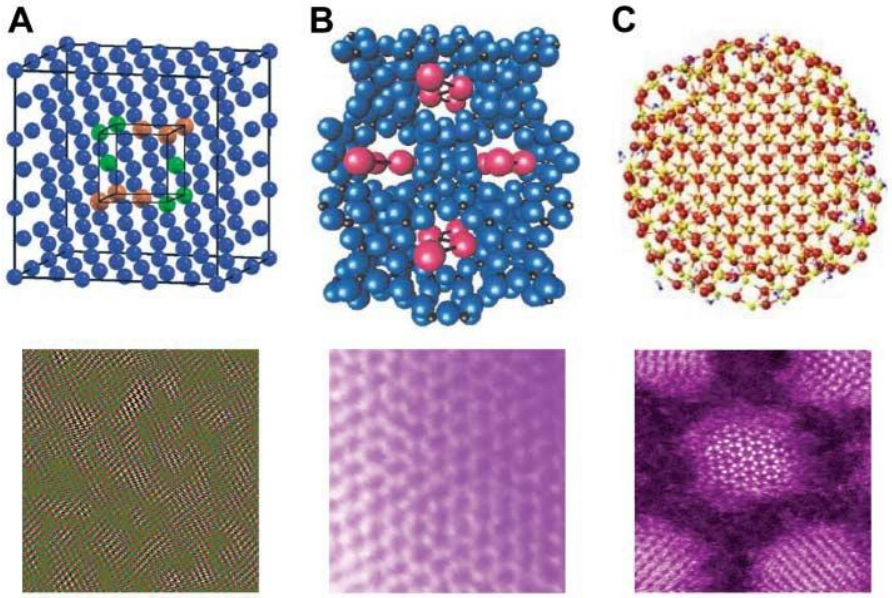
PDF approach

- Non crystalline materials
  (liquids, amorphous solids, polymers)
- Nanoscale materials
- Disordered crystalline systems with nanoscale heterogeneities

structure

PDFgui

symmetry breaking



S.J.L. Billinge and I. Levin, **The Problem with Determining Atomic Structure at the Nanoscale**, *Science* **316**, 561 (2007).

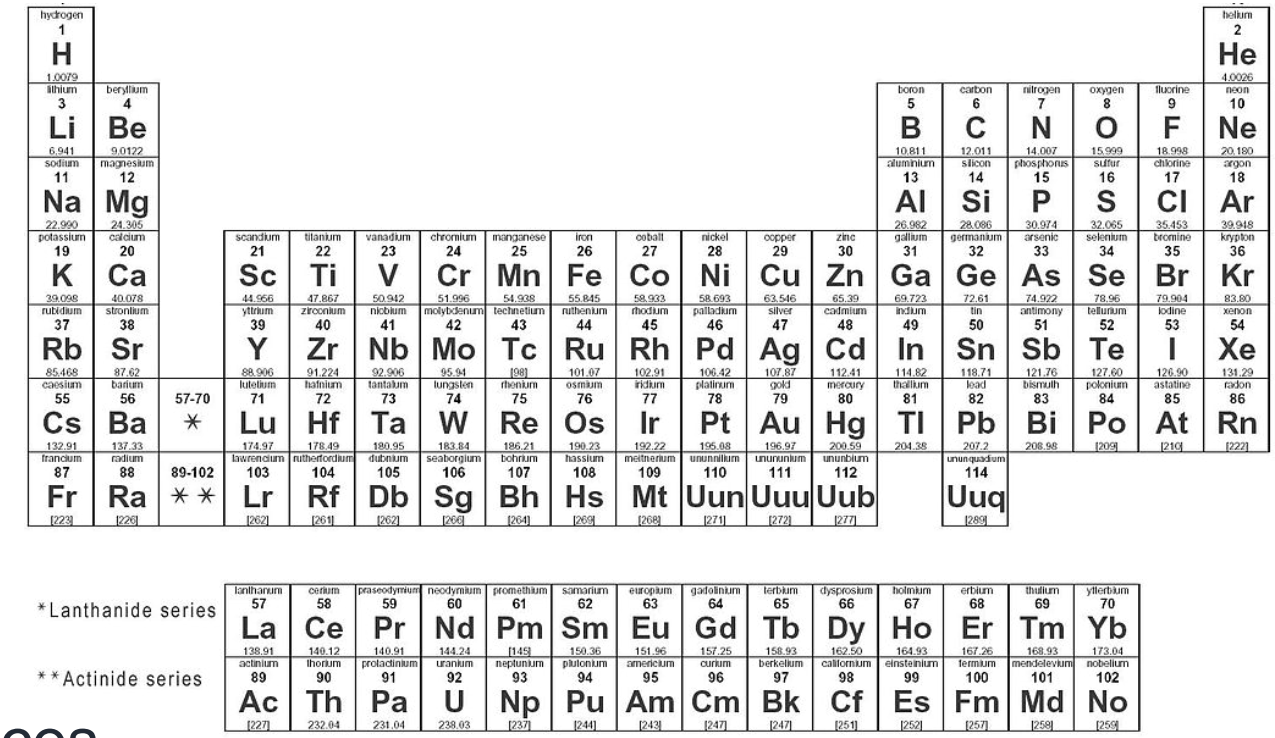
Small box PDF modeling approach

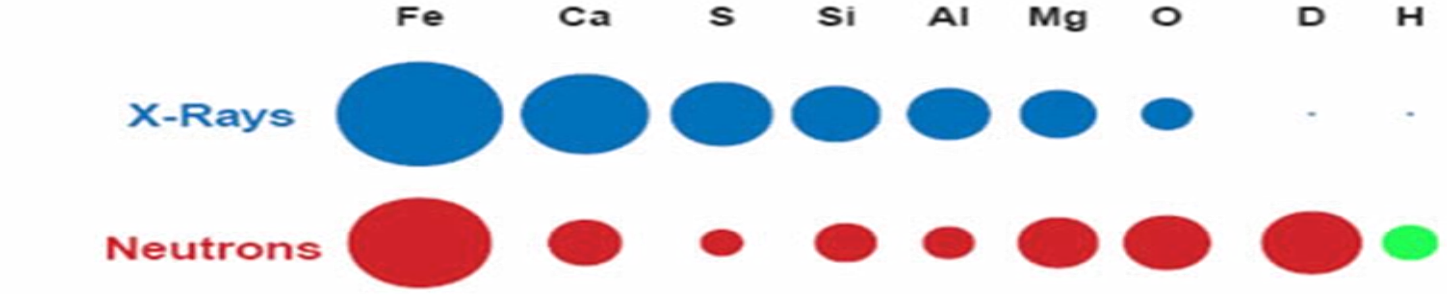
- Small box: assumption of periodic boundary conditions (P1)
- Relatively small number of atoms (up to several hundred)

- Built-in symmetry constraints with symmetry equal to or **usually lower** than the average crystal symmetry
- Involves least squares refinement over selected *r*-range (typically up to a few unit cells, translational symmetry not relevant for narrow range fits, when box size effectively provides “metrics”)

PDF approach

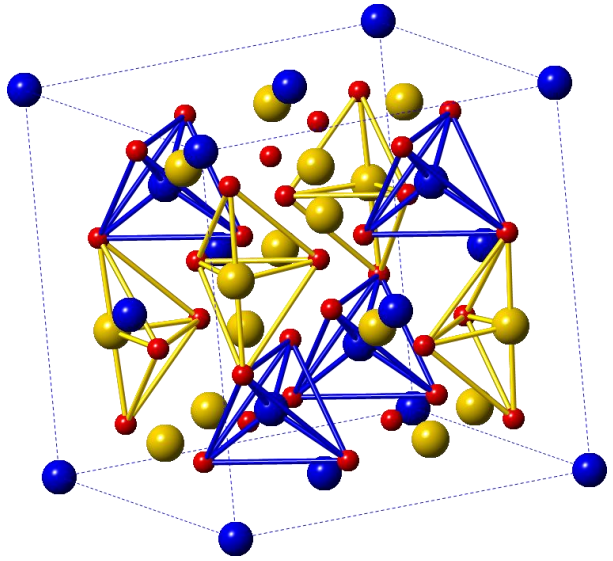
- Considering scattering contrast
- Considering absorption
- Considering isotopes and resonances

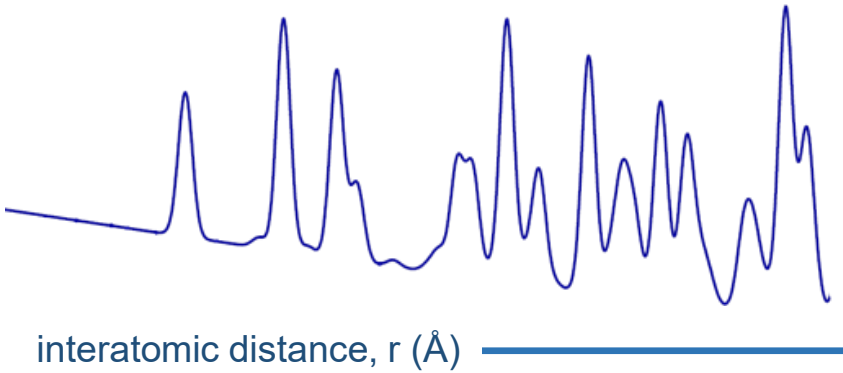




PDFgui handles
xPDF & nPDF

PDF approach





Experiment

Raw data

Data reduction

PDF

Analysis

$$I_{tot} = I_{sam} + I_{bgrd}$$

$$I_{sam} = A \cdot P \cdot [C \cdot I_{coh} + I_{inc} + I_{mul}]$$

$$S(Q) = \frac{I_{coh}(Q) - \langle b^2 \rangle + \langle b \rangle^2}{\langle b \rangle^2}$$

$$Q = \frac{4\pi \sin\theta}{\lambda}$$

experiment

$$G(r) = \frac{2}{\pi} \int_0^\infty Q\,(S(Q) - 1) \sin Qr \, \mathrm{d}Q$$

PDFgui does **NOT** fix data problems

relationship to structure

model

$$G(r) = \frac{1}{Nr\langle b \rangle^2} \sum_{i \neq j} b_i b_j \, \delta(r - r_{ij}) \; - \; 4\pi r \rho_0$$

PDFgui provides a PDF calculator

PDF approach

Relationship to structure

r_{ij} – interatomic distance between atoms *i* and *j*

$$G(r) = \frac{1}{Nr\langle b\rangle^2} \sum_{i \neq j} b_i b_j \delta(r - r_{ij}) - 4\pi r \rho_0$$

$\rho_o = N/V$ average number density
(number of atoms per unit volume)

r – running variable (distance)

scattering "scale" for pair of atoms *i* and *j*
weighted by averaged scattering "strength"

Neutrons: b_{coh}
X-rays: f(Q=0), effectively Z

Experimental setup limitations

Truncation effects in Fourier transform
(termination ripples, limited *r*-resolution)

$$G(r) = \frac{2}{\pi} \int_{\cancel{0}}^{\cancel{\infty}\ Q_{MAX}} Q\,(S(Q) - 1) \sin Qr \, \mathrm{d}Q$$

$\cancel{0}\ Q_{MIN}$

Small angle scattering
information is (usually) missing

Ideally one would like to have

- broad Q-range
- good Q-resolution

This is not always essential!

PDFgui – awareness of various effects

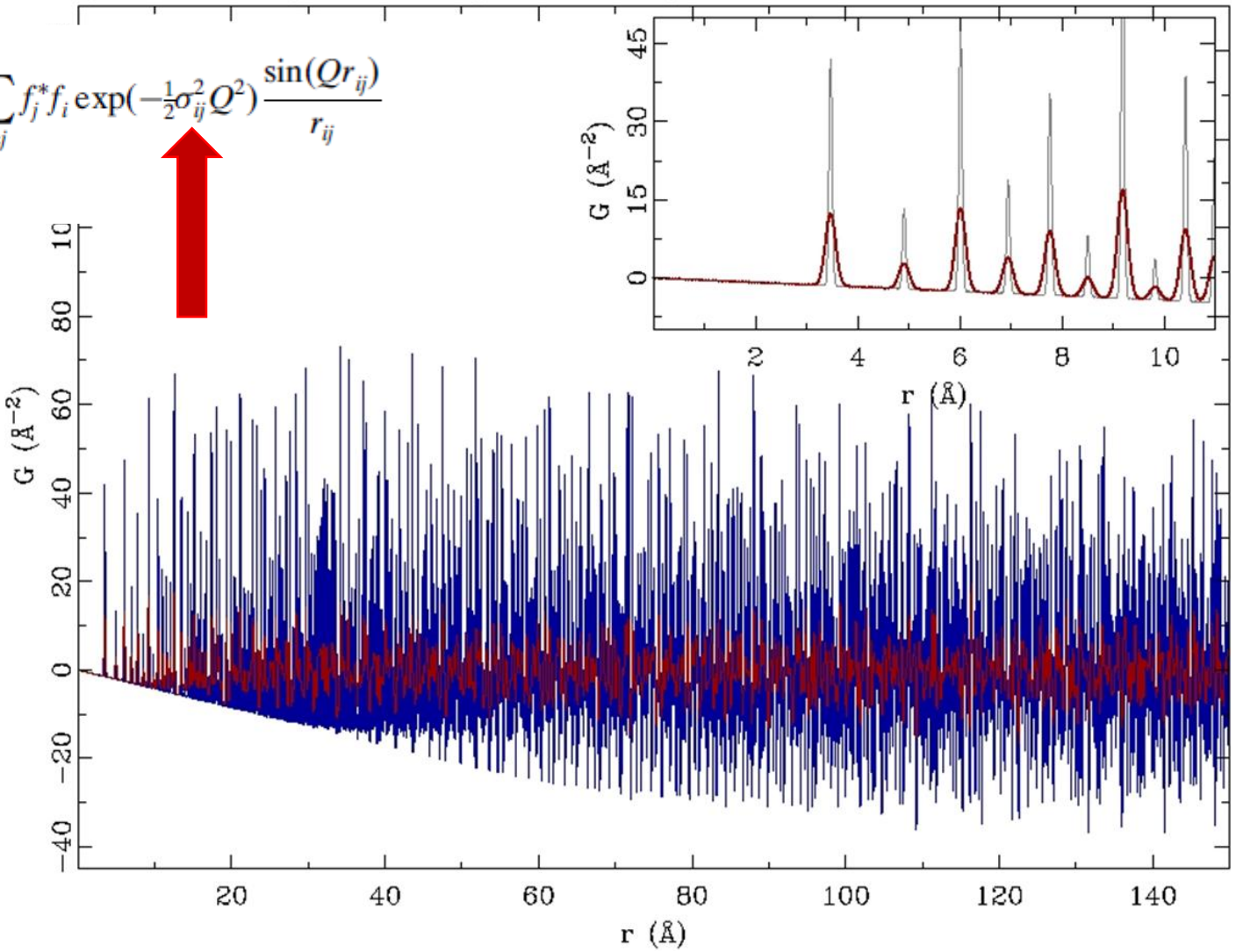
Some effects that should be accounted for in data modelling

- Thermal broadening
- Correlated motion of nearest neighbors
- Finite Q_{max} (truncation)
- Limited Q-space resolution (field of view)
- Particle size

Effect of thermal broadening

Thermal vibrations of atoms broaden PDF peaks
anisotropic atomic displacement parameters (ADPs)

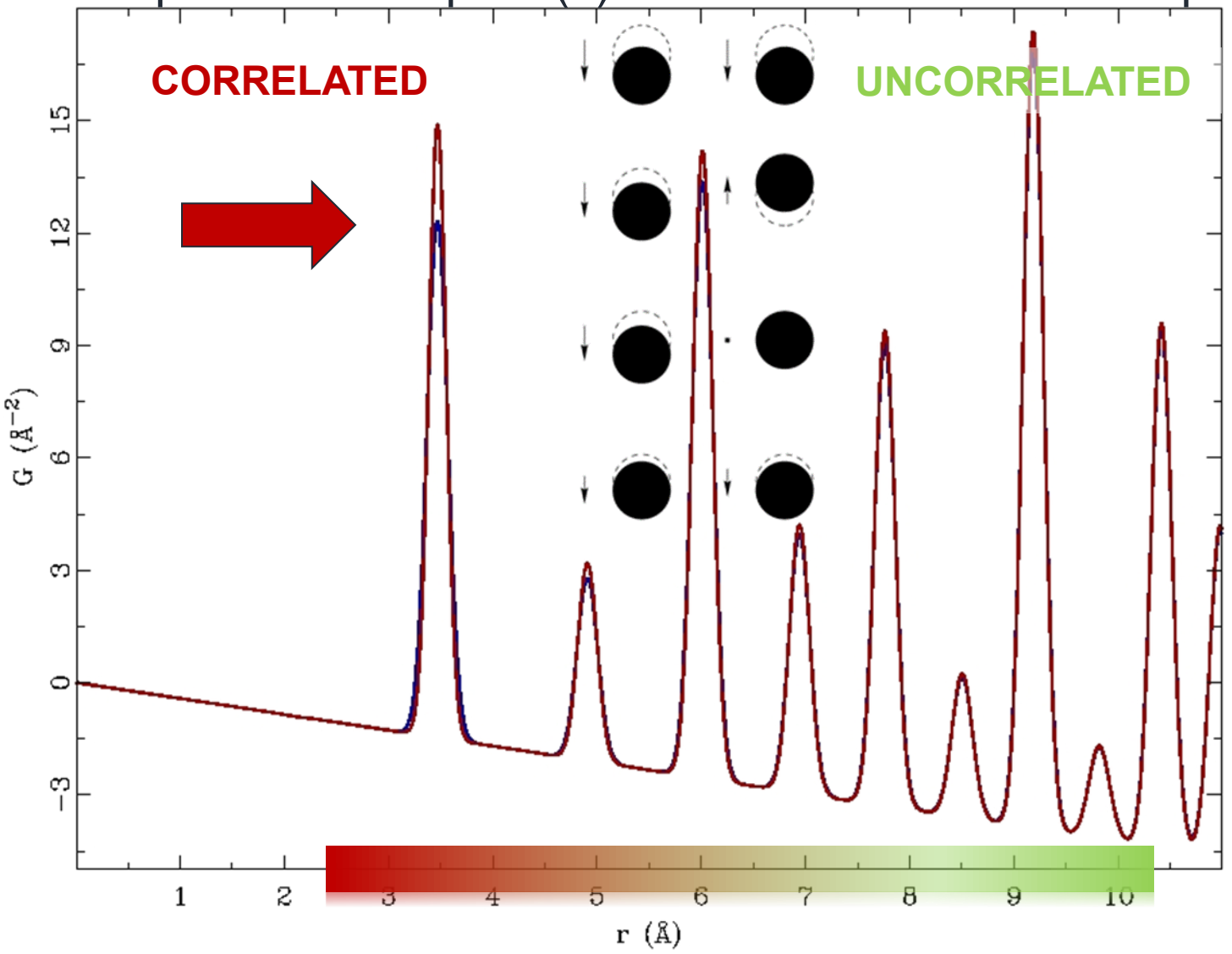
$$F(Q) = \frac{1}{N\langle f\rangle^2}\sum_{i\neq j} f_j^* f_i \exp(-\tfrac{1}{2}\sigma_{ij}^2 Q^2)\frac{\sin(Qr_{ij})}{r_{ij}}$$



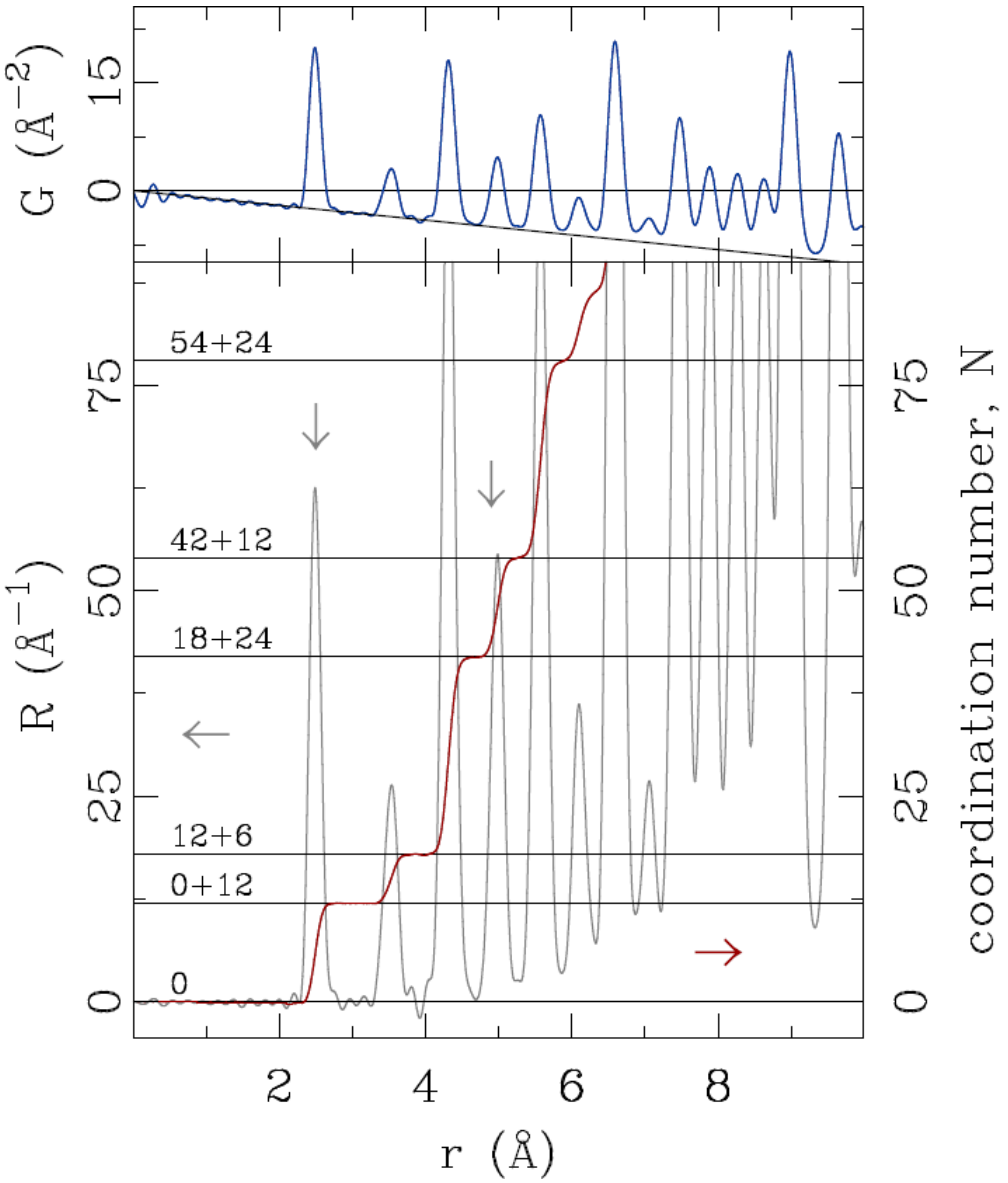
PDFgui accounts for ADPs

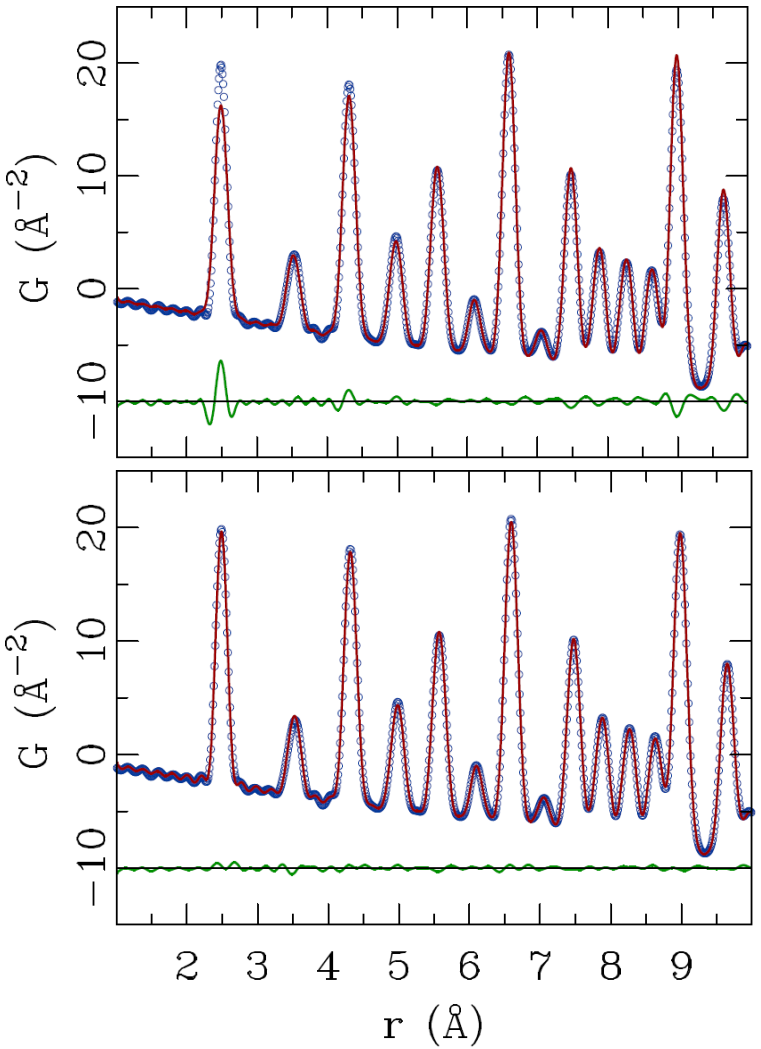
Effect of correlated atomic motion

Nearest neighbors tend to vibrate in correlated manner
this sharpens the NN peak(s) in PDF relative to distant peaks



Effect of correlated atomic motion

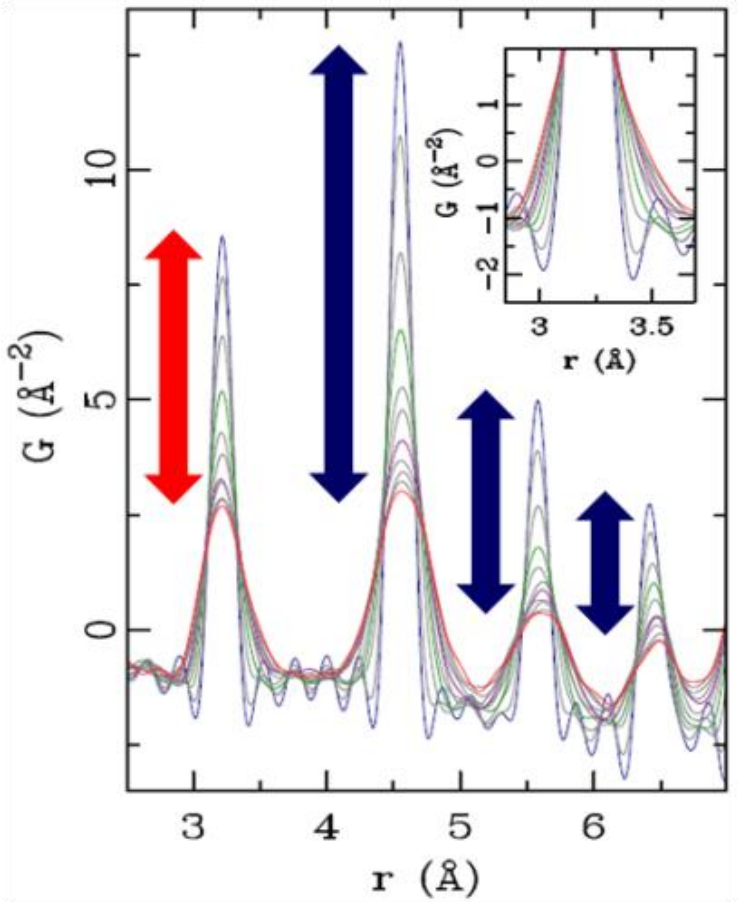




PDF correlated atomic motion outlaws

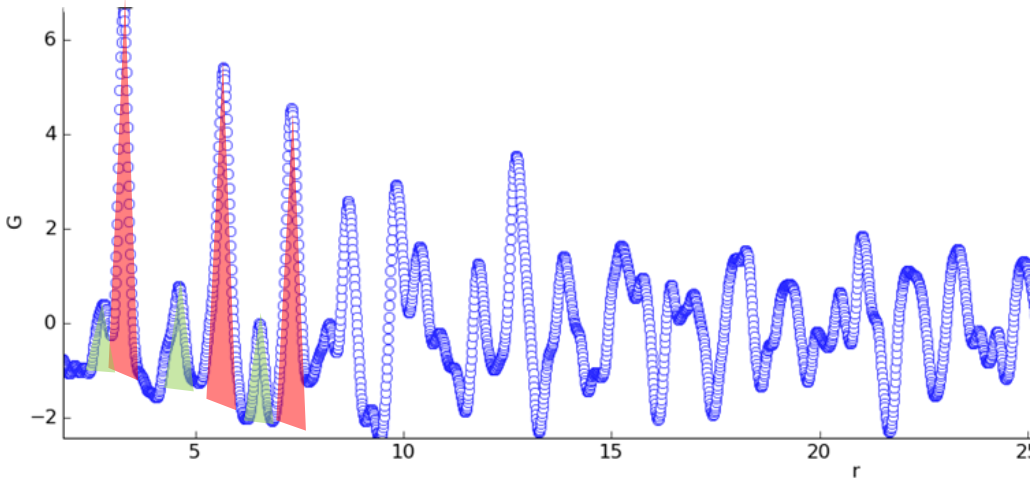
Weak effects: soft bonding
(e.g. PbTe)

break the rule



Enhanced effects: strong bonding
(e.g. in $CeCoIn_5$)

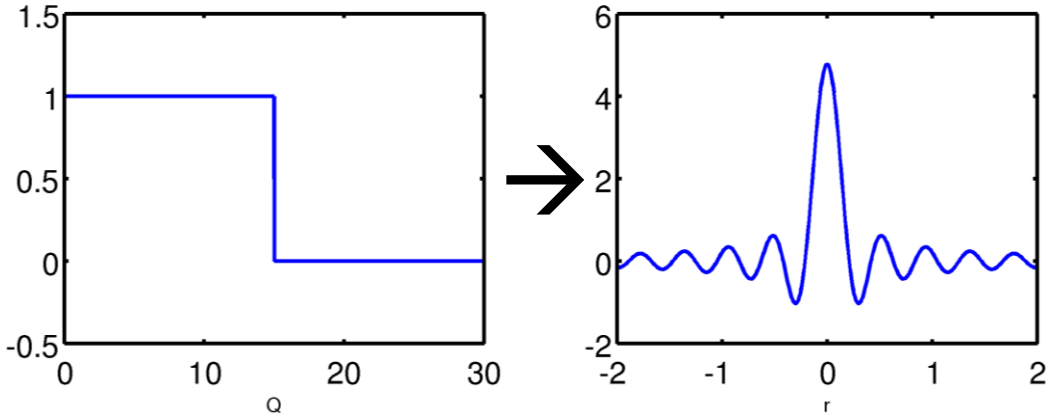
difficult to model

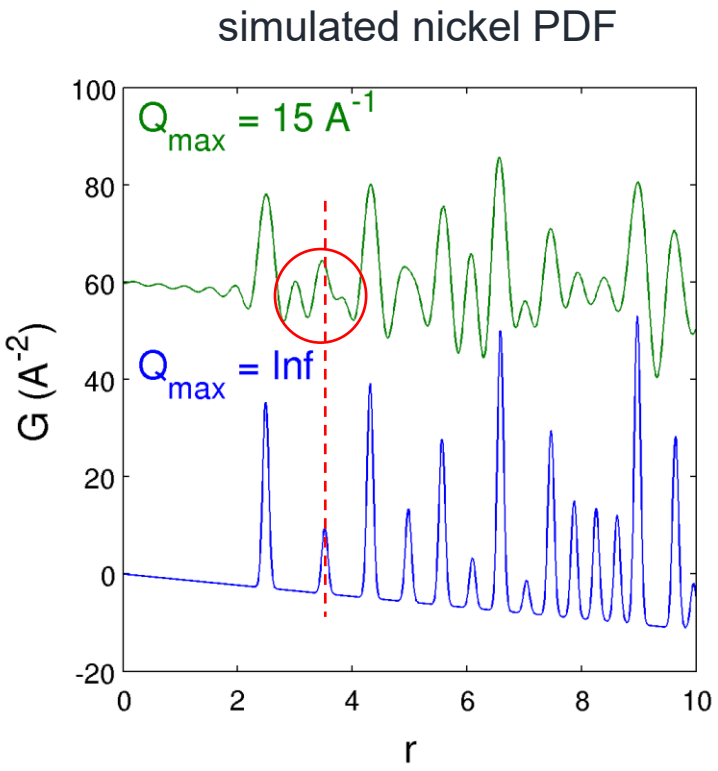


PDFgui handles most correlated motion cases (empirical corrections)

Effect of finite data range (Q_{max} truncation)

$$G(r) = \frac{2}{\pi} \int_0^{Q_{max}} F(Q) \sin Qr \, dQ$$



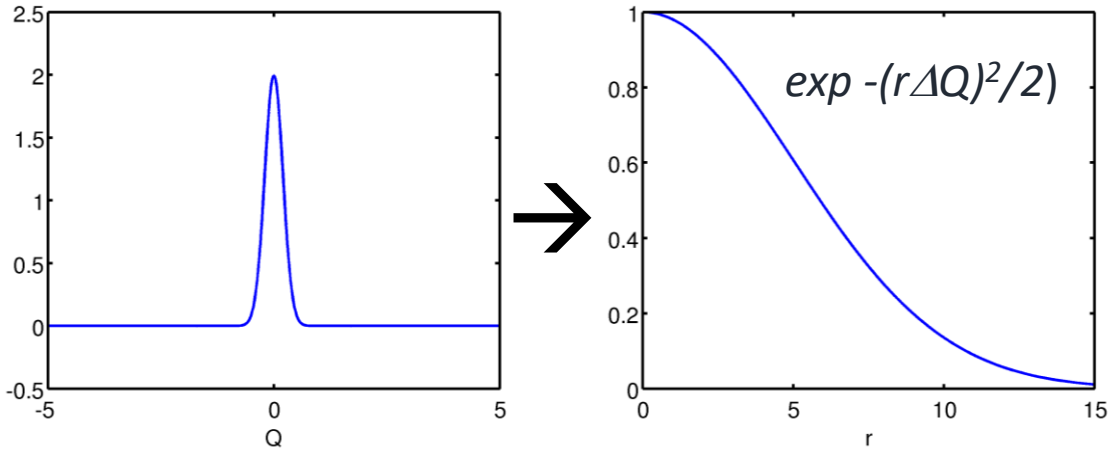


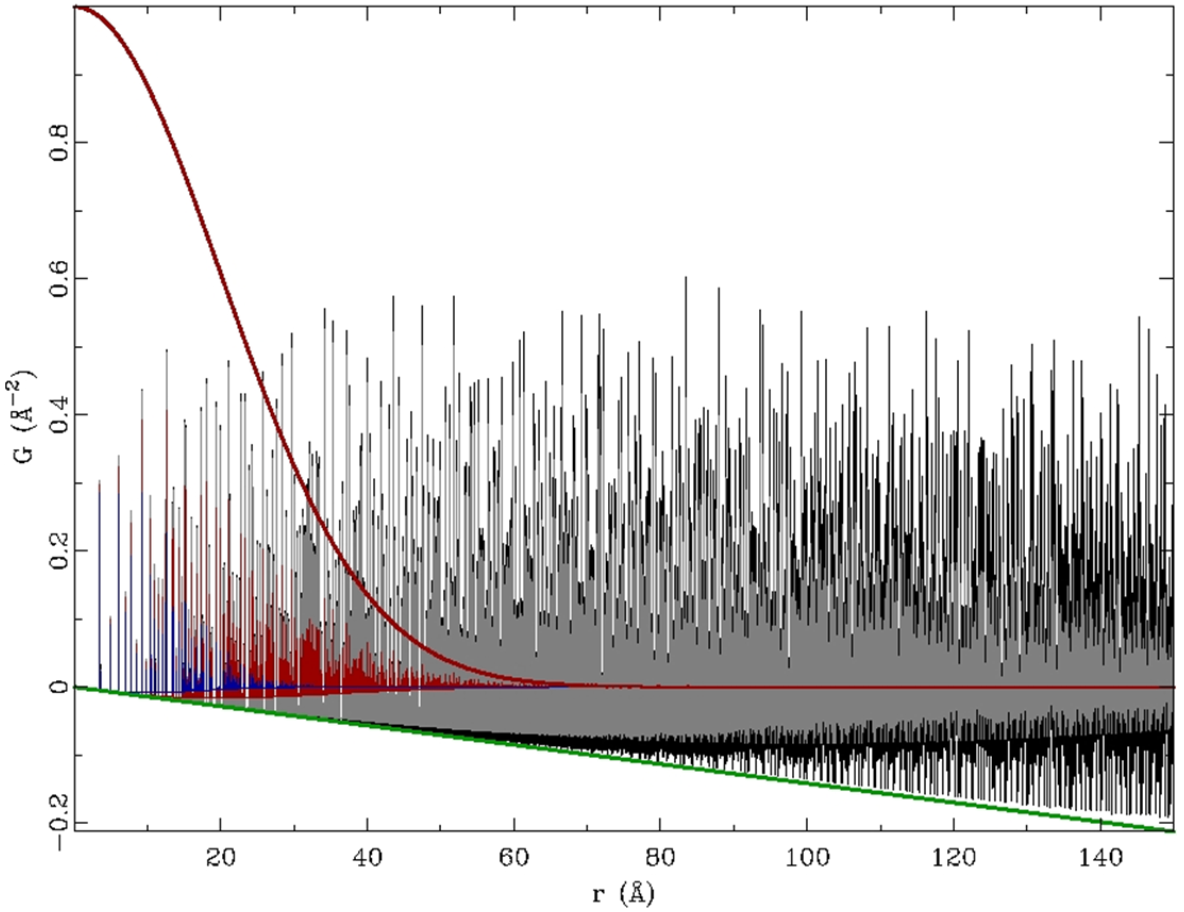
- ideal F(Q) is multiplied by a step function
- G(r) gets convoluted with a sinc function sinc(*r*) = sin(Q_{max} *r*) / *r* → *r*-resolution ≈ π/Q_{max}
- good *r*-resolution of G requires large Q_{max}
  Q = 4π sin θ/λ → best results with TOF neutrons or high-energy X-rays

PDFgui incorporates

Effect of the Q-space resolution

- Ideal F(Q) is convoluted by Gaussian to simulate finite Q resolution
- *G(r)* gets multiplied by real-space Gaussian with reciprocal width

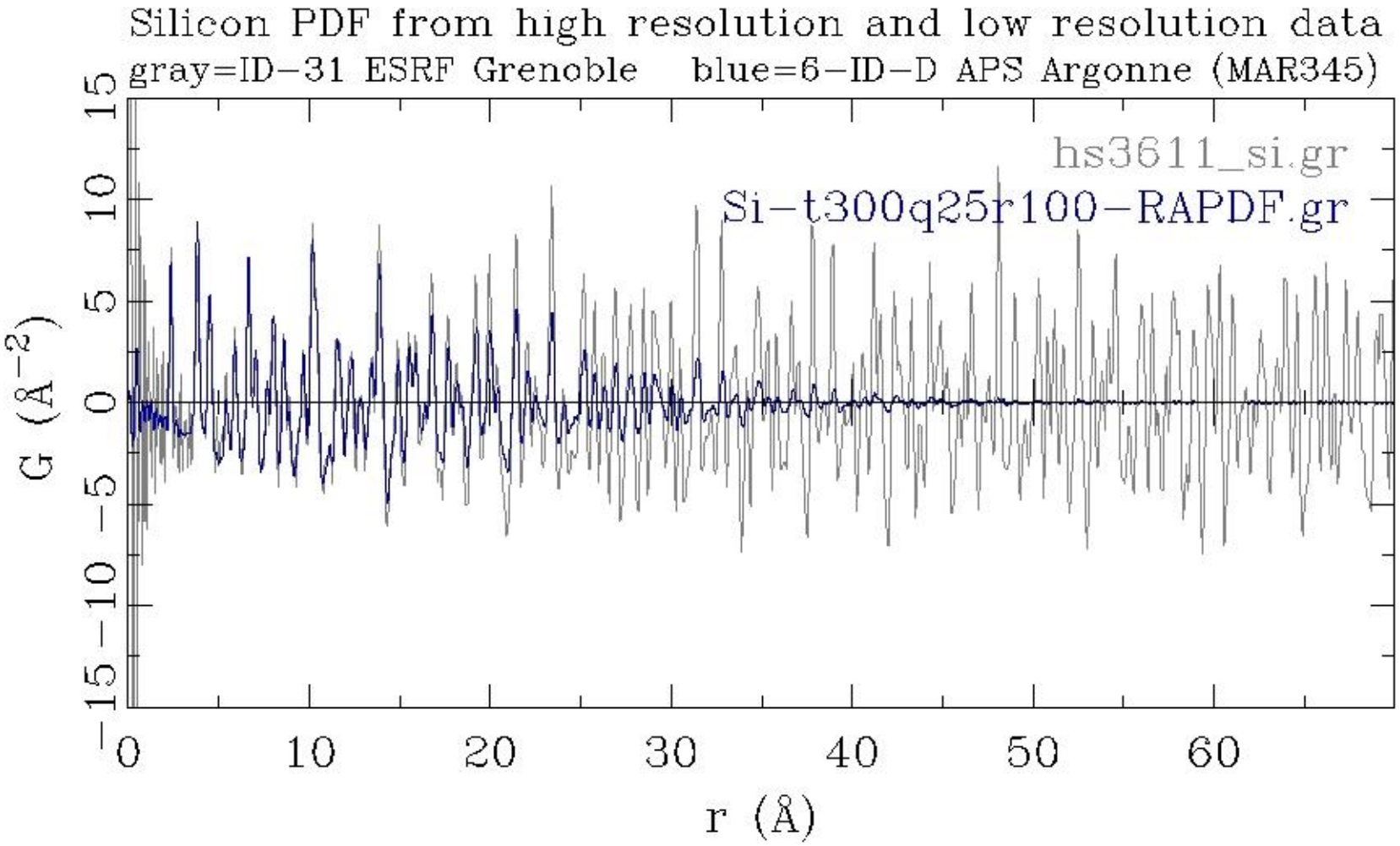




For *G(r)* to have good *r*-range high resolution in *Q* is required

*Q-resolution defines PDF "**field of view**"*

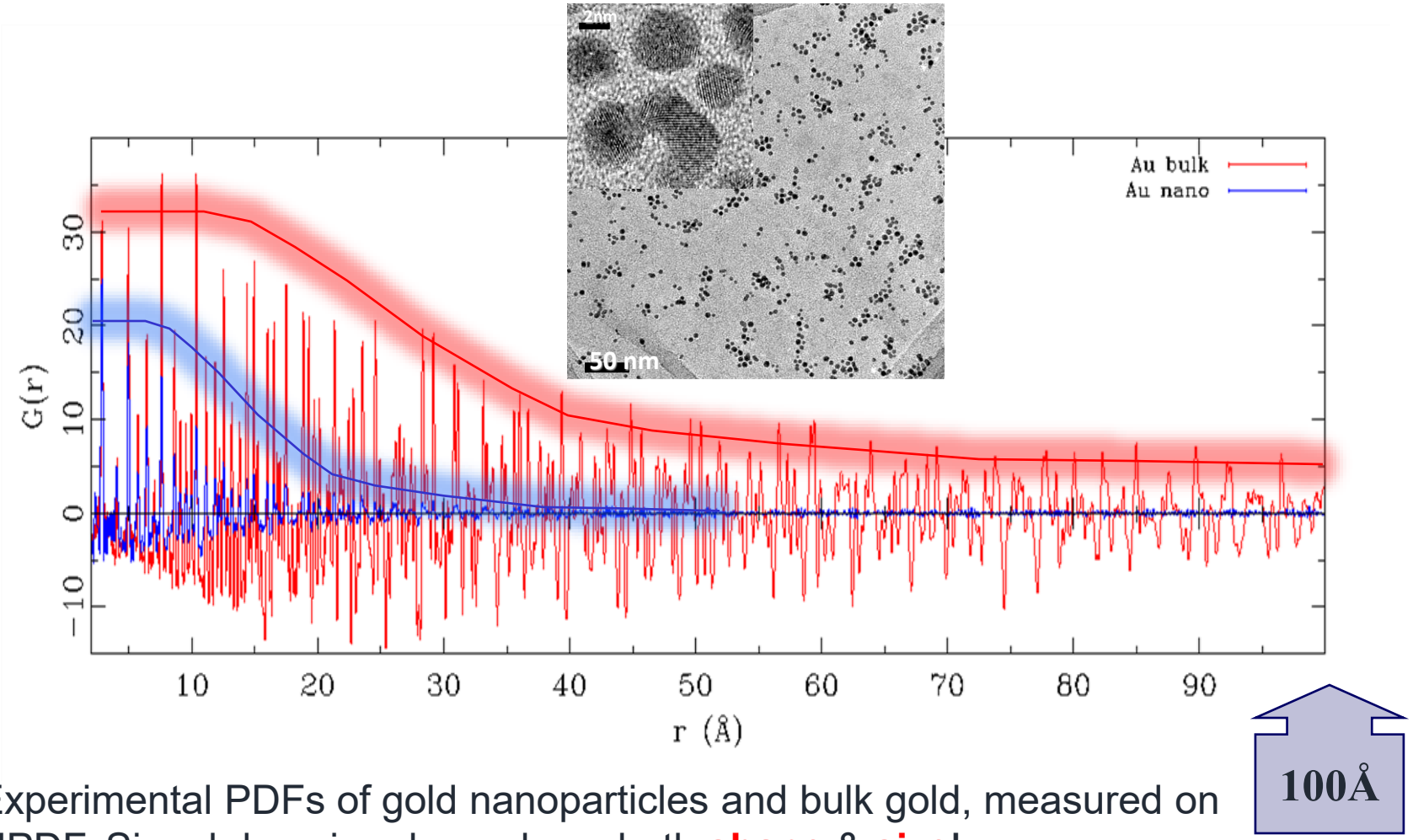
Effect of the Q-space resolution



PDFgui accounts for the effect

Effects of the finite particle size – nano vs bulk

K.L. Page *et al.*, *Chem. Phys. Lett.* **393**, 385 (2004).



Experimental PDFs of gold nanoparticles and bulk gold, measured on NPDF. Signal damping depends on both **shape** & **size**!

PDFgui accounts for spherical shapes only

Things needed for small box modeling …

- PDF **data** (*sample*.**gr** files) and associated information such as Q_{max} used, range of data, type of radiation, sample chemistry, ...

- In **small box modelling** approach, one typically starts from **a refinement** of a known/suspect structure, (thus reducing the volume of the parameter space as much as possible)
    - High-*r* region ~average structure
    - Low-*r* region ~local structure

  (biased view with bulk materials in mind)

- **Starting structure information**
    - space group and lattice parameters
    - fractional coordinates (asymmetric unit cell) & occupancies
    - having site-multiplicities handy may be helpful for crosschecking (e.g. **PDFgui works with symmetrized cells**)
    - Having an origin choice handy, if multiple are available, could matter

PDF modeling

- PDF is simulated from a known structure model

$$G_{calc}(r) = \frac{1}{Nr\langle b\rangle^2}\sum_{i\neq j} b_i b_j \frac{1}{\sqrt{2\pi}\sigma_{ij}} \exp\left[-\frac{(r-r_{ij})^2}{2\sigma_{ij}^2}\right] - 4\pi r\rho_0$$

- structure model is parameterized by a set of parameters p_i
- residuum R_w - difference between observed and simulated PDF

$$R_w(p_1, p_2, \ldots) = \sqrt{\frac{\sum_n \left[G_{obs}(r_n) - G_{calc}(r_n)\right]^2}{\sum_n G_{obs}^2(r_n)}}$$

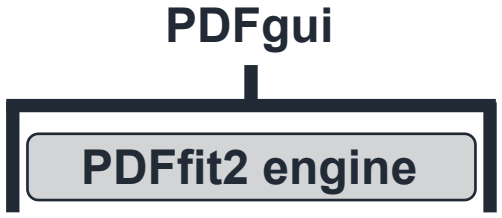
- least-squares refinement of p_i to minimize R_w
- Effects from setup (e.g. finite Q-resolution) or sample (correlated NN-motion) accounted for

$$B(r) = e^{-\frac{(rQ_{damp})^2}{2}} \qquad \sigma_{ij} = \sigma'_{ij}\sqrt{1 - \frac{\delta_1}{r_{ij}} - \frac{\delta_2}{r_{ij}^2} + Q_{broad}^2 r_{ij}^2}$$

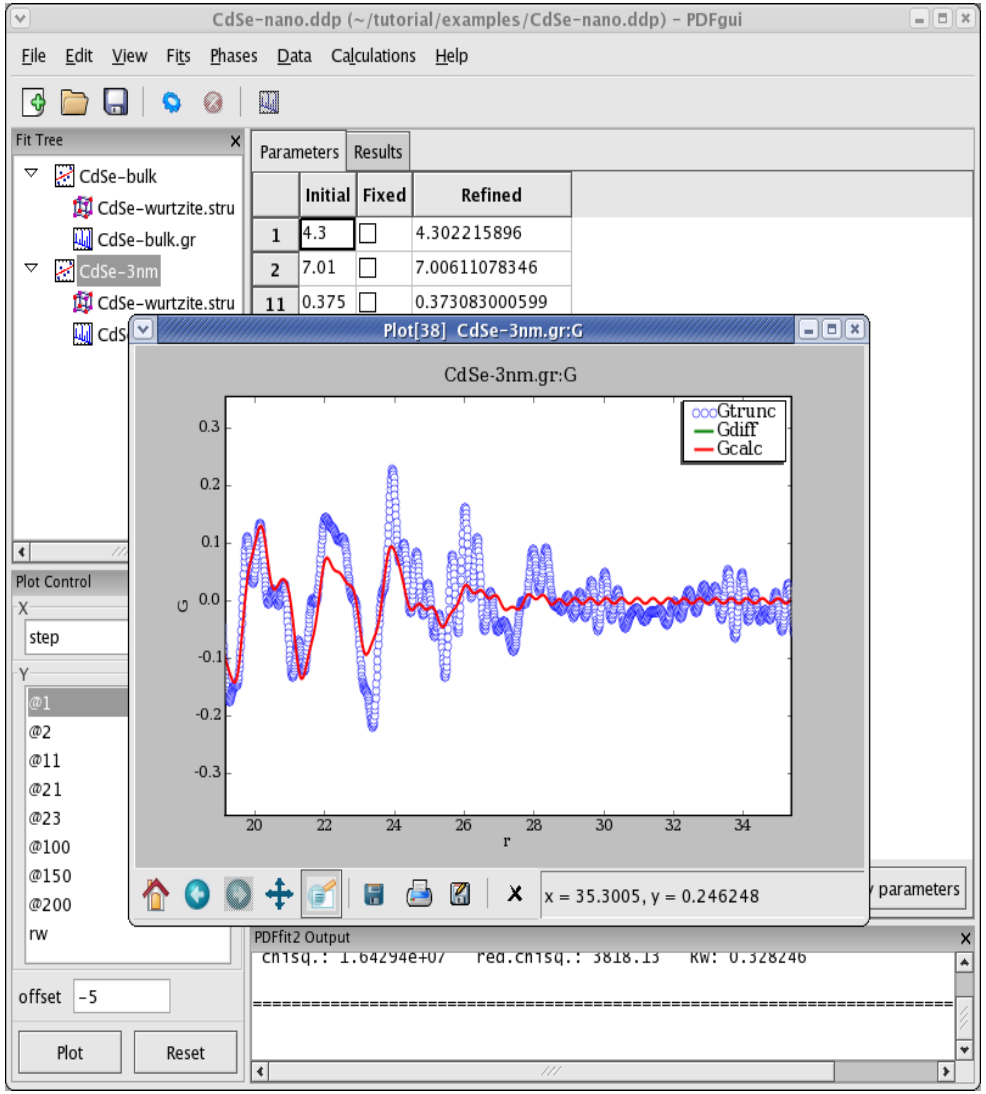
ADP

PDFgui overview

- PDFgui is a graphical interface (modeling platform) built on the PDFfit2 engine
- The PDFfit2 engine written in C++, Python accessible, can be prompt operated
- PDFgui organizes fits and simplifies many data analysis tasks, such as configuring, exectuing, and plotting multiple fits, adding functionality to script driven PDFfit2
- PDFfit2 calculates & fits a theoretical 3D-structure to PDF data (does all the work!)
- The fit system accounts for lattice constants, atomic positions and anisotropic ADPs, correlated atomic motion, as well as standard experimental factors affecting the data
- The atomic positions amd thermal coefficients can be constrained to follow symmetry requirements of an arbitrary space group, enabling studies of local broken symmetry



PDFgui overview



PDFgui

- GUI interface to PDFfit2 is user friendly modelling environment that **can be used for quick simulations (useful for experiment planning and sensitivity tests)**
- can **organize** multiple related fits in a single **project file** (.ddp file) easily shareable with colleagues, with **journal facility**
- powerful **visualization facilities**
  - live plotting of refined PDF profiles
  - parametric plots of variables from multiple fits
  - 3D structure visualization (optional connection to structure plotting)
- **structure model manipulation**
  - supports xyz, PDF, CIF and PDFfit formats
  - supercell expansion
  - expansion of asymmetric unit
  - generation of symmetry constraints for coordinates and atomic displacement factors, ADPs (“thermals”)
- **wizards** for T-series, doping-series, r-series (smart extraction of meta-data from files)

Easy set up for “on the fly” refinements of incoming data helps making experimental decisions

PDFgui parameters and program structure

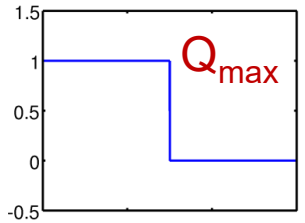
- **PDFgui parameters associated with DATASET**

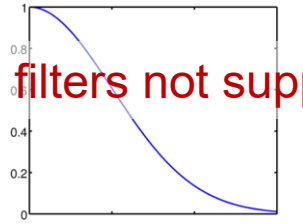
Fit range (r_{MIN}, r_{MAX})
fixed in refinement

user selected refinement *r*-range

Q_{max}
fixed in refinement

upper limit of integration used in Fourier transform
defines r-space resolution, predetermined





Lorch filters not supported

Q_{damp}
refined for calibrant
fixed for sample

Gaussian dampening (due to limited Q-resolution)

Q_{broad}
refined for calibrant
fixed for sample

High-*r* peak broadening (due to increased refined intensity noise at high Q and other sources, only significant when r_{MAX} is large)

dscale
refined

scale factor associated with dataset

PDFgui parameters and program structure

- **PDFgui parameters associated with PHASE**

pscale refined	phase scale factor NOTE: could be redundant/correlated with dscale
a, b, c, α, β, γ refined	lattice parameters
x[n] y[n] z[n] occ[n] u[*1*..*6*,n] refined (per symmetry)	x-position (fractional coordinates) y-position z-position site occupancy anisotropic displacement parameters U_{ij} [$Å^{-2}$]

NOTE: Refinement parameters can be correlated, particularly when a model is refined over a narrow r-range of data. PDFgui reports on correlations > |0.8|

PDFgui parameters and program structure

- **PDFgui parameters associated with PHASE for correlated atomic motion**

delta1
refined

1/r contribution to peak sharpening [Å^{-1}]

delta2
refined

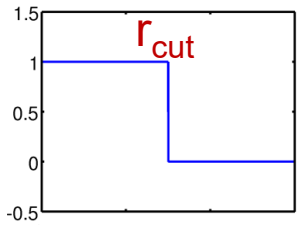
1/r^2 contribution to the peak sharpening [Å^{-2}]

sratio
refined

applied | not applied

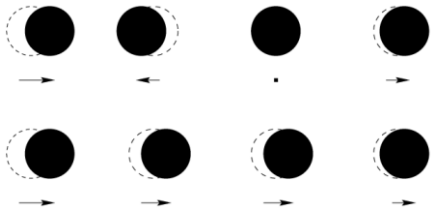
peak width reduction for correlated motion
(special cases of rigid structural units)

rcut
constat



radius cuttoff for applying the sratio
sharpening factor [Å]

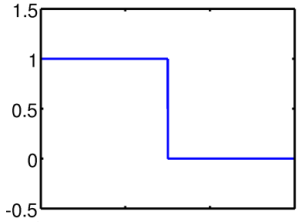
Note: Empirical correlated motion parameters are selected depending on material, they are very strongly correlated and affect other parameters



$$\sigma_{ij} = \sigma'_{ij}\sqrt{1 - \frac{\delta_1}{r_{ij}} - \frac{\delta_2}{r_{ij}^2} + Q_{broad}^2 r_{ij}^2}$$

PDFgui parameters for nanoparticles

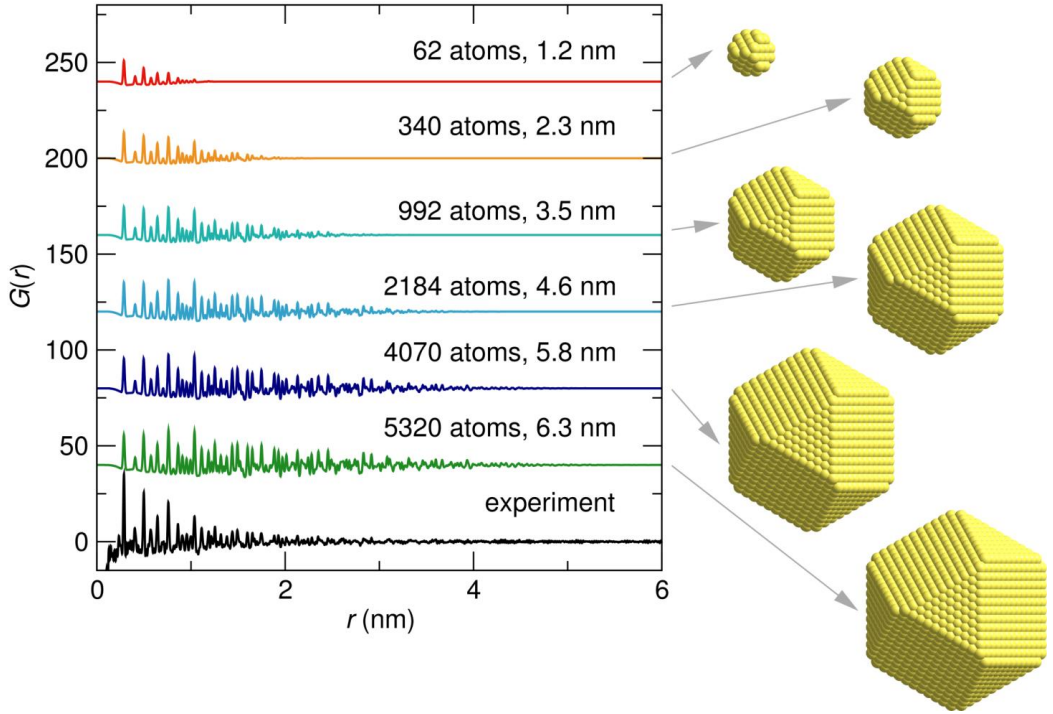
spdiameter
refined



spherical particle diameter for PDF shape damping function [Å]

stepcut
constant

r value above which the PDF is truncated to zero [Å]



PDFgui parameters and program structure

- **PDFgui declarations associated with PHASE**

X declaration	atom type associated with given site (all sites) e.g. Ni/Ta/Ca (label used to read scattering info from lookup tables of b_{coh} and Z).

- **PDFgui declarations associated with DATASET**

Neutron/X-ray declaration	scatterer type (used to determine lookup table)

NOTE: In rare instances one may experience the following
In case of X-ray radiation Z_X is used for element X. If ions present one can change X from original element to a fellow element with adequate electron count.

In case of neutron radiation b_{cohX} is used for element X. Lookup table contains information per natural isotope abundance. If isotope substitution is present, lookup table has to be modified with adequate *b* specified for a dummy element with made-up alphabetical code that will then be declared in the phase using that alphabetical code.

PDFgui parameters and program structure

Parameters are assigned by the handle syntax **@pn**,where **pn** is the parameter number

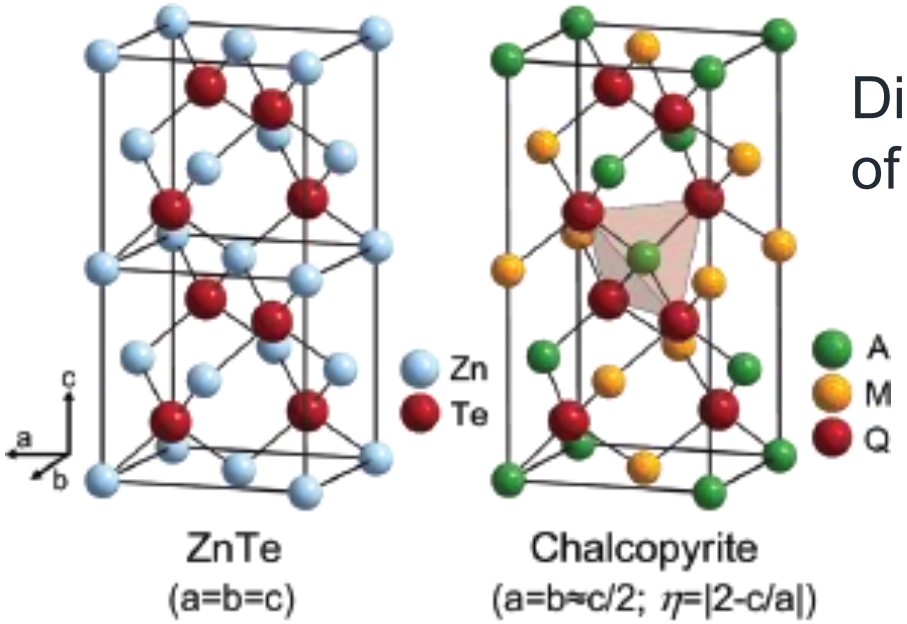
For example, @1, @55, @321, etc, note that numbers do not have to be consecutive

Variables that are assigned the same handle will be described by the same parameter

Caution should be exercised to avoid unintentional assignment of the same parameter number to incompatible variables (variables of different type)

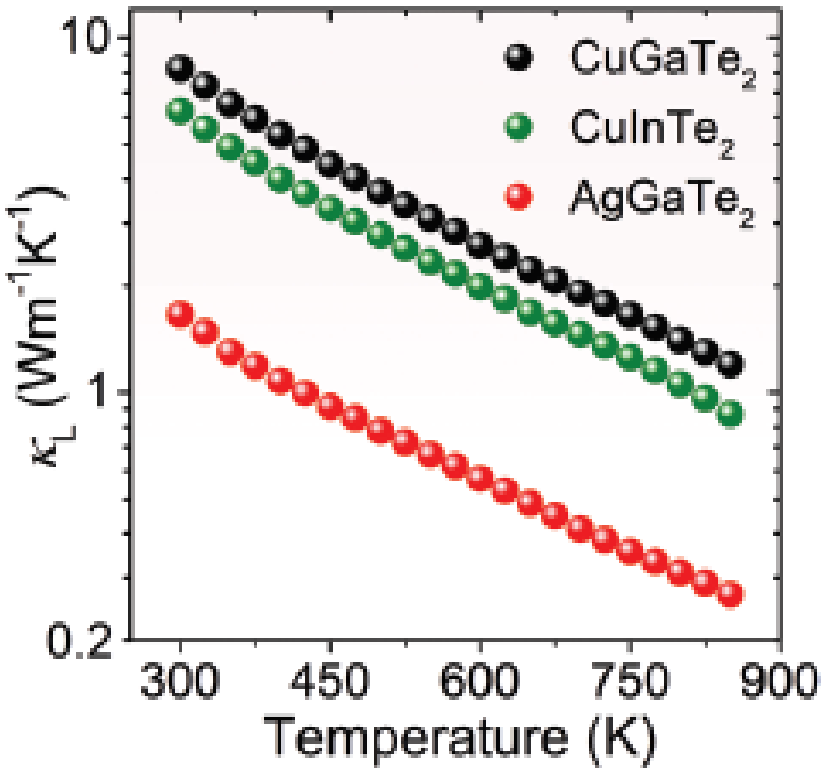
…more details at the tutorial ☺

Illustrative example: emerging state in $AgGaTe_2$

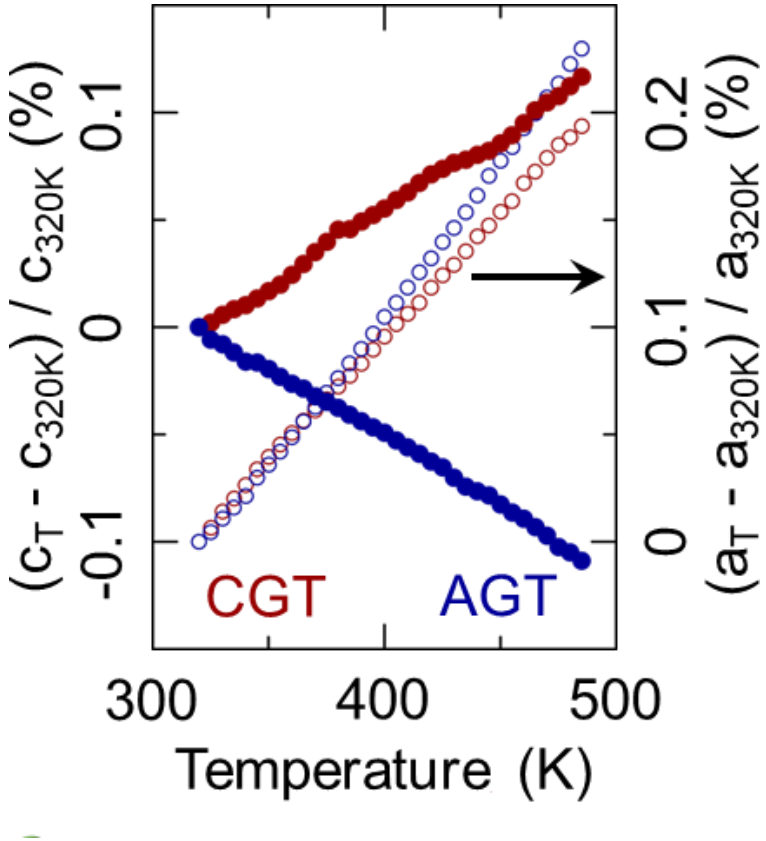


Diamondoid structures (zincblende doubled along c-axis) of intense interest in high-performance thermoelectricity

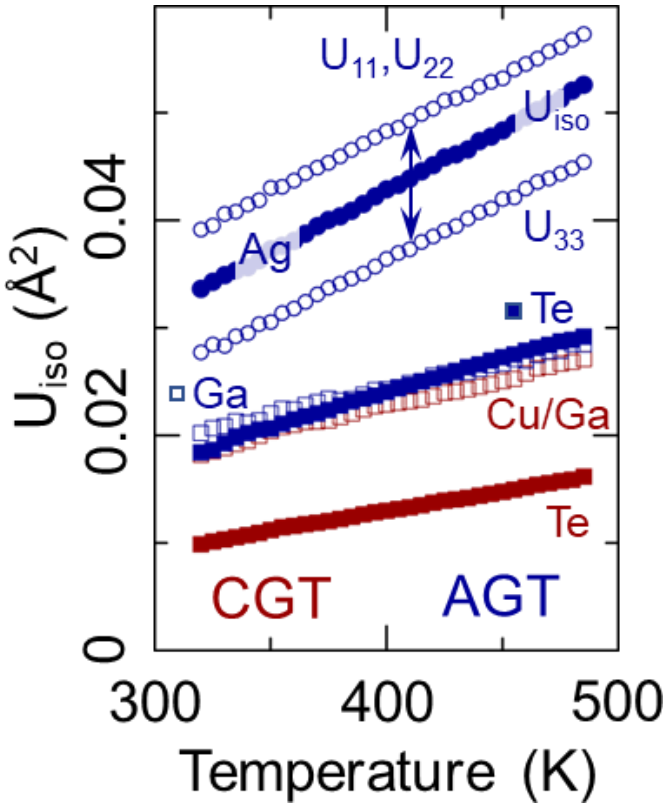
Why is lattice thermal conductivity at high temperature Ultralow, and smaller in Ag- than in the Cu-variants?



Illustrative example: emerging state in $AgGaTe_2$

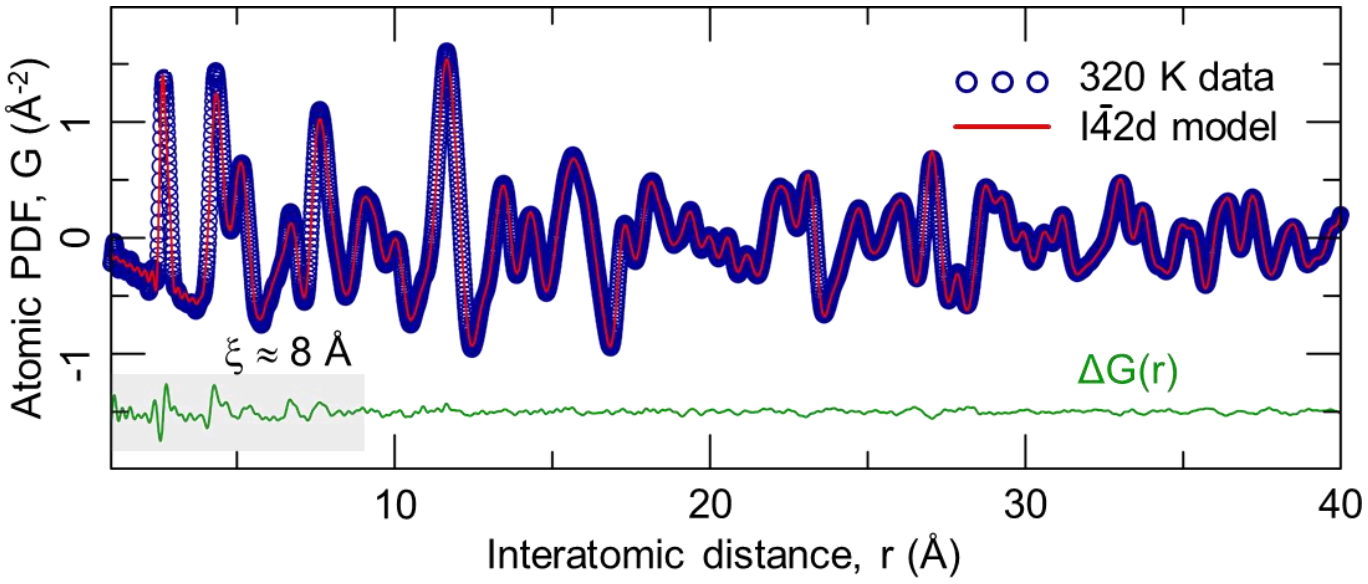


Rietveld finds that in average structure c-axis of **Ag-variant** has negative thermal expansion (**NTE**) while **Cu-variant** behaves normally. Why?



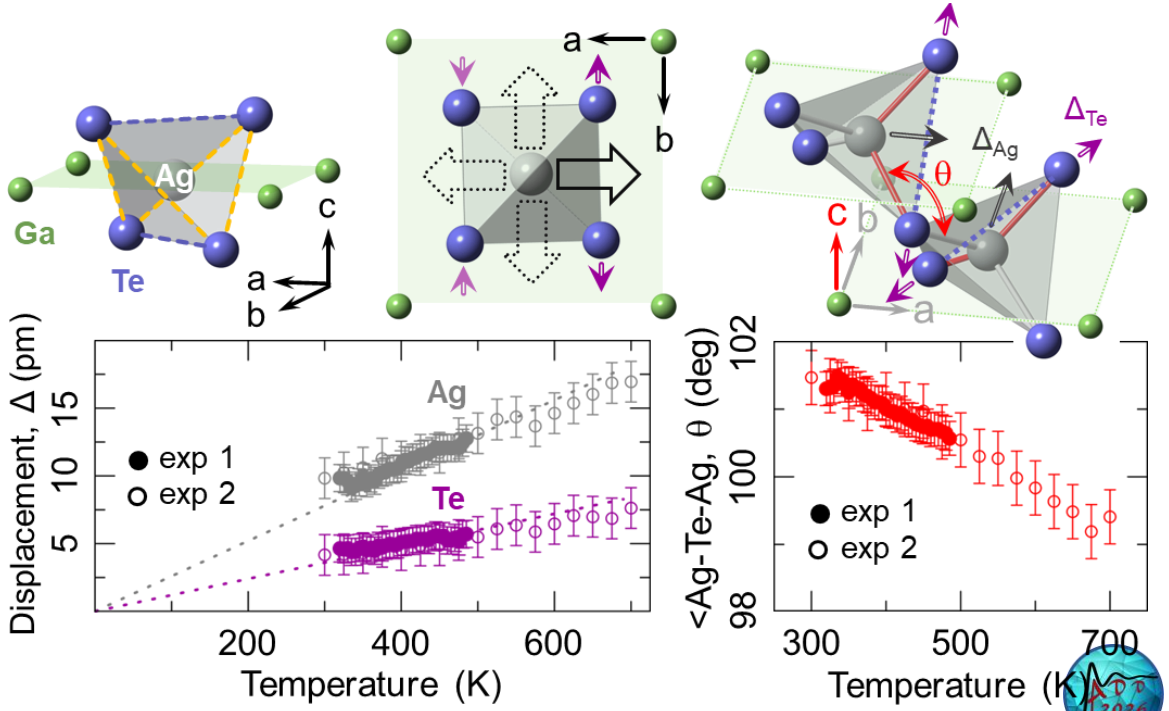
Atomic ADPs of **Ag** are abnormally enlarged in the ab-plane (perp. to c-axis) indicating in-plane nanoscale disorder!

Illustrative example: emerging state in $AgGaTe_2$



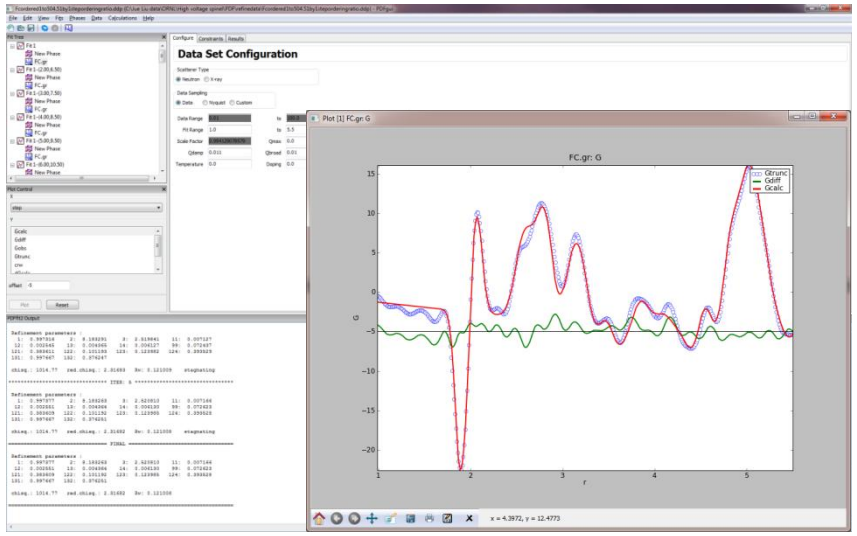
PDFgui modeling evidences short-range distortion of $AgTe_4$ tetrahedra, correlated over nanometer length-scale!

- discovery via elaborate PDFgui modeling
- distortion involves local $AgTe_4$ rotations leading to global NTE effect
- continuous emergence on warming
- NTE a macroscopic measure of nanoscale distortion and its evolution on warming



Hongyao Xie et al, Adv. Mater. ***34***, 2202255 (2022)

“Small Box” software comparison

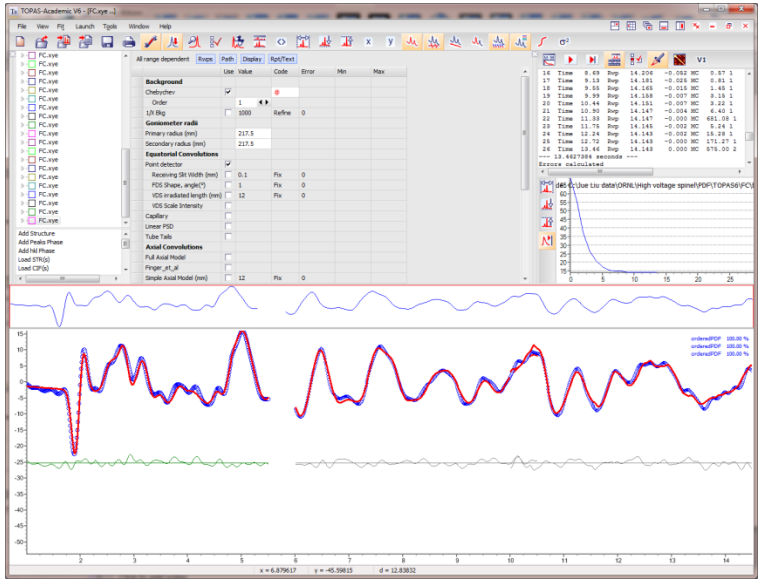


PDFgui http://www.diffpy.org/

- Slow refinement, e.g. for high-*r*
- Cumbersome outputs

+ Open Source and Free

+ GUI is Simple and User-friendly



TOPAS PDF

- Commercial ($$$$)
- Steeper learning curve
- Have to write your own macro

+ Very Fast

+ Easy manipulation of fitting parameters and plotting

+ Can easily customize output functions

http://www.topas-academic.net

slide courtesy Katharine Page