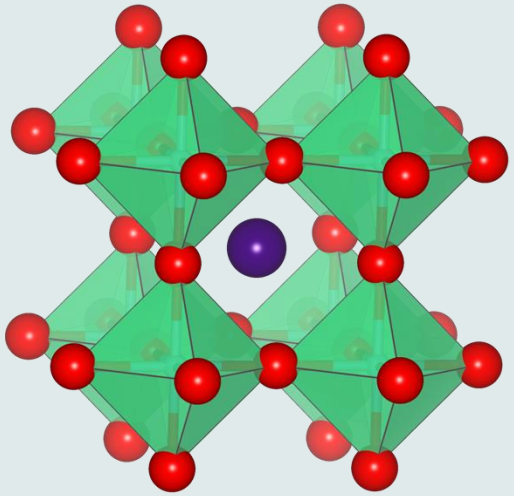
https://github.com/rsc-solid-data/data_diffn_crystals

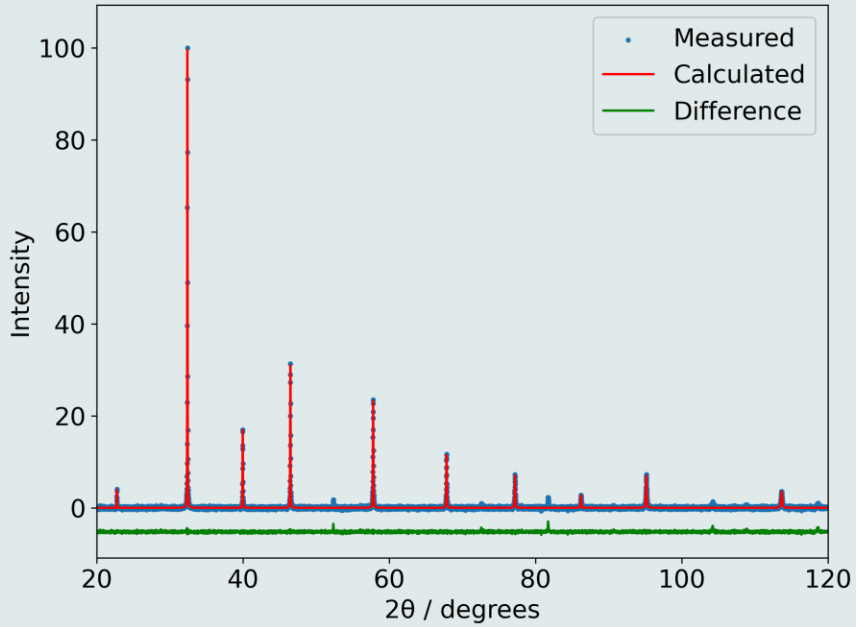


Data and Models of Solid State Structures

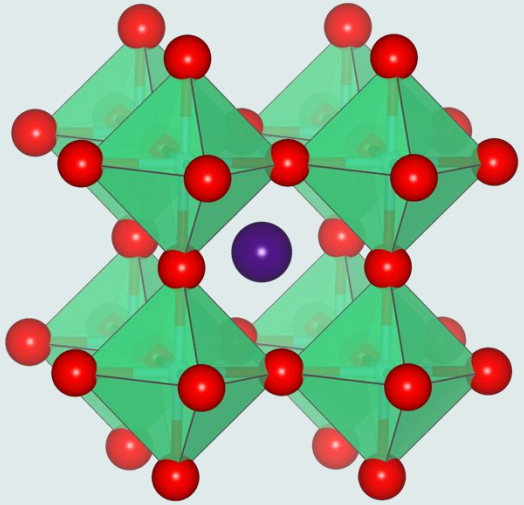
RSC INTEREST GROUP
SOLID STATE CHEMISTRY

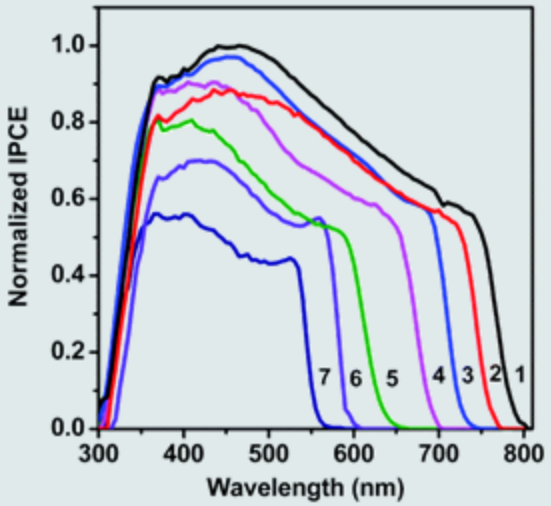
LEVERHULME RESEARCH CENTRE
FUNCTIONAL MATERIALS DESIGN

CCDC



Structure: Foundation of materials “design”





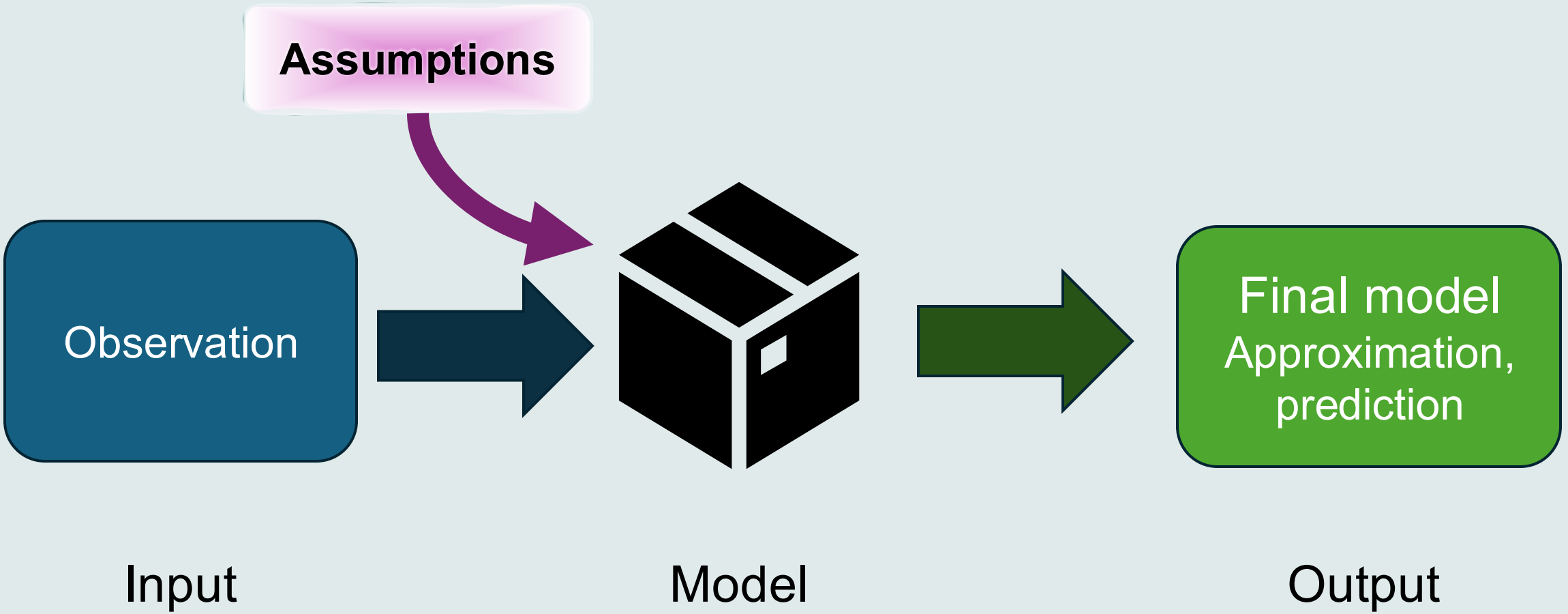
Structure **Properties** **Function** and **application**

- Solid state structure underpins function in materials, from conductivity to colour.
- Representations of solid-state structure are **models**, not direct observations.
- Accurate, truly representative structural models are vital

  → Valid **structure-property relationships**

S. A. Kulkarni *et al.*, *J. Mater. Chem. A*, 2014, **2**, 9221.

Models are everywhere



Models are everywhere

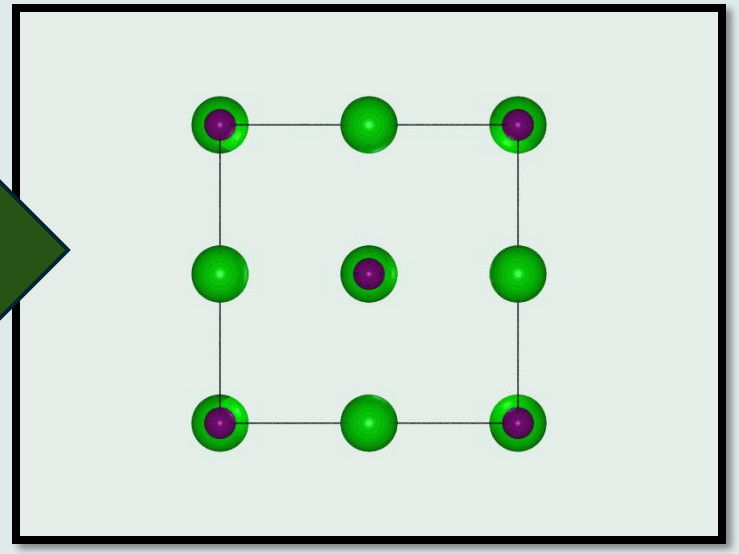
Assumptions

Ionic model
Hard spheres, known radii, electrostatic interactions

Crystalline solid
Regular repeat defined by periodic lattice

Crystal structure
Approximation of true atomistic arrangement (another model)

Real crystal



Input

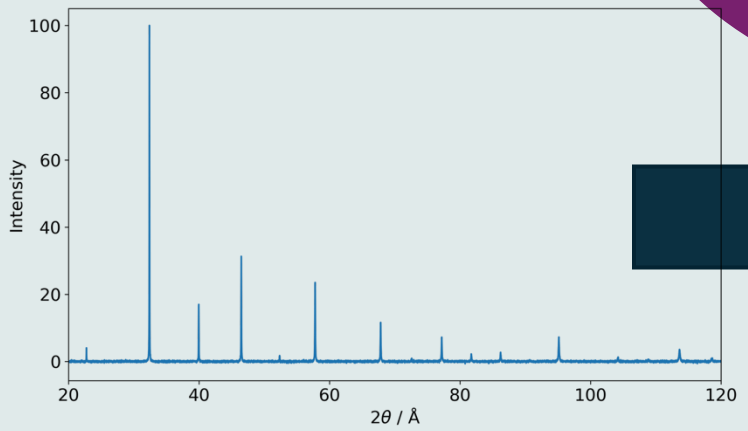
Model

Output

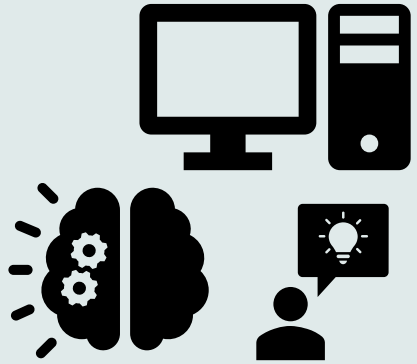
Models are everywhere

Assumptions

e.g.
Data reliability
(sample, measurement, processing…)
Long-range, perfectly ordered structure
Time-averaged "static" structure
represents dynamics adequately



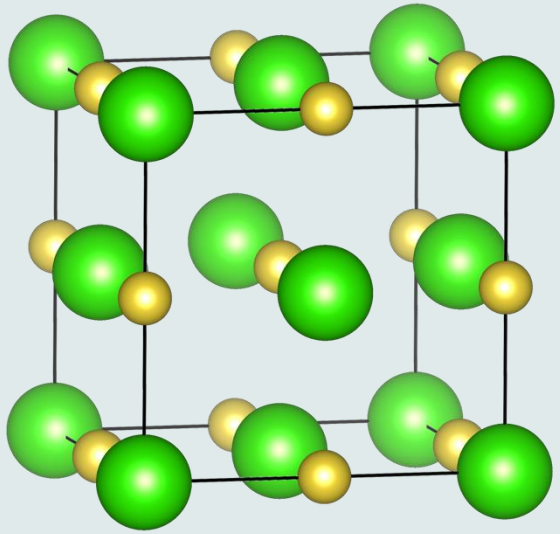
Input



Model

Bragg's law
Intensities
Structure factors
Thermal effects…

Crystal structure
(another model)



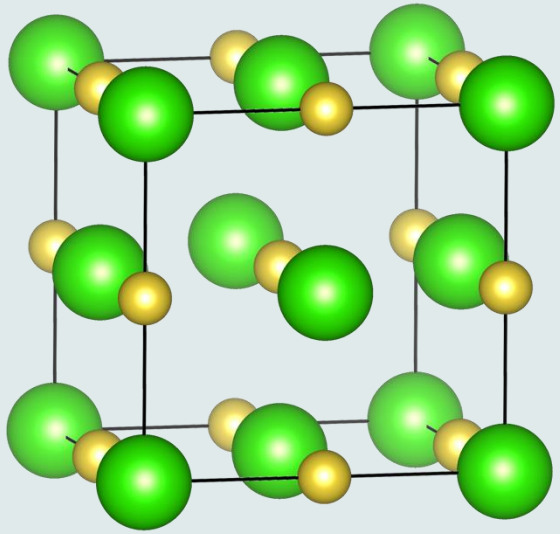
Output

Models are everywhere

Assumptions

e.g.
Model reliability
(physically realistic, precision on refined parameters)
Long-range, perfectly ordered structure
Time-averaged "static" structure represents dynamics adequately

Crystal structure
(another model)



Measurable data
- XRD
- IR
- Raman

Properties
- Optical
- Porosity
- Guest affinity
- Conduction

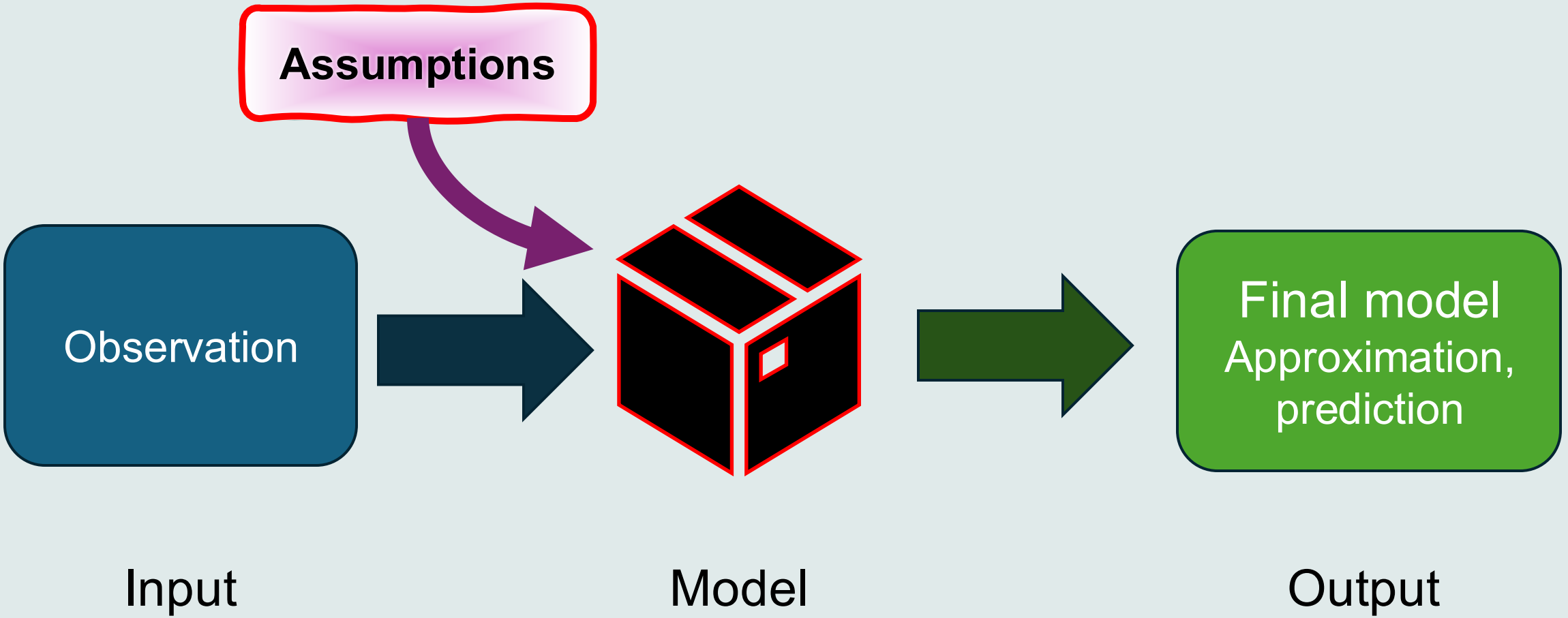
Output

Model

Bragg's law
Intensities
Structure factors
Thermal effects…

Input

Models are everywhere



“All models are wrong, but some are useful”

– George E. P. Box

2.3 Parsimony

Since all models are wrong the scientist cannot obtain a “correct” one by excessive elaboration

In applying mathematics to subjects such as physics or statistics we make tentative assumptions about the real world which we know are false but which we believe may be useful nonetheless.

2.4 Worrying Selectively

Since all models are wrong the scientist must be alert to what is importantly wrong. It is inappropriate to be concerned about mice when there are tigers abroad.

https://en.wikipedia.org/wiki/All_models_are_wrong

G.E.P. Box, *Journal of the American Statistical Association*, Vol. 71, No. 356. (1976), pp. 791-799.

Crystallography

Modelling crystals

XRD data to structural model

Assumptions

Ionic model
Hard spheres, known radii, electrostatic interactions

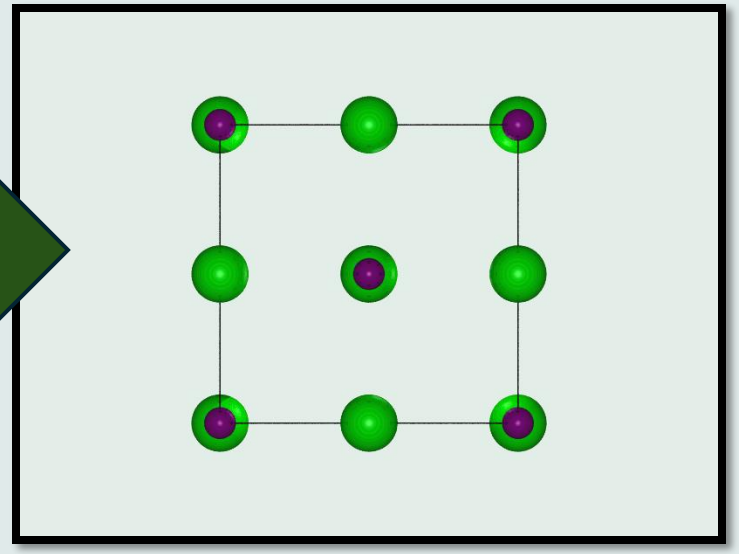
Crystalline solid
Regular repeat defined by periodic lattice

Real crystal

Input

Model

Approximation of true atomistic arrangement (another model)

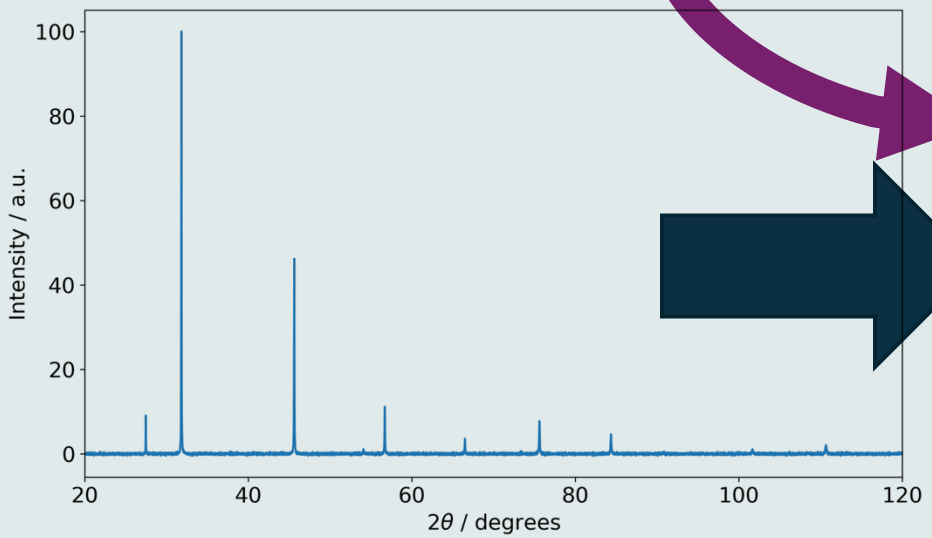


Output

XRD data to structural model

Assumptions

e.g.
Data reliability
(sample, measurement, processing…)
Long-range, perfectly ordered structure
Time-averaged "static" structure
represents dynamics adequately

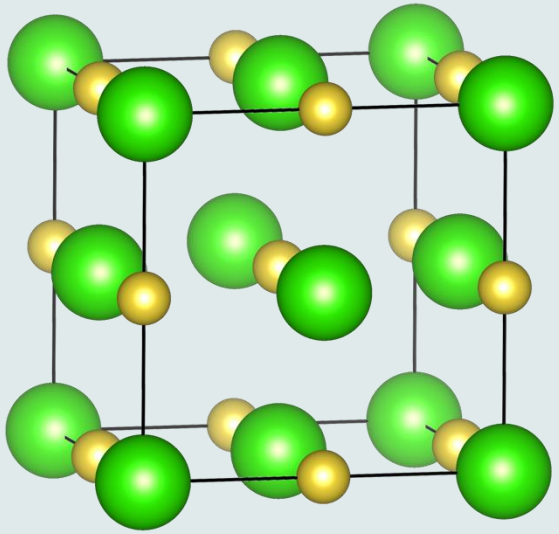


Input

Model

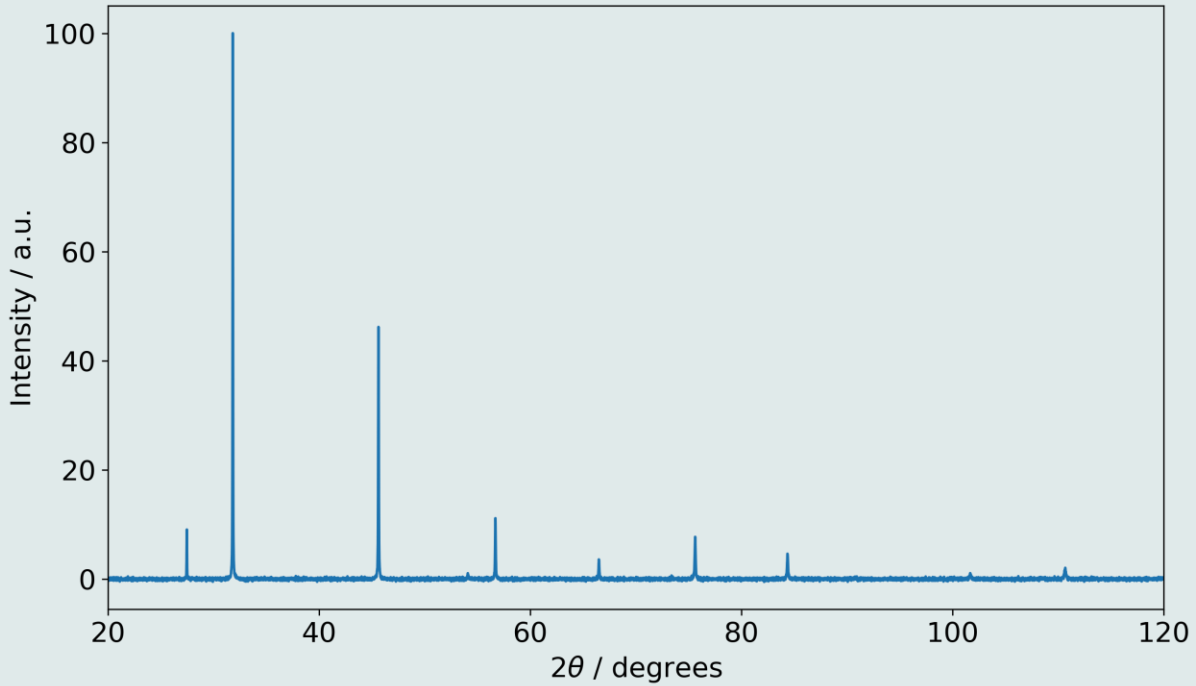
Bragg's law
Intensities
Structure factors
Thermal effects…

Crystal structure
(another model)



Output

XRD data to structural model

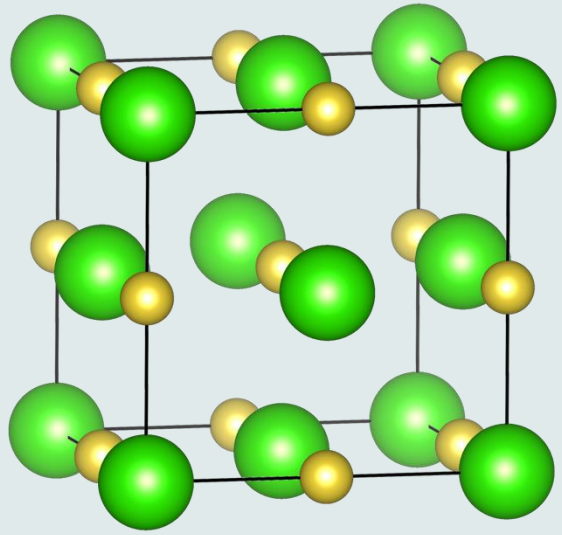


Positions **Areas**
Widths

Input → Model

Bragg's law
Intensities
Structure factors
Thermal effects…

Crystal structure
(another model)



Output

XRD data to structural model

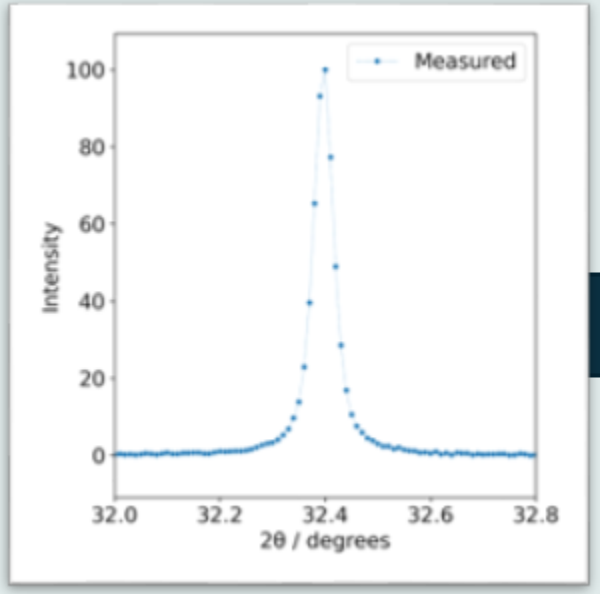
Assumptions

e.g.
Data reliability
(sample, measurement, processing…)
Background model
Diffraction peak shapes

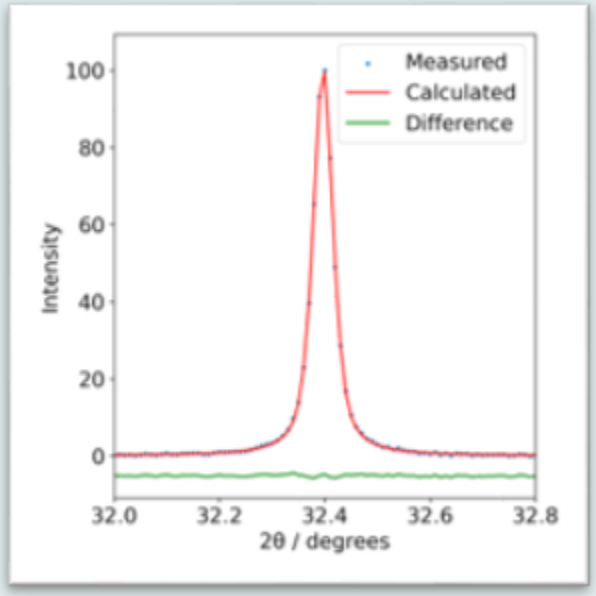
Positions
Areas
Widths

more models!

Interpretation



$$y = \eta L(x) + (1 - \eta)G(x)$$

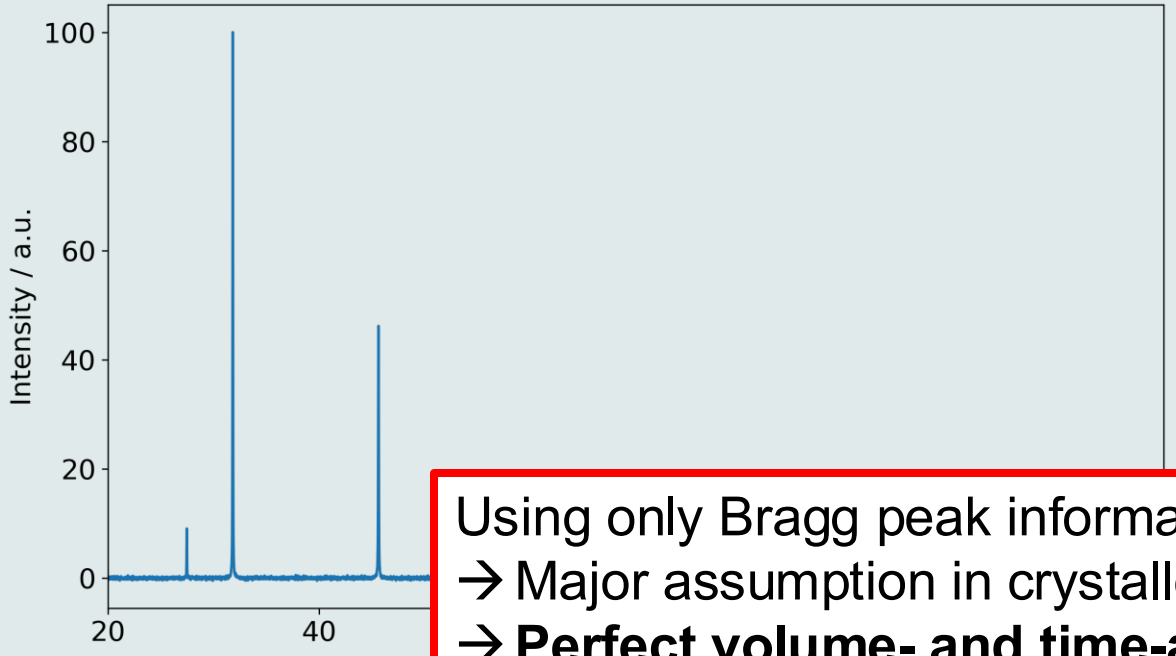


Input

Model

Output

XRD data to structural model



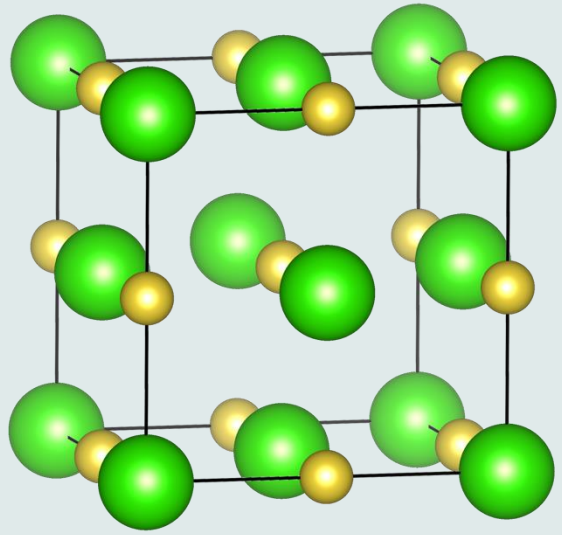
Using only Bragg peak information
→ Major assumption in crystallography
→ **Perfect volume- and time-averaged model**

Positions Areas
Widths

Input → Model

Bragg's law
Intensities
Structure factors
Thermal effects…

Crystal structure
(another model)



Output

Modelling the diffraction pattern

- Bragg's law

$$\lambda = 2d_{hkl} \sin \theta_{hkl}$$ reflection (*hkl*) positions

- Scattering intensities

$$F_{hkl} = \sum_j f_j \exp 2\pi i\left(hx_j + ky_j + lz_j\right)$$

f_j is the scattering factor for atom *j* at the position (x_j,y_j,z_j) in the unit cell.
For X-rays, it is related to the atomic number of *j*

$$I_{hkl} \propto |F_{hkl}|^2$$

Analysis of diffraction intensities

$$\rho_{xyz} = \frac{1}{V}\sum_{hkl} F_{O,hkl}\exp[-2\pi i(hx+ky+lz)]$$

$$= \frac{1}{V}\sum_{hkl} \left|F_{O,hkl}\right|\exp(i\Phi_{hkl})\exp[-2\pi i(hx+ky+lz)]$$

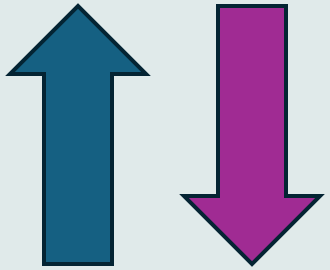
$$\ldots$$

$$= \frac{1}{V}\sum_{hkl} \left|F_{O,hkl}\right|\cos[2\pi(hx+ky+lz)-\Phi_{hkl}]$$

Reciprocal space: **diffraction**

$$F_{hkl} = \sum_j f_j \exp 2\pi i\left(hx_j + ky_j + lz_j\right)$$

Fourier transform
FT



Inverse
Fourier transform
$(FT)^{-1}$

$$\rho_{hkl} = \frac{1}{V}\sum_{hkl}\left|F_{O,hkl}\right| \cos[2\pi(hx + ky + lz) - \Phi_{hkl}]$$

Real space: **crystal structure**

Structure determination from PXRD

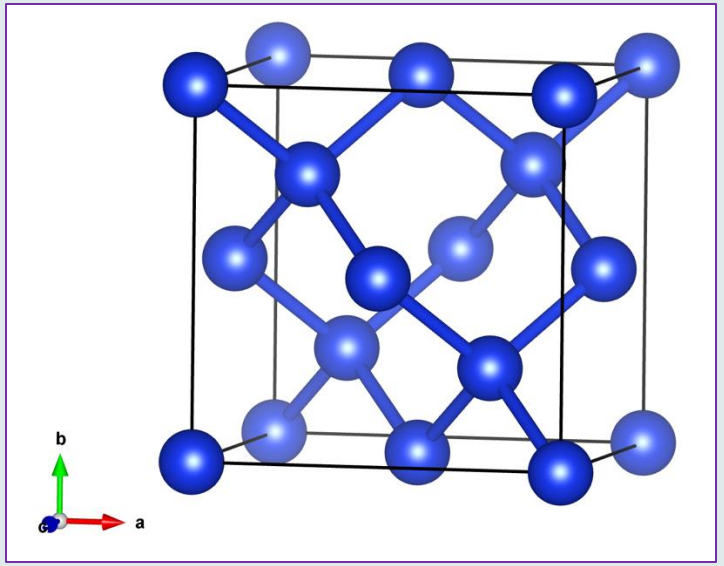
- Reliable intensities I_{hkl} from PXRD profiles can be challenging
  - Reflections often overlap and cannot be fully resolved
- If an approximate model of the crystal structure is available:
  - Calculate diffraction pattern for the model
  - Use least square refinement (minimisation) to vary the model parameters to produce a good match for the experimental data
- Optimisation of a structural model to fit the whole experimental powder profile:

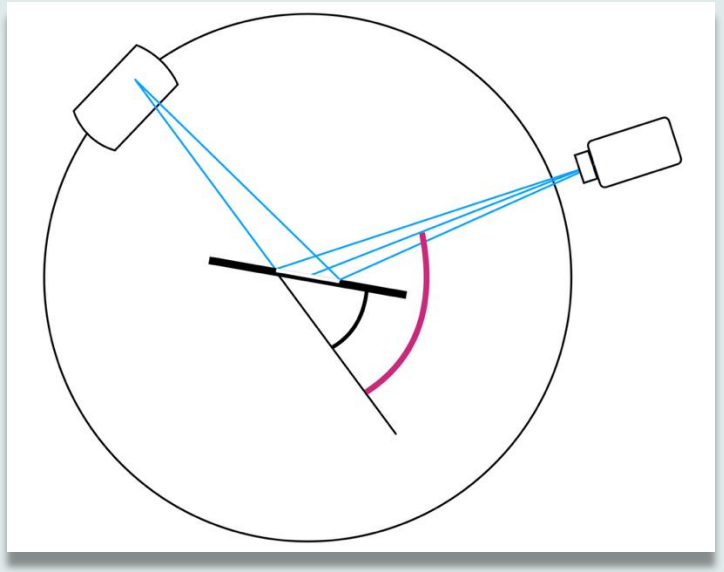
Rietveld refinement

Model refinement

Initial structural model

- Unit cell and symmetry (crystal system, lattice centring, space group)
  - a, b, c, α, β, γ ↔ Positions of reflections
  - P, I, F → Systematic absences
- Atomic contents and positions
  - For each atom: species (f_j)
    position x, y, z
    atomic displacement b
    → Intensities





Experimental model

- Diffractometer
- Data collection
  - ↔ Intensities
  - ↔ Peak positions

Sample model

- Crystal size
- Sample format
  - ↔ peak shapes

Least squares minimisation

$$S_y = \chi^2 = \sum_i w_i \left(y_{o,i} - y_{c,i}\right)^2$$

Usual weighting scheme for PXRD: $w_i = {}^1/_{\sigma_{o,i}^2} = {}^1/_{y_{o,i}}$

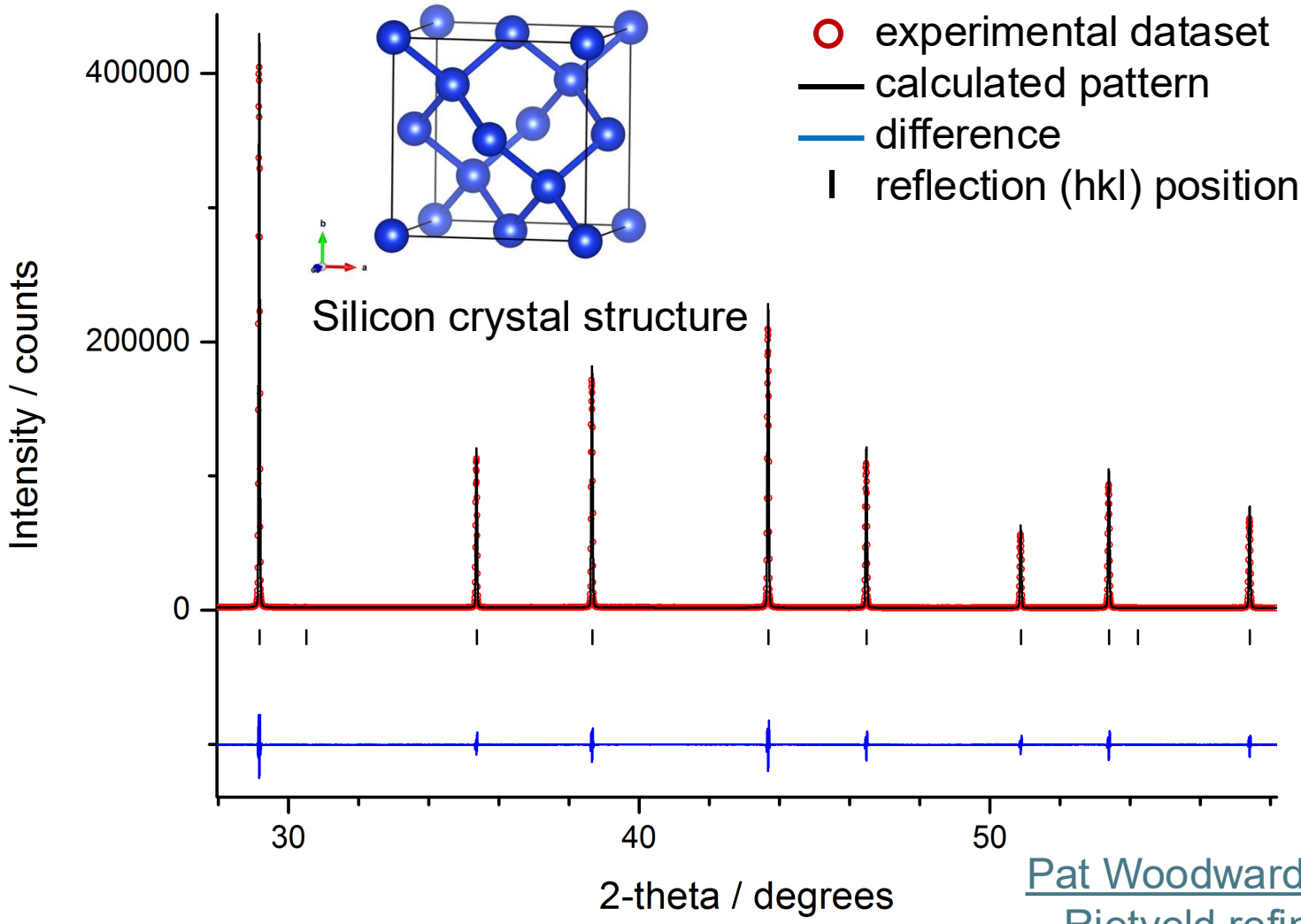
σ is standard uncertainty on the measured data point

where

$$y_{c,i} = s \sum_K L_K |F_K|^2 \varphi(2\theta_i - 2\theta_K) P_K A_i + y_{b,i}$$

S – goodness of fit
i – data points in profile
w – weight
y_o – observed intensity
y_c – calculated intensity
y_b – background intensity

F – structure factor
K – *hkl*
s – scale factor
L – Lorentz polarisation factor
$\varphi(2\theta_i - 2\theta_K)$ – peak shape function
P – preferred orientation function
A – absorption (function)



Pat Woodward's (OSU)
Rietveld refinement

Accessing structural data

- Crystal structures
  - BYO – determined from diffraction
  - ICSD
  - CSD
  - COD
  - PDB, minerals, incommensurate
  - Materials Project
  - PCOD
  - OQMD
- Diffraction data
  - ICDD (PDF-4/5+, etc.)
  - P2D2 (from PCOD)
  - BYO
    - collect data
    - from crystal structures

https://github.com/rsc-solid-data/data_diffn_crystals

Comparing XRD profiles

- Be aware of where the cifs are from
  - Experimental conditions
  - Comparability of molecular simulations

- Be aware of where the experimental data is from
  - Experimental conditions
  - Sample compositions (guest inclusion, compositional variation)
  - Sample effects (size, preferred orientation, absorption, phase purity)
  - Instrumental considerations (broadening, Lp, optics)
  - Structural effects (order/disorder, symmetry, phase transitions)

Comparing XRD profiles

- Try the unit cell and space group
  - Structureless refinement – no atoms, arbitrary intensities
  - Le Bail or Pawley
- Try a refinement
  - Rietveld refinement is local minimisation
  - Need a good starting model

Some extra resources

Fundamentals of Powder Diffraction and Structural Characterization of Materials, Springer US, Boston, MA, 2009.

B. H. Toby, *R* factors in Rietveld analysis: How good is good enough?, *Powder Diffr.*, 2006, **21**, 67–70.

L. B. McCusker, R. B. Von Dreele, D. E. Cox, D. Louër and P. Scardi, Rietveld refinement guidelines, *J Appl Cryst*, 1999, **32**, 36–50.

J. H. O'Donnell, R. B. Von Dreele, M. K. Y. Chan and B. H. Toby, A scripting interface for GSAS-II, *J Appl Cryst*, 2018, **51**, 1244–1250.