From Data to Structure

in Neutron Total Scattering at the CSNS Multi-Physics Instrument （CSNS-MPI)

Juping Xu, Wen Yin*, Huaican Chen, Yuanguang Xia

Institute of High Energy Physics, Chinese Academy of Sciences

China Spallation Neutron Source Center

ICANS XXV March 2026 Malmö

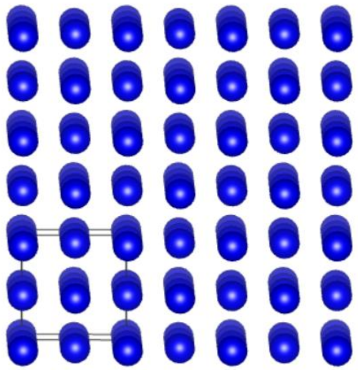
Outline

- Neutron Total Scattering and PDF
  Physical Foundations for Probing Multi-range Structure
- Data Processing and Modeling
  From Raw Scattering Data to Reliable Structural Models
- Intelligent Enhancement
  Advanced Algorithms and AI Empowering Structure Analysis
- Summary and Outlook

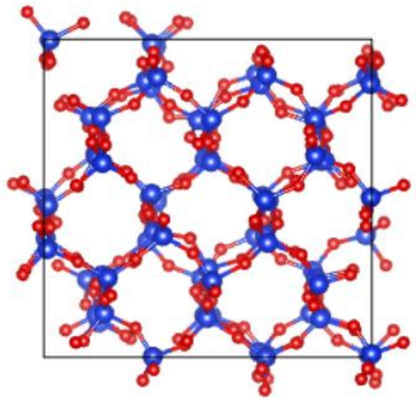
Neutron Total Scattering **Diffraction + Diffuse Scattering**

Multis-order structures

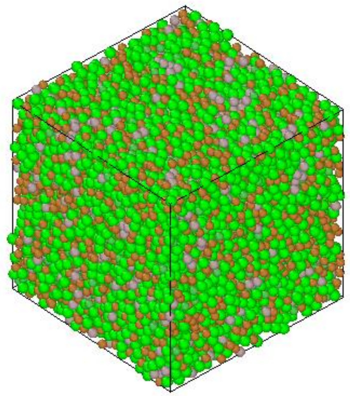
Long range ordered crystal



Local ordered/disordered structure

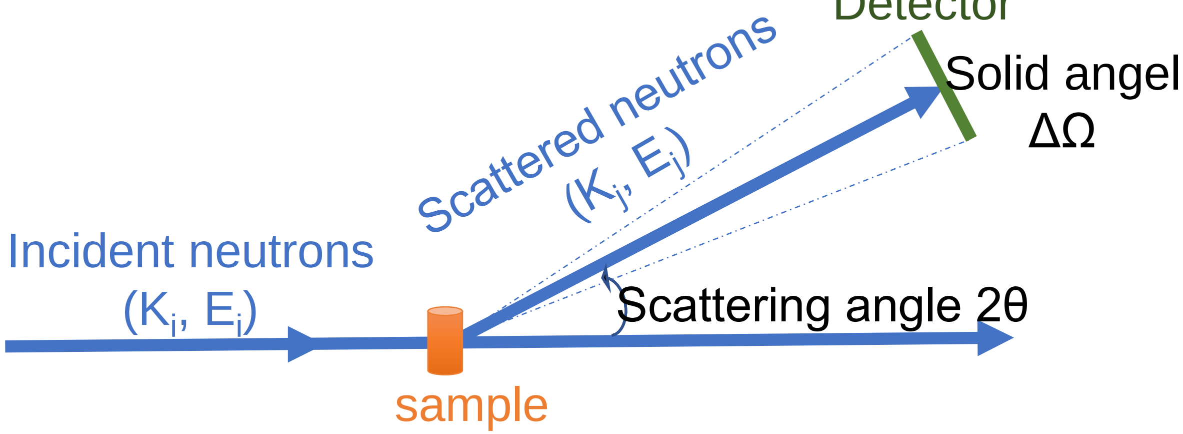


Disordered material



Can analysis matter structure in different order, from atomic scale to nanoscale.

Elastic interaction of neutrons with matter



Elastic scattering section:

$$\left(\frac{d\sigma}{d\Omega}\right)_{c,el} = \frac{(2\pi)^3}{v_0}\sum_{\boldsymbol{\tau}} |F(\boldsymbol{\tau})|^2 \delta(\mathbf{Q}-\boldsymbol{\tau}),$$

$$\left(\frac{d\sigma}{d\Omega}\right)_{c,el} = \frac{1}{N_c}\left|\sum_{njk}\left[c_{jk}+P_{njk}\right]\langle\overline{b_k}\rangle e^{i\mathbf{Q}\cdot(\mathbf{l}_n+\mathbf{d}_j)}\left[1+i\mathbf{Q}\cdot\sum_{n'j'k'}\Delta P_{n'j'k'}\mathbf{u}_{njk}^{n'j'k'}\right]\right|^2$$

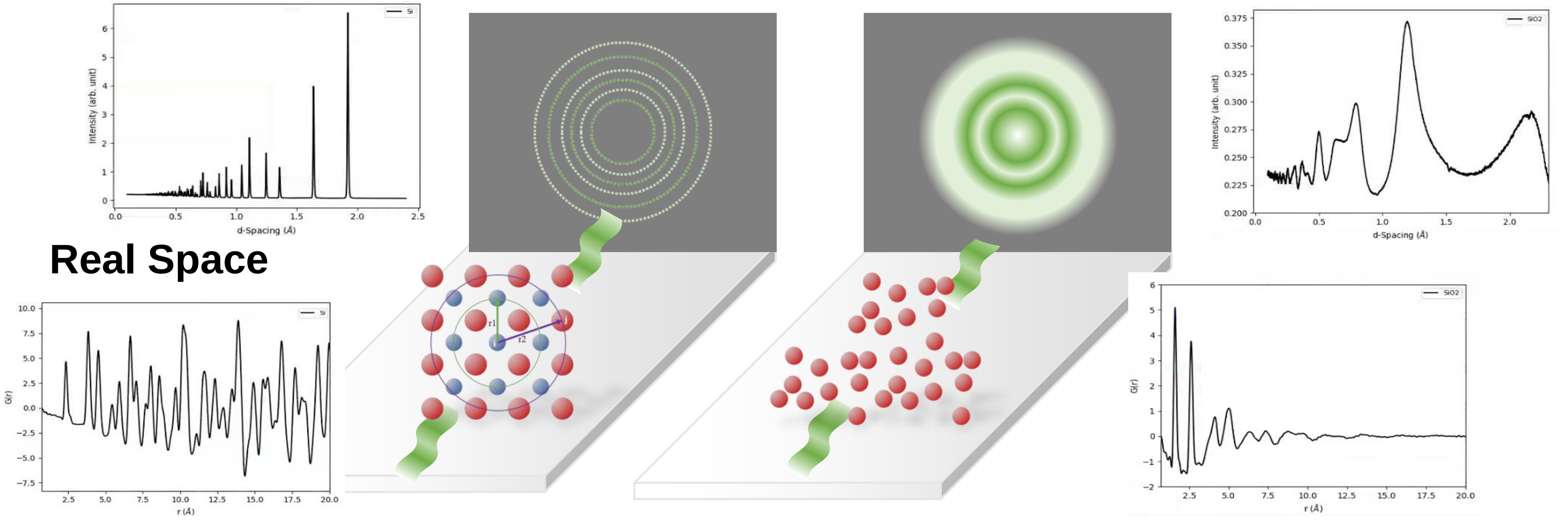
Neutron Total Scattering Pair Distribution Function (PDF)

From pattern to atomic structure

$$s(Q) = \frac{1}{N}\sum_{j,k} \overline{b_j}\,\overline{b_k}\,\sin(Qr_{ij})/Qr_{ij} = \frac{1}{N}\left(\sum_j \overline{b_j}^2 + \sum_{j\neq k} \overline{b_j}\,\overline{b_k}\,\sin(Qr_{ij})/Qr_{ij}\right) \qquad S_{dist}(\mathrm{Q}) = \rho_0\int_0^\infty 4\pi r^2 G(r)\frac{sinQr}{Qr}dr$$

Reciprocal Space

Real Space



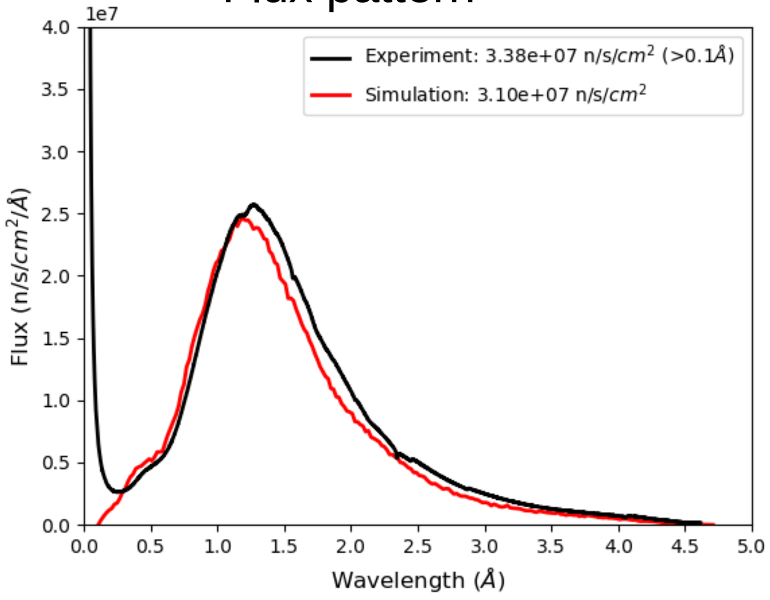
$$G(r) = \sum_{m,n} c_m c_n \overline{b_m}\,\overline{b_n}(g_{mn}(r) - 1) \qquad G(\mathrm{r}) = \frac{1}{(2\pi)^3\rho_0}\int_0^\infty 4\pi Q^2\,\mathrm{S(Q)}\frac{sinQr}{Qr}dQ$$

Peak positions in the PDF spectrum yield bond lengths, while peak areas yield coordination numbers.

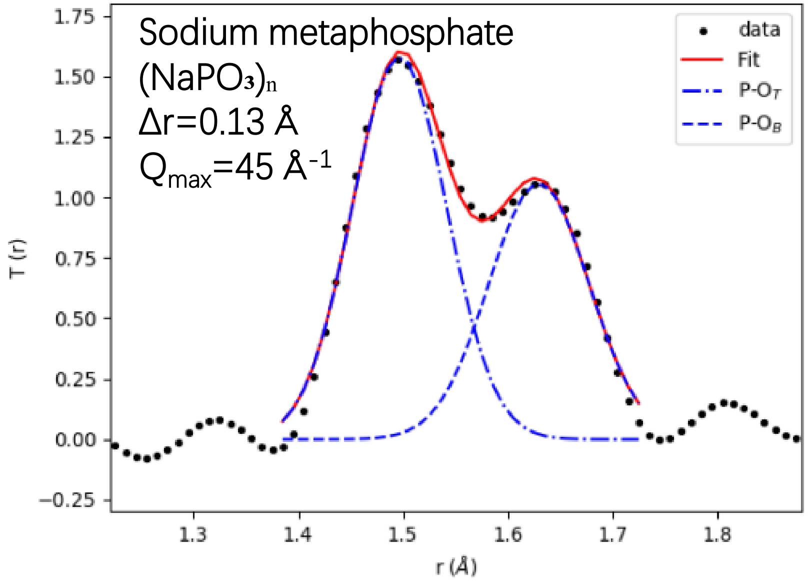
Neutron Total Scattering Multi-Physics Instrument (MPI)

CSNS-MPI a diffractometer has a wide Q range to get a high resolution in real space

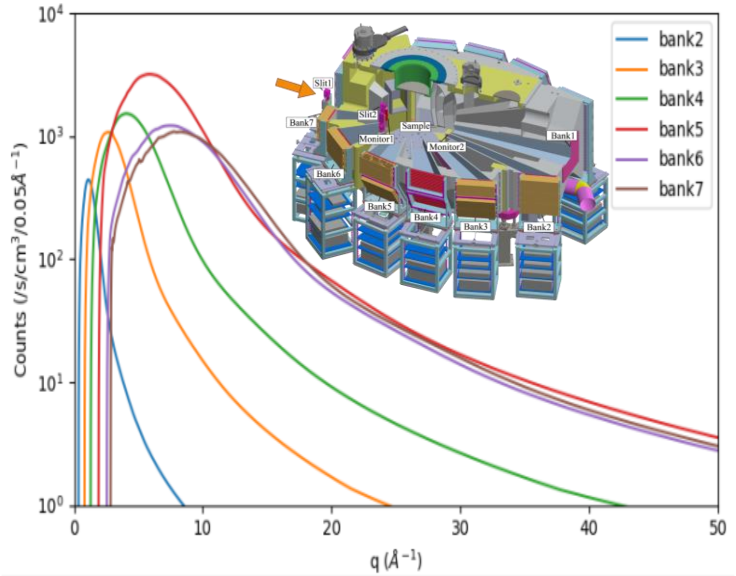
Flux pattern

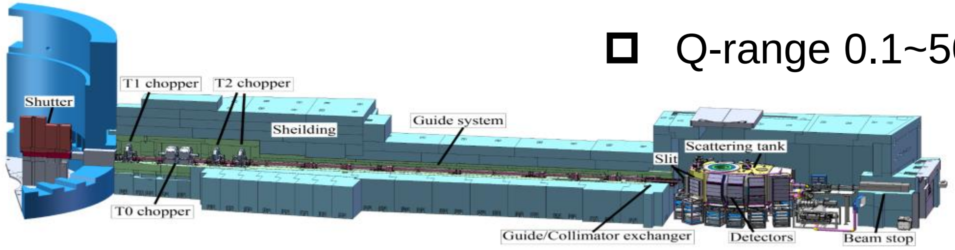


Real space resolution measurement



Detector C-number of Vanadium





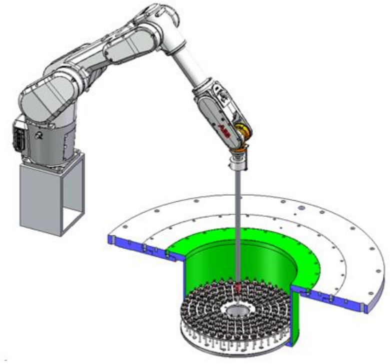
- Flux at sample position： 4.9×10^{7}/n/s/cm²@160kW （High Neutron Flux）
- Q-range 0.1~50 Å $^{-1}$， best resolution dQ/Q<0.4%， 0.1Å （High Real Space Resolution）

Neutron Total Scattering **Multi-Physics Instrument (MPI)**

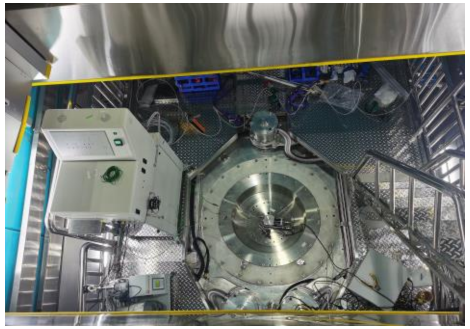
- **Challenges** Background, Multi-scattering effect, Absorption

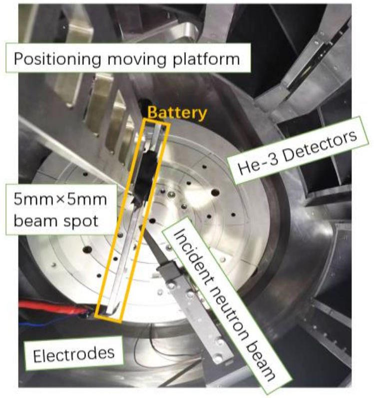
Multi-field coupled sample environment

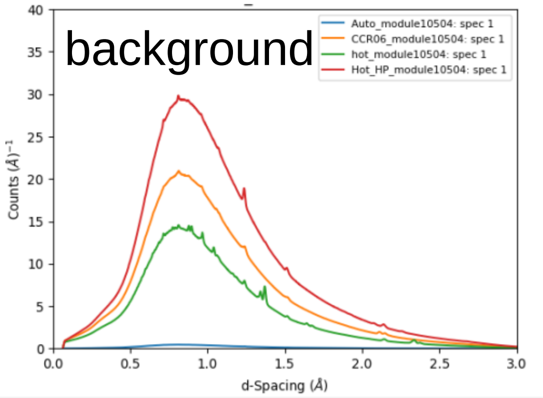
Robot Sample change



Gas absorption +CCR

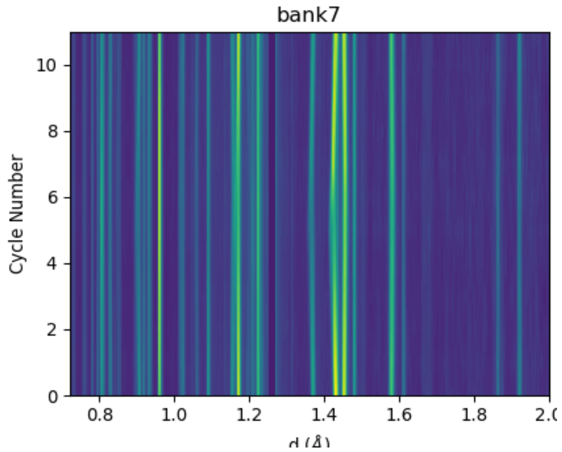




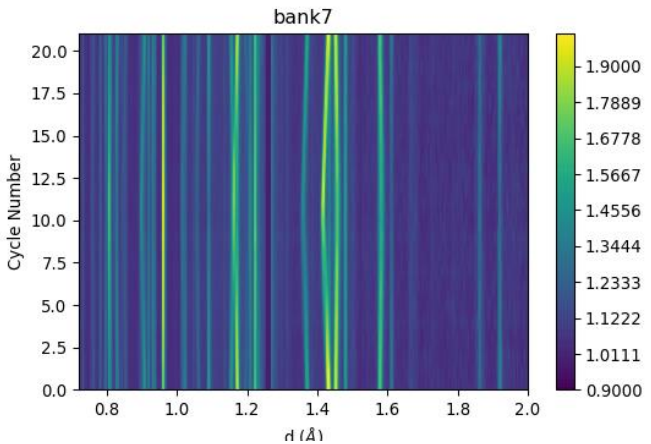


Fast-charging structural evolution

10C



5C



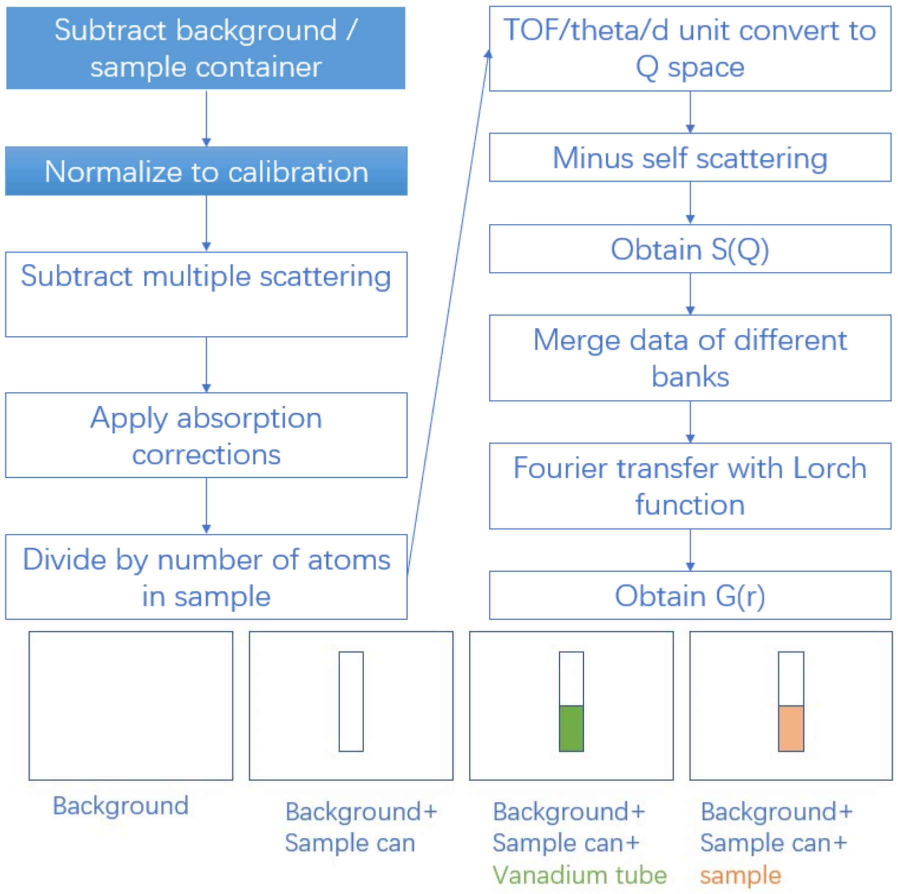
Through a data processing approach a fast time resolution of sub-minute level was achieved at MPI.

A reasonable data strategy not only yields reliable spectra, but also expands the functional boundaries of the diffractometer.

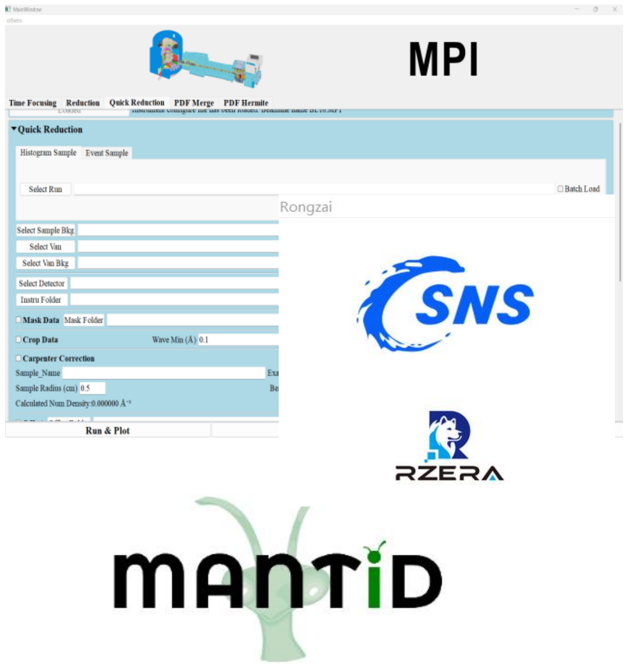
Data Processing and Modeling

From Raw Counts to Reproducible S(Q)

Logical flow diagram of data reduce

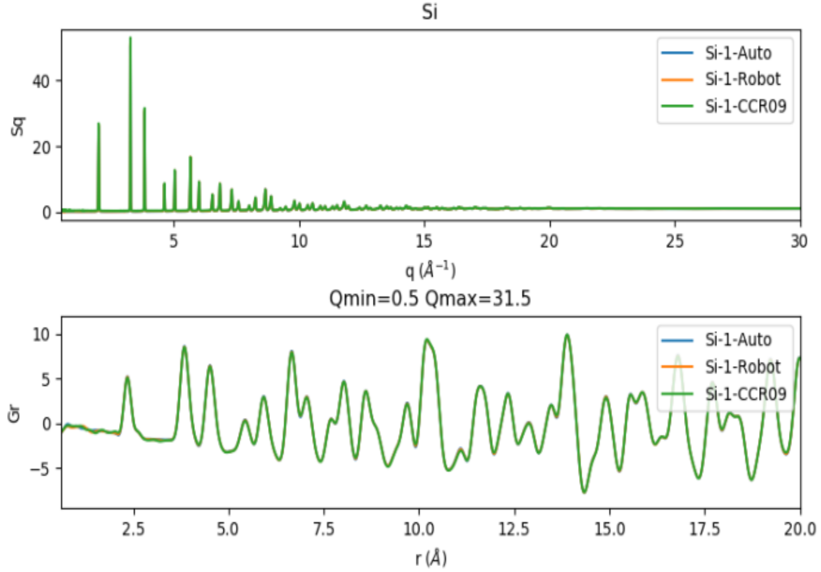


Reduce Software



$$I = \frac{(S - S_B) - \alpha(C - C_B)}{V - V_B}$$

consistent spectra



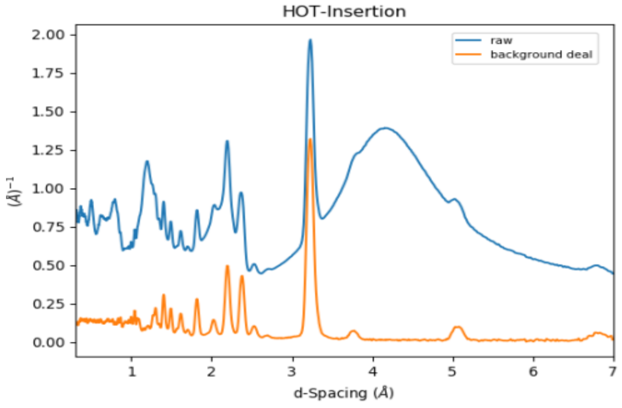
dBOT - Classify experiment types and establish a robust data correction method tuning to automatic reduction.

Standardized parameter processing ensures consistency across different sample environments and repeated measurements.

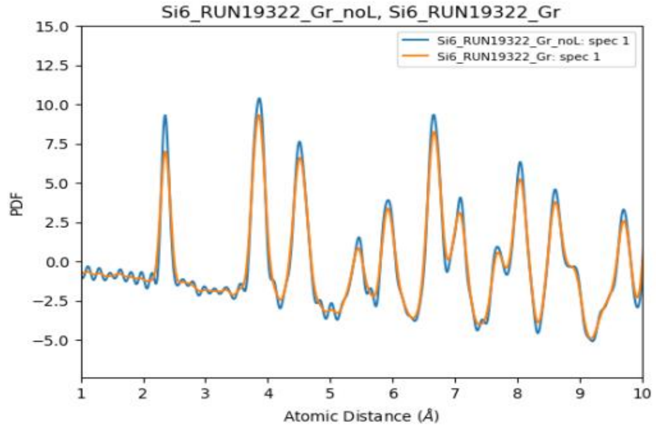
Data Processing and Modeling **Key Corrections**

Quality Control Ensuring Fidelity

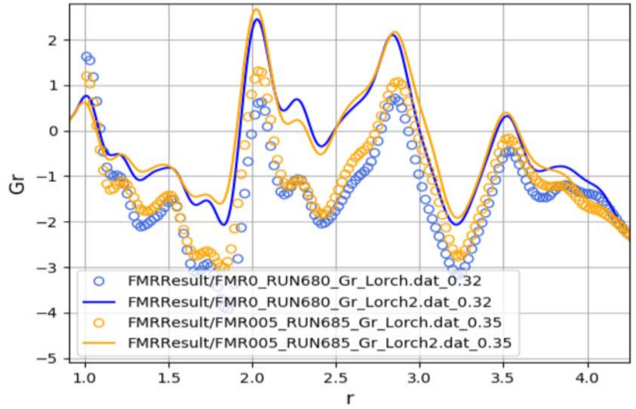
Decomposition Signals



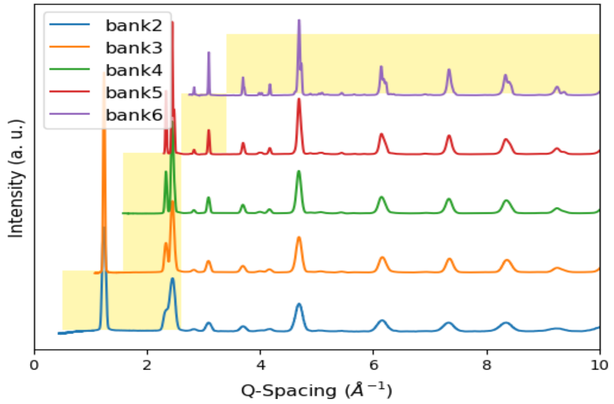
Lorch Modification



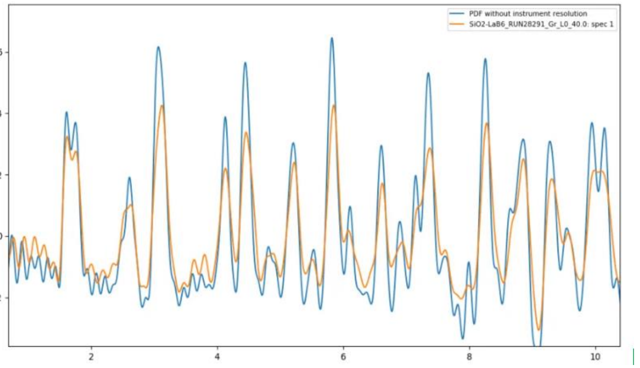
Small r correction



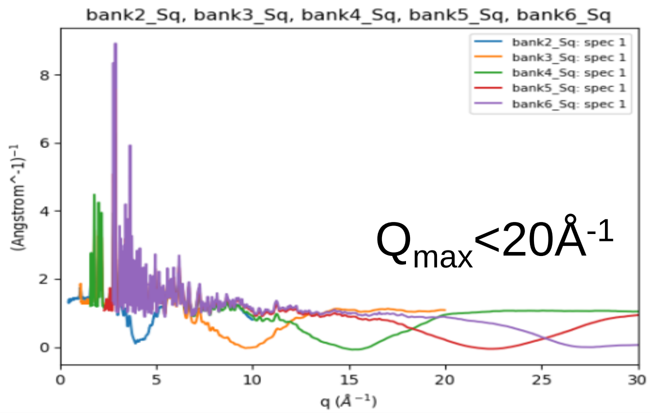
Merge data of different banks



Resolution improve



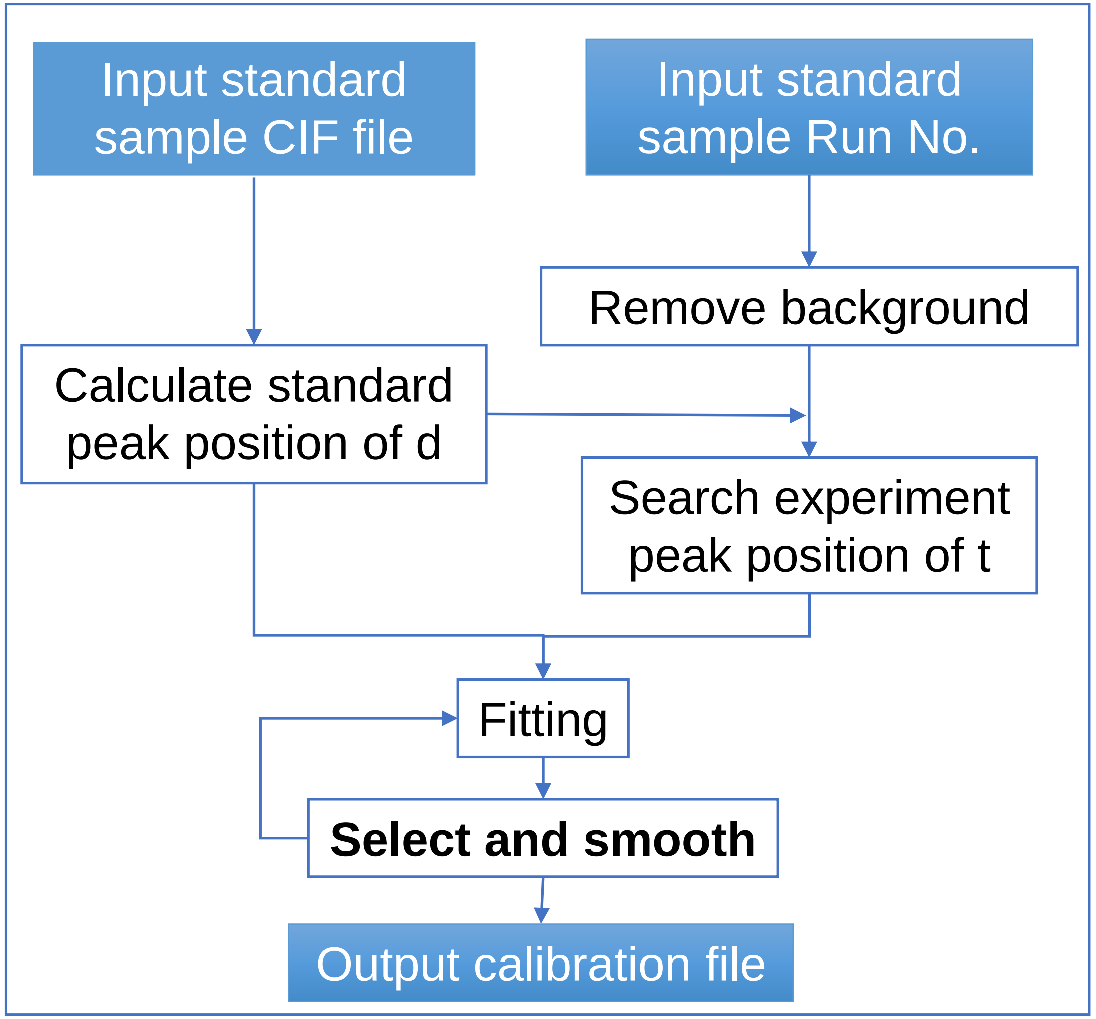
Strong absorption



Structure analysis starts with disciplined correction of systematic effects.

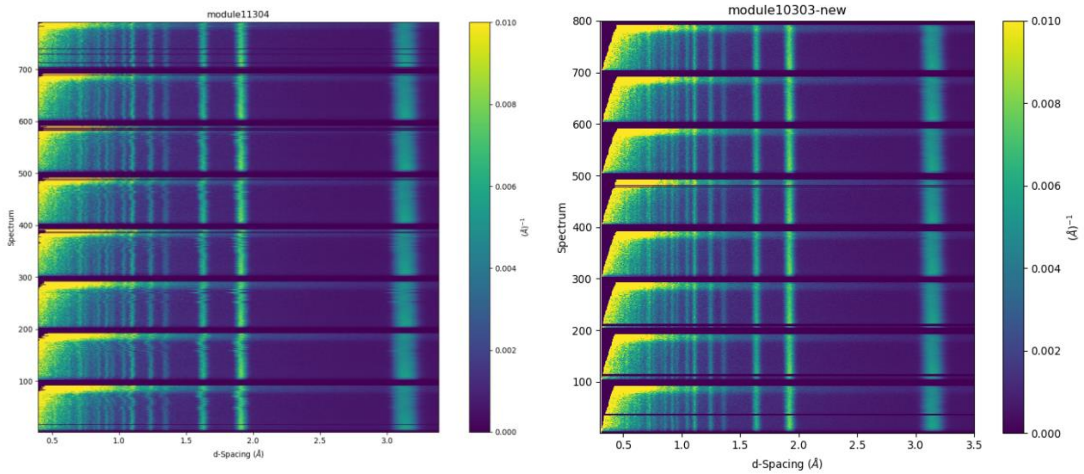
Data Processing and Modeling **Automated Calibration**

Neutron flight time focus



Wide-Q obtained through the automatic time focusing of combined multiple standard samples

Combined calibrated by the Si、LaB6、MOF…



Once set, it runs. The automated calibration converts instrument stability into guaranteed data quality.

Robust, Rapid, Repeatability

Data Processing and Modeling **Forward Modeling**

Refinement Philosophy

<table>
<tr><td rowspan="7">NPD</td><td rowspan="3">Background</td><td rowspan="2">Sample</td><td>Diffuse Scattering</td><td>Local structure,
amorphous ratio,
lattice dynamics,
magnetic disorder scattering</td></tr>
<tr><td colspan="2">Inelastic Scattering</td></tr>
<tr><td colspan="3">Scattering from Sample Environment, Air, etc.</td></tr>
<tr><td rowspan="4">Peaks</td><td>Position</td><td>Unit Cell & Magnetic</td><td>Crystal lattice parameters
& space group,
magnetic unit cell parameters
& space symmetry group,
qualitative phase analysis,
macroscopic strain</td></tr>
<tr><td>Intensity</td><td>Structural Parameters</td><td>Chemical composition,
atomic positions,
elemental distribution,
occupancy,
temperature factor,
texture,
quantitative phase analysis,
magnetic moment magnitude
& direction</td></tr>
<tr><td rowspan="2">Peak Shape</td><td colspan="2">Instrumental Parameters</td></tr>
<tr><td>Sample Broadening</td><td>Microstrain,
crystallite size,
magnetic domain size</td></tr>
</table>

Support popular software

Diffraction Refinement

Minimize
$$R_{wp} = \sqrt{\frac{\sum_i w_i \left[I_{obs}(Q_i) - I_{calc}(Q_i)\right]^2}{\sum_i w_i I_{obs}^2(Q_i)}}$$

PDF Refinement

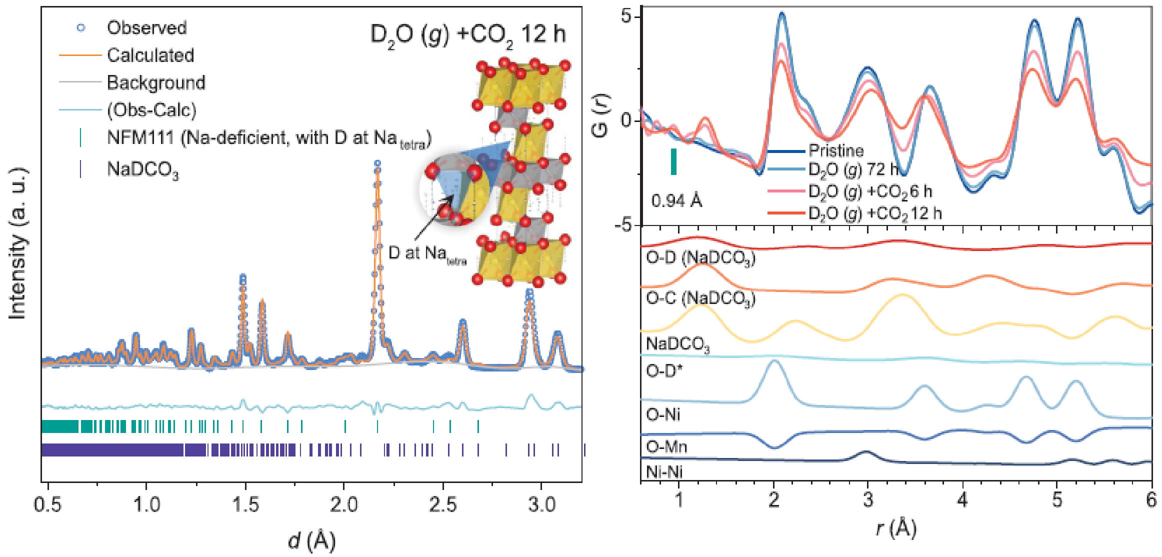
Minimize
$$R_{pdf} = \frac{\sum |G_{obs}(r) - G_{calc}(r)|}{\sum |G_{obs}(r)|}$$

Forward modeling seeks the best structural model whose calculated NPD/PDF matches the experimental data by Least Squares Method.

Data Processing and Modeling Application

Local structure in energy field

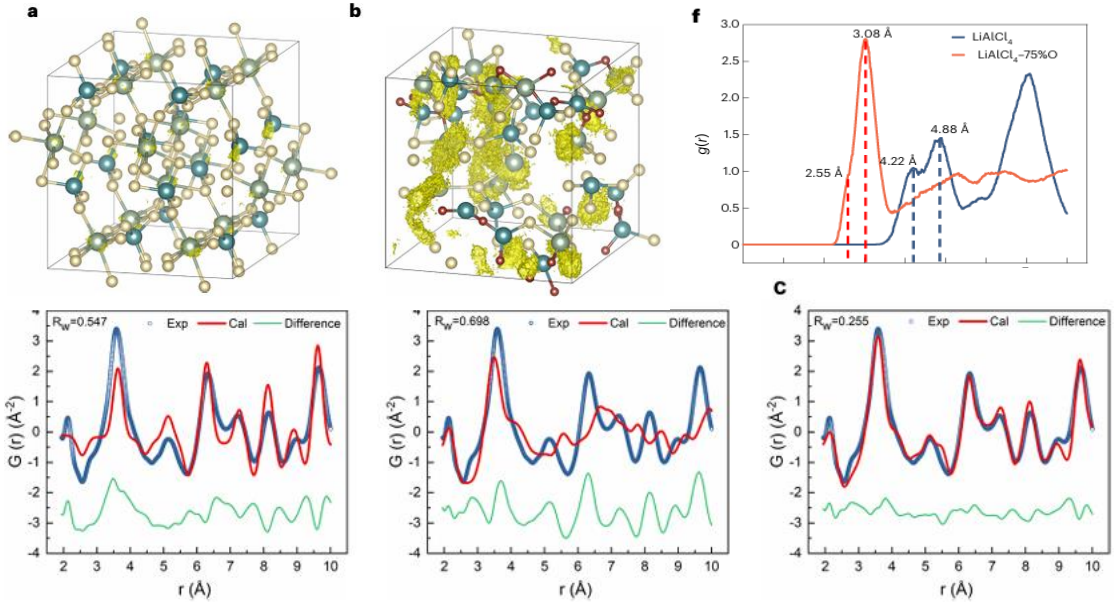
Air sensitivity mechanism of Na-layered oxides



Under the interaction of water and carbon dioxide, the exchange of H^+ (D) from water with Na^+ in the Na layer of NLO leads to acid hydrolysis.

Yongsheng Hu, et al, Science 385, 744–752 (2024)

Pressureless inorganic solid-state batteries



Pressure-free Li/Na-based solid-state batteries have been achieved in a class of viscoelastic inorganic glasses by substituting oxygen for chlorine in tetrachloroaluminate.

Yongsheng Hu, et al, Nature Energy 8, 1221–1228 (2023)

Intelligent Enhancement Can AI help analysis?

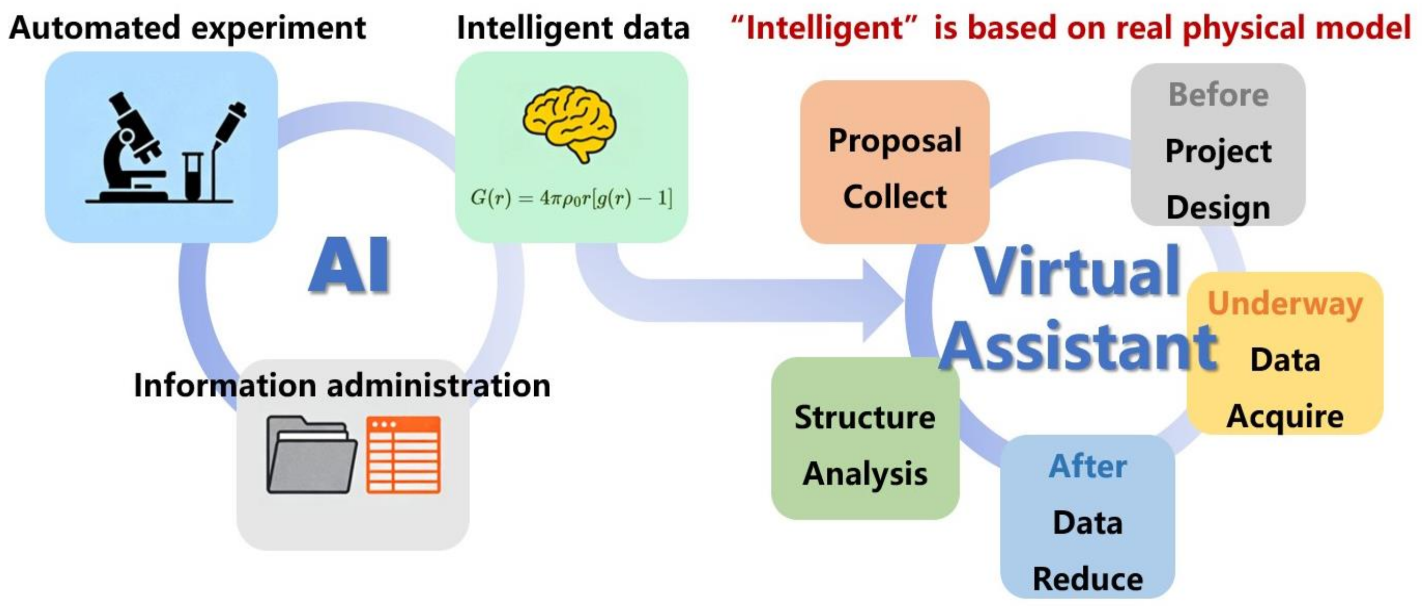
➢ Limitations

The traditional "forward modeling" method has trouble with complex, disordered materials. It needs a good starting guess and takes a lot of computing time.

Structure model-dependent
Strong parameter correlations
Incapable of batch processing
Heavily relies on empirical experience
Prone to local minimum traps
Computationally expensive

➢ Opportunities

Intelligent MPI Workflow



Classify space groups
 -Convolutional Neural Networks
Model atomic graphs
 - Graph Neural Networks
Accelerate refinement
and so on···

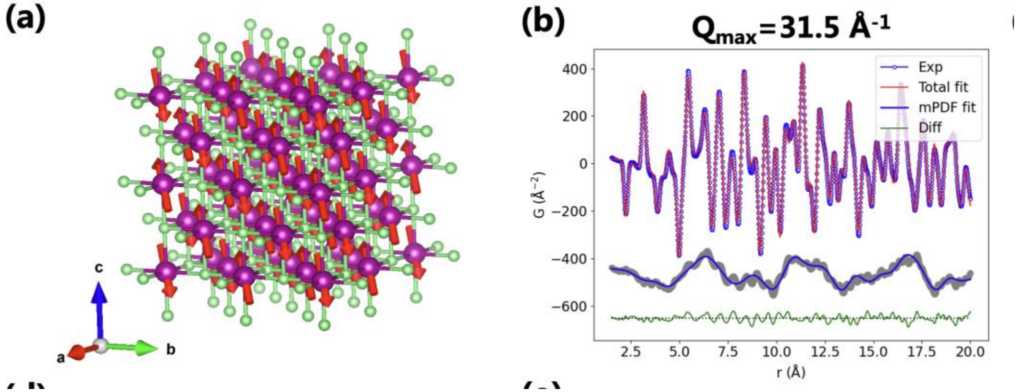
End-to-end mapping:
 S(Q) or G(r) → structural parameters
physical constrained neural networks
Real-time, unattended structural determination

Data are no longer just output—they drive the discovery of new structures.

Intelligent Enhancement Algorithms Develop

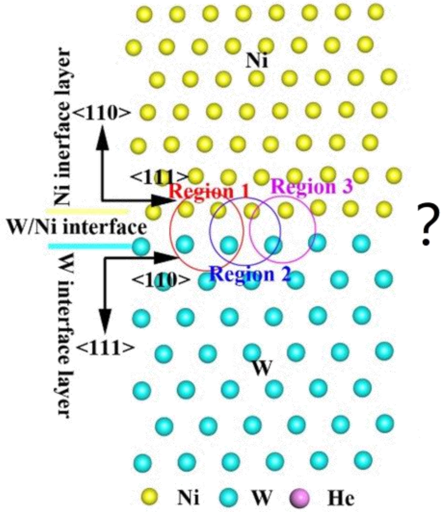
AI helps speed up data analysis—but it's new physics-based algorithms that really decide what we can discover.
New algorithm ideas keep pushing the MPI field forward.

Magnetic pair distribution function method：mPDF@MPI



Short-range magnetic correlations at the subnanometer scale are revealed in real space.

Interface analysis method :iPDF@MPI



MMCs exhibit a narrow near-interface region and weak scattering signals.

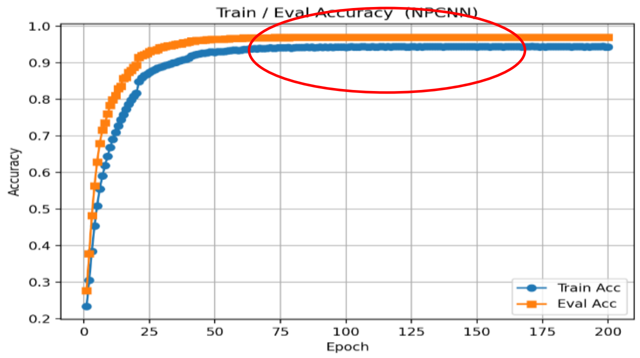
Fundamental algorithmic innovation continues to propel advances in the field of MPI.

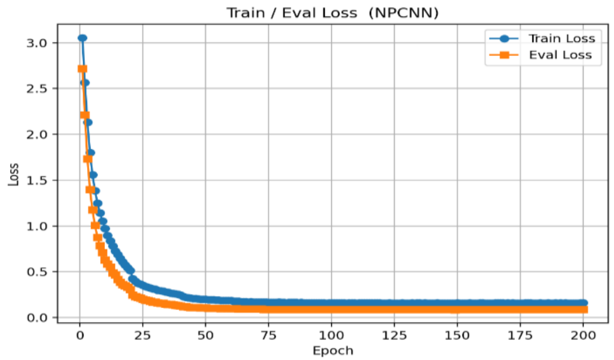
Nuclear Inst. and Methods in Physics Research, A 1069 (2024) 169967

Intelligent Enhancement AI / ML Method

Rapid data analysis for battery materials

Space group retrieval based on Convolutional Neural Networks (CNN)

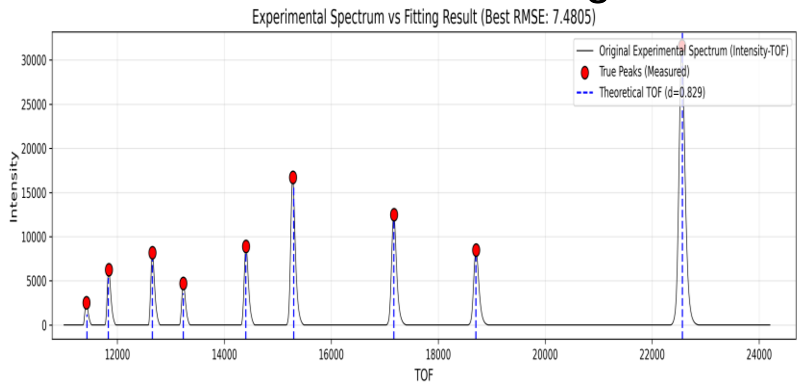




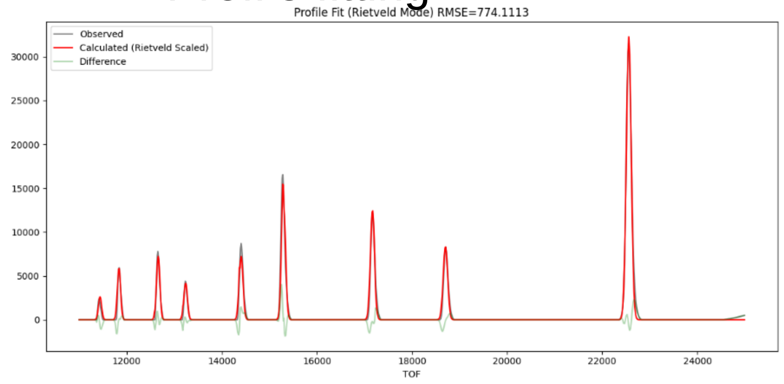
Loss and accuracy curves for the training and validation of space group classification

Crystal structure fitting based on Reinforcement Learning (RL)

Peak fitting

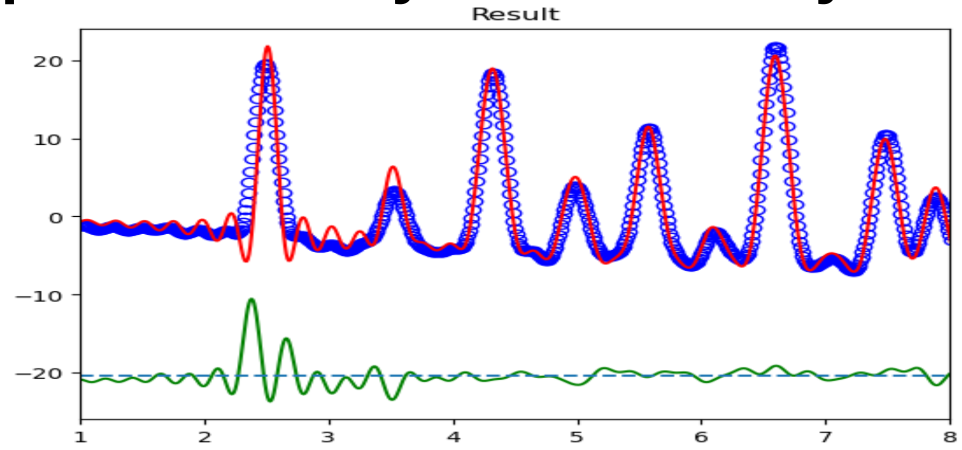


Profile fitting

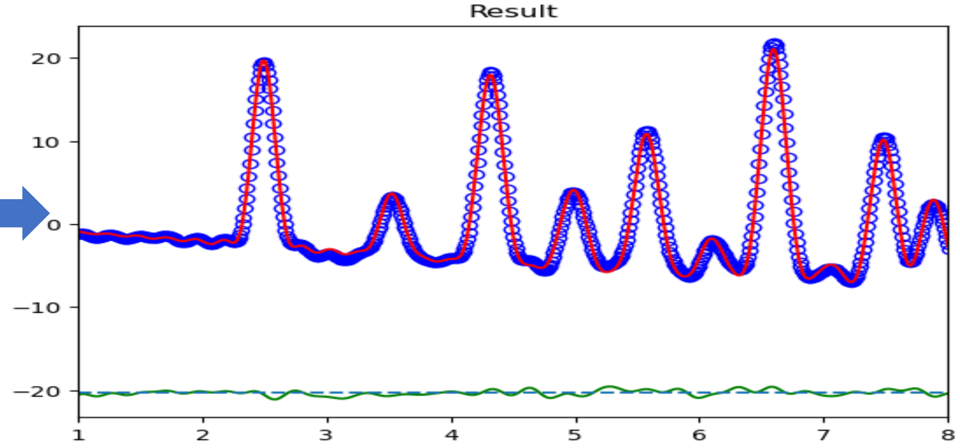


Intelligent Enhancement AI / ML Method

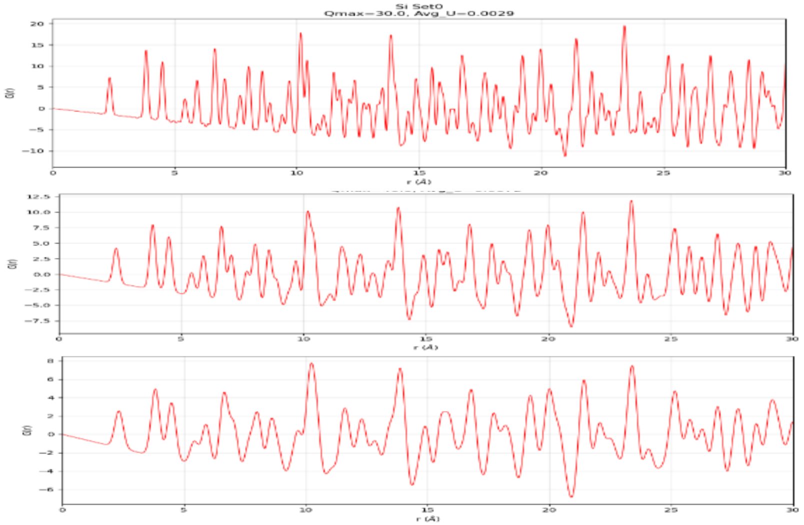
Rapid data analysis for battery materials



Least squares method:
Rw= 0.16268



Reinforcement learning based on the PPO algorithm:
Rw= 0.05181



dataset augmentation

Local optimum trap:

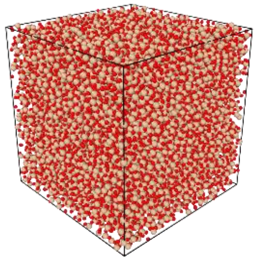
Global optimization methods (genetic algorithm, simulated annealing) are required to avoid convergence to wrong solutions.

Compared with the least squares method, it achieves a better goodness of fit and avoids getting stuck in local minima.

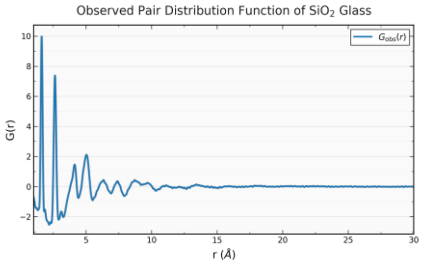
Intelligent Enhancement AI / ML Method

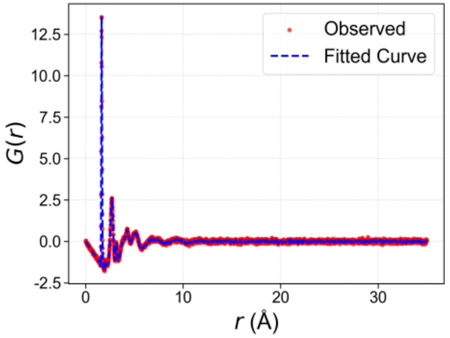
CNN-based PDF study of glass system

create physically reasonable initial structural using molecular dynamics

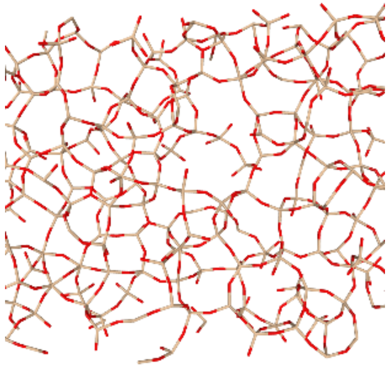


Experimental PDF data collection





Reverse Monte Carlo (RMC) method



Obtain the material structure

PDF→ CNN→Structure info

Fast and Simple

For measured data of the PDF data of glass, the proportions obtained by CNN are in good agreement with the actual values.

Summary and Outlook

1 Rigorous data engineering — from automated calibration and standardized processing to critical corrections

2 Innovative intelligent algorithms — advanced agents based on traditional analysis, physics-driven new methods, or AI tools including CNN and ML, expanding perceptual boundaries and analytical efficiency.

Outlook

1 Multimodal data fusion — combining neutron, X-ray data to construct a more complete structural picture

2 Collaborating with users and professional team— promoting the development of open-source toolchains.

Thank you!