

PAPER • OPEN ACCESS

Direct prediction of inelastic neutron scattering spectra from the crystal structure^{*}

To cite this article: Yongqiang Cheng *et al* 2023 *Mach. Learn.: Sci. Technol.* **4** 015010

View the [article online](#) for updates and enhancements.

You may also like

- [Testing the sign-changing superconducting gap in iron-based superconductors with quasiparticle interference and neutron scattering](#)
Tanmoy Das and A V Balatsky
- [Low-cost GNSS/INS integration for enhanced land vehicle performance](#)
Chao Chen and Guobin Chang
- [Identification of normal modes responsible for ferroelectric properties in organic ferroelectric CBDC](#)
Sanghamitra Mukhopadhyay



PAPER

OPEN ACCESS






RECEIVED
31 October 2022REVISED
19 December 2022ACCEPTED FOR PUBLICATION
13 January 2023PUBLISHED
1 February 2023

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Direct prediction of inelastic neutron scattering spectra from the crystal structure*

Yongqiang Cheng^{1,**} , Geoffrey Wu², Daniel M Pajeroski¹, Matthew B Stone¹ , Andrei T Savici¹ , Mingda Li³  and Anibal J Ramirez-Cuesta¹ 

¹ Neutron Scattering Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States of America

² Columbia University, New York, NY, United States of America

³ Massachusetts Institute of Technology, Cambridge, MA, United States of America

** Author to whom any correspondence should be addressed.

E-mail: chengy@ornl.gov

Keywords: inelastic neutron scattering, autoencoder, symmetry-aware neural network, structure-property relationship

Supplementary material for this article is available [online](#)

Abstract

Inelastic neutron scattering (INS) is a powerful technique to study vibrational dynamics of materials with several unique advantages. However, analysis and interpretation of INS spectra often require advanced modeling that needs specialized computing resources and relevant expertise. This difficulty is compounded by the limited experimental resources available to perform INS measurements. In this work, we develop a machine-learning based predictive framework which is capable of directly predicting both one-dimensional INS spectra and two-dimensional INS spectra with additional momentum resolution. By integrating symmetry-aware neural networks with autoencoders, and using a large scale synthetic INS database, high-dimensional spectral data are compressed into a latent-space representation, and a high-quality spectra prediction is achieved by using only atomic coordinates as input. Our work offers an efficient approach to predict complex multi-dimensional neutron spectra directly from simple input; it allows for improved efficiency in using the limited INS measurement resources, and sheds light on building structure-property relationships in a variety of on-the-fly experimental data analysis scenarios.

1. Introduction

The collective vibration of atoms in a material is a fundamental ‘vital sign’ that dictates many of its properties. Experimentally, it is often probed by spectroscopic methods in which a beam of probe particles (photons, neutrons, or electrons) is directed into the sample. The atomic dynamics is then measured by the energy and momentum transfer when the particles are scattered. Inelastic neutron scattering (INS) is a technique with a few unique advantages. For example, it is sensitive to light elements such as H, O, Li, etc, which are often challenging to see with other methods such as x-ray scattering. The high penetration of neutron beams allows the use of complex sample environments and allows neutrons to probe deep into bulk material. INS is also ideal for studying phonons because the energy and momentum of thermal neutrons are well matched to those of phonons, allowing high-resolution measurement of the whole dynamic structure factor $S(\mathbf{Q}, E)$, where \mathbf{Q} is the momentum transfer and E is the energy transfer. The dynamic structure factor contains rich microscopic information on correlations and interactions. For instance, neutrons interact directly with nuclei, making it possible to calculate INS spectra from lattice or molecular dynamics

* This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

simulations rigorously. In cases where researchers are probing for magnetic excitations using neutron scattering, phonons and molecular vibrations may be considered spurious signals overlapping with the desired magnetic signal. In this case, having simple planning tools to calculate the phonon spectrum would allow scientists to measure the magnetic excitations more efficiently and effectively. Unlike Raman scattering or Infrared spectroscopy, INS has no selection rules for lattice excitations or molecular vibrations, meaning that all such modes can, in principle, contribute to INS intensities. This makes INS powerful, but also the spectra more challenging to decipher and dependent on comprehensive analysis and modeling. In fact, modeling INS from first-principles calculations (e.g. using density functional theory, DFT) requires significant computing resources and expertise. It is one of the most significant barriers limiting the scientific productivity at INS experimental facilities.

With the rapid advance in high-performance computing and machine learning, a data-driven solution to overcome these barriers becomes increasingly possible and intriguing [1–4]. Recently, Chen *et al* [5] demonstrated that one could directly predict the phonon density of states (PDOS) from a material's crystal structure without explicitly calculating phonons via DFT. Fundamentally, the vibrational behavior of a material is indeed determined by its structure (i.e. the atomic coordinates and the elemental species), but the connection is rather complicated and implicit. Thus, a simple prediction seemed almost impossible to establish. The key to success by Chen *et al* is using an Euclidean neural network (as implemented in the e3nn library) [6, 7], a highly efficient way to represent the crystal structure with intrinsic equivariance to three-dimensional (3D) rotation, translation, and inversion. They showed that when the crystal structure represented by e3nn is trained with the corresponding pre-calculated PDOS (represented by a scalar of 51 numbers in the range of [0,1]), the resulting neural network can make reasonably accurate predictions of PDOS. This achievement opens the door for directly predicting other related properties, such as INS spectra, from the crystal structure. There are, however, two additional barriers before this can be applied to INS. First, we need an INS spectra database for training. For a rigorous calculation, one needs the PDOS and complete phonon information, including eigenfrequencies and polarization vectors, directly coupled with performing high-throughput calculations of INS spectra for different experimental scenarios. Second, although the prediction of the one-dimensional (1D) $S(E)$ spectra would be helpful, the two-dimensional (2D) $S(Q,E)$ map contains much more information and is what most INS users are interested in when measuring powder samples (the 4D $S(\mathbf{Q},E)$ for single crystals will be deferred for examination in future studies). A 2D map, however, contains significantly more variables (e.g. for a 300×300 map in wave-vector, Q , and energy transfer, E , space, that is 90 000 variables to predict), and thus direct prediction is unlikely to work. In this paper, we demonstrate how the above issues can be solved by implementing a workflow for high-throughput INS spectra calculation, as well as an autoencoder to compress and represent the 2D $S(Q,E)$ in latent space. The hybrid network proposed in this work, combining an autoencoder for data compression and a symmetry-aware neural network for feature prediction in the latent space, can be used for not only INS prediction, but more generally connecting the crystal structure directly with other multi-dimensional properties of the material.

2. Results

To develop the machine learning model for prediction, a training dataset of INS spectra covering all common elements and typical crystal structures needs to be produced. We have adopted the phonon database by Togo [8–10] as it provides the calculated force constants needed to obtain the eigenfrequencies and polarization vectors (via Phonopy) [11]. The OCLIMAX [12] program, which allows streamlined, rapid, and accurate calculations of INS spectra, was used to obtain the 1D $S(E)$ and 2D $S(Q,E)$ from the Phonopy output. Following this workflow, we created an INS database containing 10 023 inorganic crystals. With the DFT-derived force constants available (i.e. the most time-consuming DFT calculations already performed by Togo), the high-throughput workflow allows the construction of such an INS database in hours (in this case, on a computer cluster equipped with 50 compute nodes or 1600 CPU cores). This makes it possible to customize the database training for a specific application (e.g. a specific instrument, sample environment, resolution, Q/E range, or material class). In this study, the 1D spectra were simulated for the VISION spectrometer [13], and the 2D $S(Q,E)$ were simulated in the energy transfer (E) range of 0–150 meV and momentum transfer (Q) range of 0–15 \AA^{-1} . All simulations were done at 0 K (phonon population terms at finite temperatures can be applied later following the Bose–Einstein distribution).

A straightforward extension of the original work by Chen *et al* [5] is to train a Euclidean neural network to predict 1D $S(E)$ spectra, which is, in principle, the same as the prediction of the PDOS. Indeed, we have achieved similar performance, as shown in figure 1, using essentially the same training parameters. The accuracy of the prediction can be seen in the distribution of the mean squared errors (MSE), as well as the example spectra with MSE ranging from the smallest 25% (top row) to the largest 25% (bottom row).

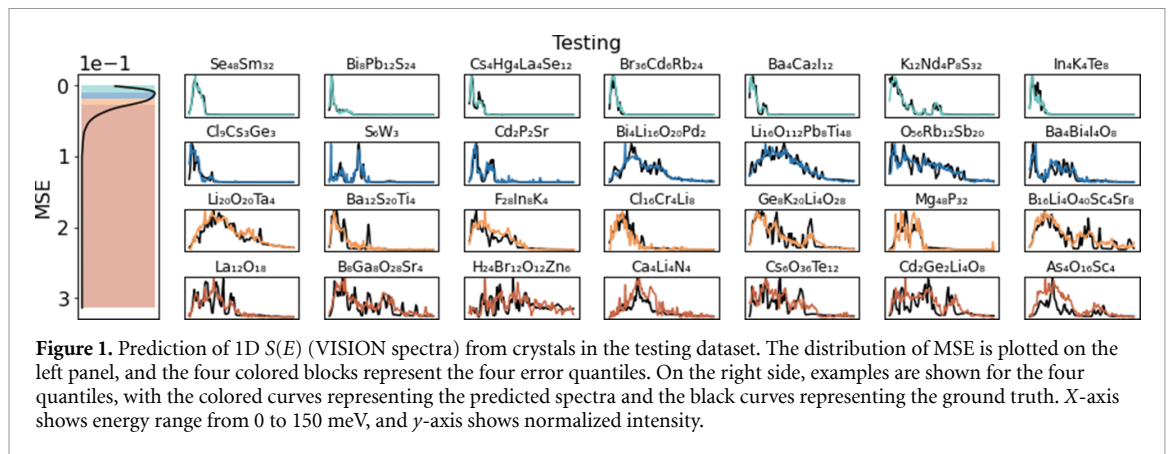
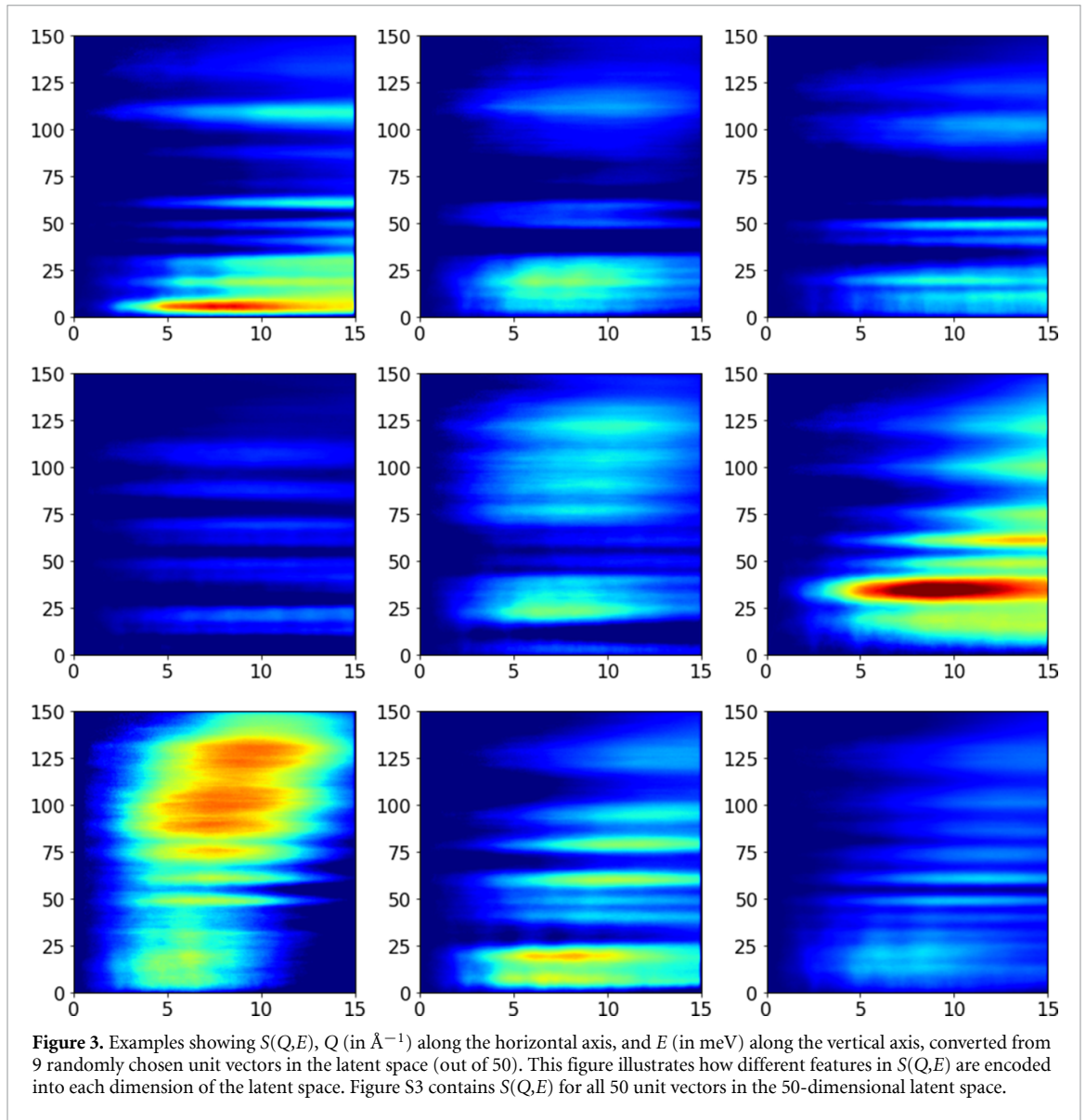
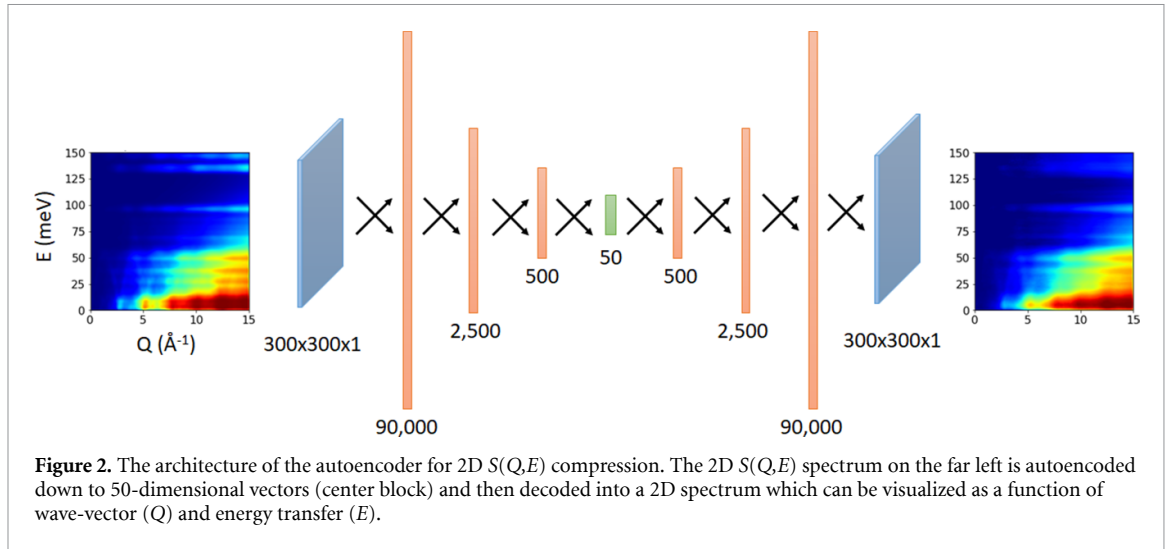


Figure 1. Prediction of 1D $S(E)$ (VISION spectra) from crystals in the testing dataset. The distribution of MSE is plotted on the left panel, and the four colored blocks represent the four error quantiles. On the right side, examples are shown for the four quantiles, with the colored curves representing the predicted spectra and the black curves representing the ground truth. X-axis shows energy range from 0 to 150 meV, and y-axis shows normalized intensity.

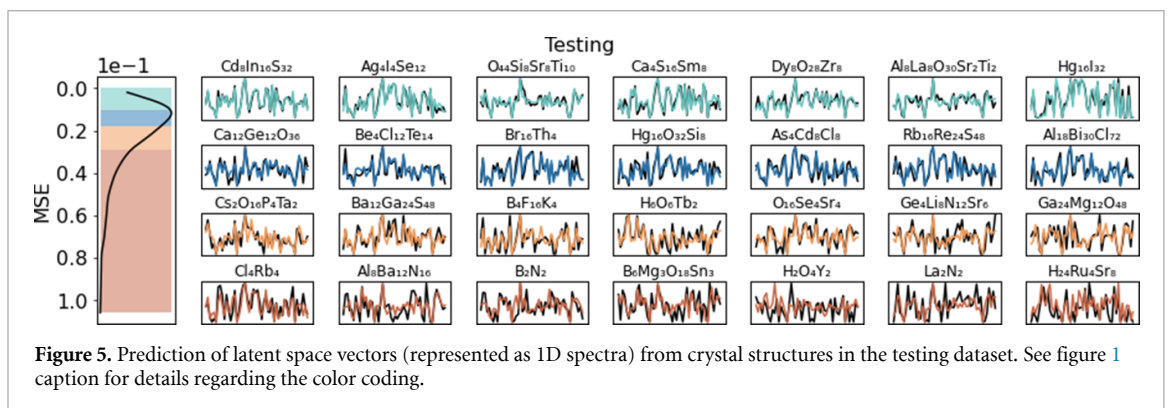
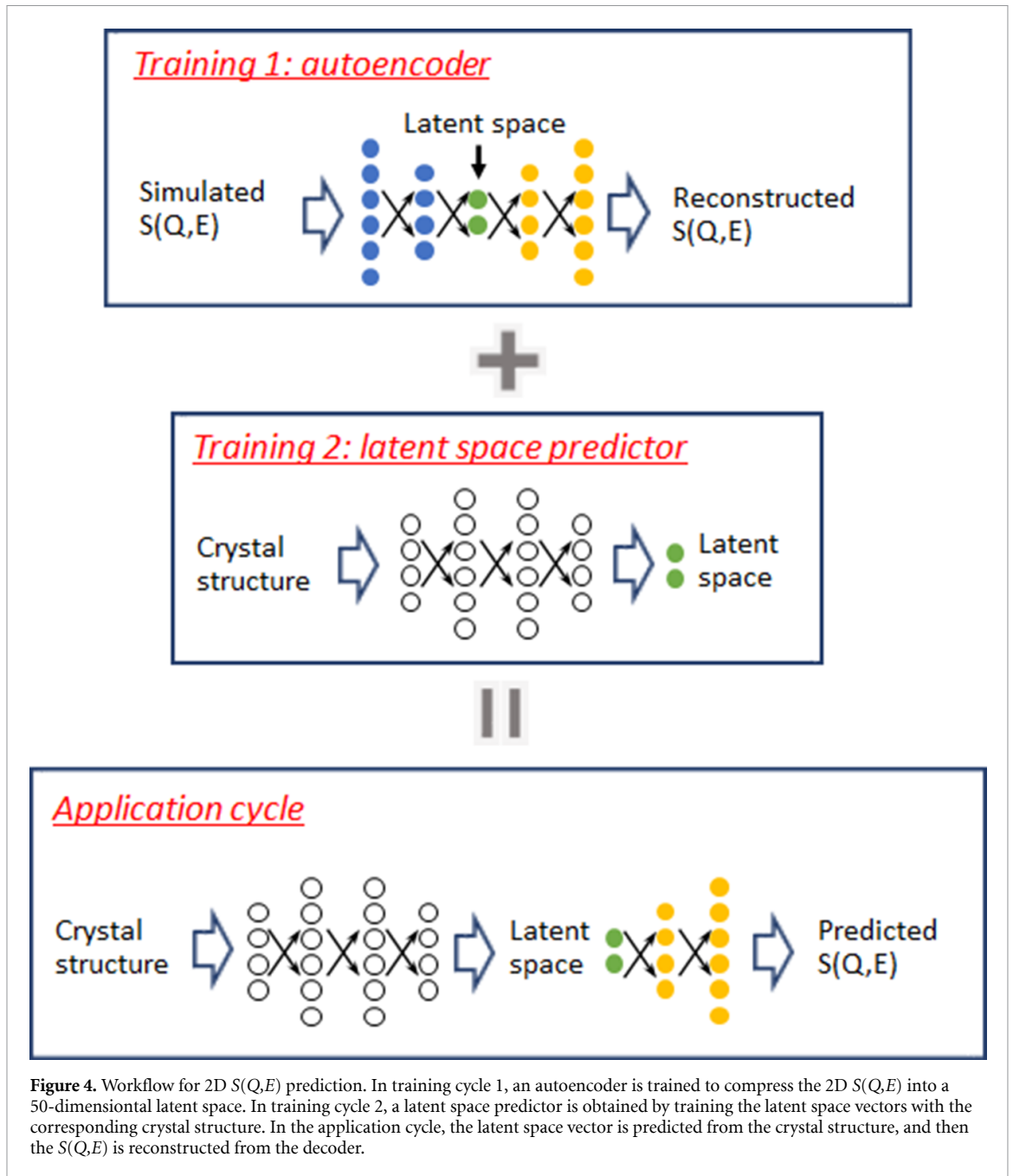
Overall, even in the worst-case scenarios, the general features of the spectra, such as the energy range of fundamental excitations, distribution of phonon bands and band gaps, position and relative intensity of the major peaks, are well predicted. Some discrepancies such as the sharpness of the peaks and some fine structures are observed. These observations are important for the users to decide whether the prediction accuracy is sufficient for their specific applications. In addition to what has been presented in figure 1, different INS features can also be trained as the prediction output. For example, one can train one neural network to predict the fundamental excitations and another neural network for all the higher order combinations and overtones. One can also obtain partial contributions to the total spectra for a specific element or atom. This will be very useful in guiding the decomposition of the experimentally measured total spectra.

While the prediction of 1D $S(E)$ is helpful, to fully take advantage of neutron scattering, which can measure the 2D $S(Q,E)$ for powder samples with high resolution and accuracy, we would like to be able to perform the same kind of prediction, from structure directly to 2D $S(Q,E)$. However, while a 1D $S(E)$ in figure 1 only has 50 variables to predict, a 2D $S(Q,E)$, assuming in a typical case of 300×300 bins, needs 90 000 variables to reconstruct. This is beyond what can be achieved with the current neural network. Fortunately, many 2D $S(Q,E)$ maps show certain similarities in the overall pattern. For example, bands and stripes indicating energy levels and phonon branches, common intensity profiles with increasing Q , etc. This means the 90 000 variables are far from uncorrelated, and one should be able to significantly compress the 2D $S(Q,E)$ to a much smaller dimension without losing the essential features. Recently, Samarakoon *et al* [4] demonstrated that an autoencoder could efficiently compress single crystal diffuse neutron scattering data, a 3D dataset, into a small latent space. Enlightened by this work, here we train an autoencoder to compress the synthetic 300×300 $S(Q,E)$ into a 50-dimensional latent space vector, which will later be associated with the crystal structure and another neural network trained for predictions. In fact, in two follow-up publications [14, 15], Samarakoon *et al* also demonstrated the usefulness of this approach in exploring the parameter phase space in magnetic materials. The architecture of the autoencoder is illustrated in figure 2 (details in section 4), together with an example of the reconstruction. The successful reconstruction (including some fine details and coherent scattering features) indicates that we can ‘decompose’ any 2D $S(Q,E)$ onto 50 ‘principal components’, which are the unit vectors in the 50-dimensional latent space. Figure 3 shows a few examples of $S(Q,E)$ decoded from the unit vectors, which illustrate how different patterns in various (Q,E) ranges are highlighted by different unit vectors. The full spectra of all 50 unit vectors are shown in figure S3.

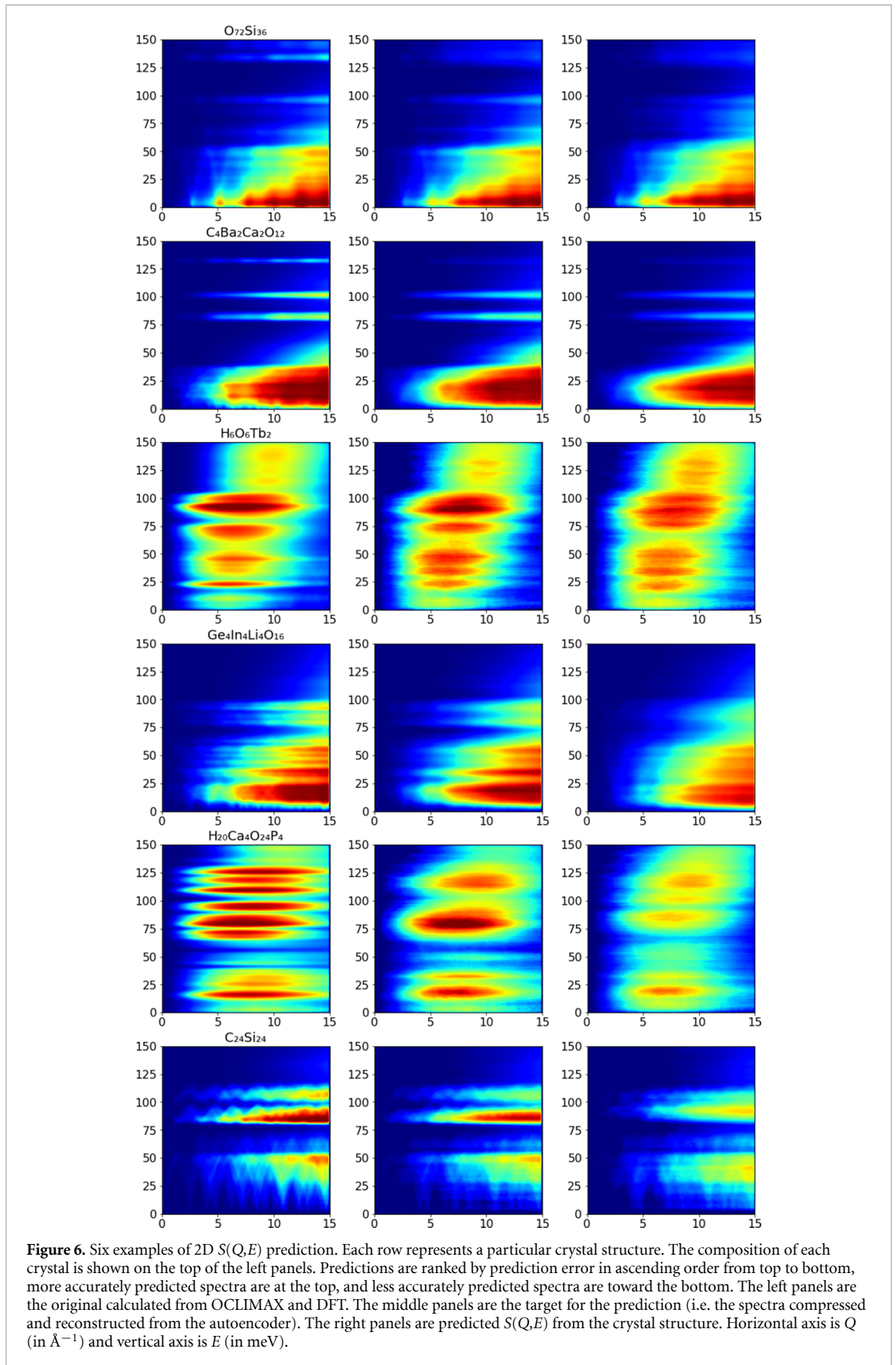
With the compressed representations of the 2D $S(Q,E)$, the next step is to connect the crystal structures with the latent space vectors. The trained encoder and decoder will be used separately in this step. Specifically, we run all the 10 023 $S(Q,E)$ through the encoder and convert them into latent space representations. Then we perform another training of a Euclidean neural network to predict the latent space representations from the crystal structure (i.e. a latent space predictor). Here the 50-dimensional latent space vectors are treated the same as the 1D $S(E)$. An important detail in this training is that a relatively large cutoff radius of 6 Å was used to include more neighbors into consideration. This is intended to capture the coherent effects present in some of the 2D $S(Q,E)$, especially for crystals with small unit cells (simple and high symmetry structure) and coherent scattering elements. In the application cycle, we take a crystal structure as input with the autoencoder and the latent space predictor. The latent space predictor will predict the latent space vector. The predicted latent space vector can then be passed through the decoder to reconstruct the predicted 2D $S(Q,E)$. The workflow is illustrated in figure 4, and the predicted latent space vectors for some crystals that are not part of the training dataset are shown in figure 5, in the same style as figure 1, with the latent space vectors presented as 1D spectra.



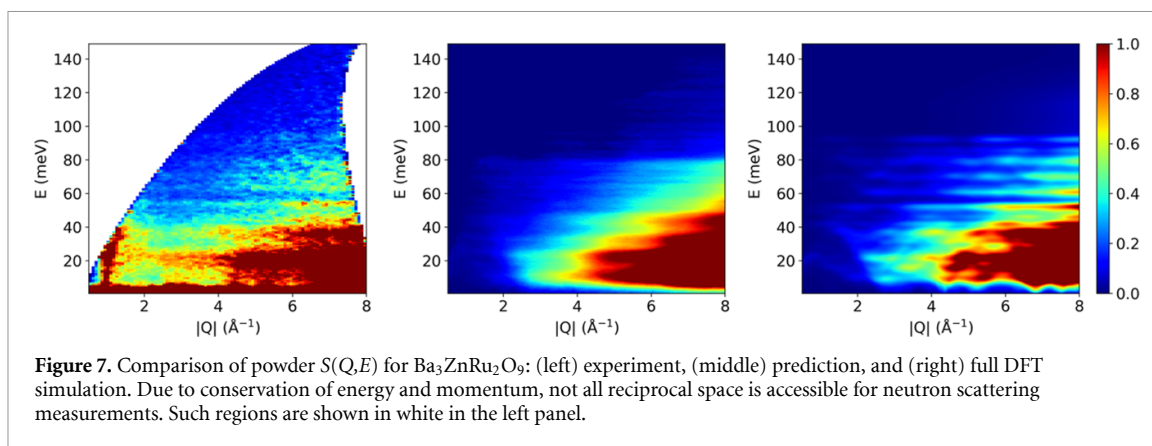
The results of the 2D $S(Q,E)$ prediction are shown in figure 6. The left column shows the calculated $S(Q,E)$ based upon DFT and OCLIMAX calculations. The middle column shows the corresponding compressed-and-reconstructed $S(Q,E)$ from the ‘true’ latent space vector as obtained from the autoencoder.



The right column shows the predicted $S(Q,E)$ from the crystal structure (i.e. reconstructed from the 'predicted' latent space vectors). Therefore, a comparison between the left and middle columns indicates how faithfully the autoencoder compresses and reconstructs the $S(Q,E)$; and a comparison between the middle



and right columns reveals the performance of the latent space predictor. Among the six chosen examples, the top two have the MSE in the lowest 25%; the MSE of the middle two is in the range of 25%–75%, and the bottom two have the highest MSE (in the highest 25%). More examples can be seen in figures S6–S8, and the distribution of MSE can be found in figure S9, which contains contributions of MSE from both the



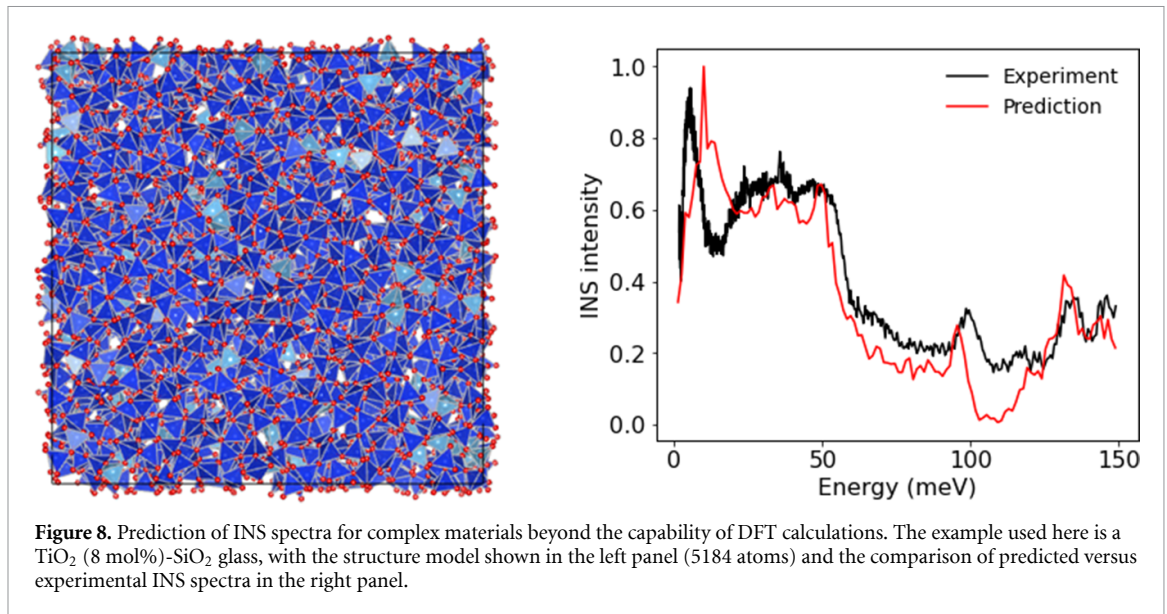
autoencoder and the latent space predictor. For most systems, the predictions are fairly accurate, reproducing the most distinctive features and sometimes even the details such as the coherent scattering patterns. It is also observed that the worst predictions are often on systems containing hydrogen (thus having multiple strong and sharp peaks in the spectra) or very strong and unique coherent effects (with unique phonon dispersion patterns that are difficult to learn if such patterns are rare/non-existent in the training data). One possible way to alleviate this issue is to train multiple neural networks for different classes of materials based upon composition and structure, i.e. we sort the data into different categories according to the distinctive features they may have and train a dedicated neural network for each category. For example, hydrogen containing materials can be treated separately because they usually have their distinctive features in $S(Q,E)$.

Thus far, we have been working with synthetic INS data to train and test the model. In the following example we will show a direct comparison with experimental data that demonstrates how the machine learning model is used in a real application. A research team is studying the magnetic properties of $\text{Ba}_3\text{ZnRu}_2\text{O}_9$, which is not found in our training database. The powder INS data was collected at SEQUOIA [16] at 5 K with an incident energy of 160 meV. Figure 7 compares the $S(Q,E)$ from experiment, machine learning model prediction, and full DFT simulation. While the full DFT simulation reveals more details associated with coherent scattering, the rapid prediction from the machine learning model is similar to the experimentally measured spectrum. We emphasize that the full DFT calculation took days on a computer cluster using hundreds of CPU cores while the machine learning prediction was done on a single graphics processing unit (GPU) in a fraction of a second. Moreover, for materials with more complicated crystal structures, full DFT calculations may not be feasible. The comparison in figure 7 immediately shows anomalous scattering at low Q (below 2 \AA^{-1}) in the experimental spectra, which is assigned to magnetic correlations. The level of accuracy from the rapid prediction is sufficient for many application scenarios such as experiment planning, ad-hoc decision making, and quick analysis where full DFT simulation is either impractical or unnecessary. One thing can be learned from the rapid prediction in this case is that because of overlapping contributions of magnetic scattering and phonon scattering in the region of 40 meV energy transfer, one may need to adjust the measurement to measure lower ranges of wave-vector transfer to fully quantify the magnetic scattering in this portion of the spectrum.

To further demonstrate that our approach can handle much more complex systems beyond the crystals in the training dataset (or in general beyond crystals that can be calculated by DFT), in figure 8 we show the predicted INS spectrum for a TiO_2 (8 mol%)- SiO_2 glass. Since the structure is amorphous and the distribution of Ti is also random, a large structure model must be used (in this case it contains 5184 atoms with 3456 O, 1590 Si, and 138 Ti, as depicted in figure 8). Phonon calculation with DFT would be impractical for such a complex system. However, our machine learning model can still make rapid predictions in the same way as what it does for other simpler crystals, and the accuracy of the prediction is decent when it is compared with experimental data reported in literature [17]. This example clearly demonstrates how our method can enable and greatly accelerate INS data analysis, especially as additional complex materials are being measured at the neutron scattering facilities.

3. Discussion

The availability of the synthetic INS database and the capability to perform on-demand, effortless prediction of $S(Q,E)$ are expected to have significant value to the neutron scattering and other spectroscopic research communities. It will advance scientific research using neutron scattering by enabling better designed and more efficient experiments given the extremely limited experimental resources. The ability to rapidly



understand entire regions of the phonon spectrum will significantly enhance the cycle of structural design and optimization of materials' properties. The developed framework in this work will be deployed to the data analysis servers at neutron scattering facilities (e.g. analysis.sns.gov at the Spallation Neutron Source) to support the facility users in the stages of experimental planning, execution, and data analysis. The users will be able to upload their crystal structure (e.g. in the standardized format of crystallographic information file (CIF) files) to the server and they will receive the predicted INS spectra immediately. A graphic user interface will be developed to further facilitate visualization for a better user experience. The machine learning model on the server will be retrained and updated periodically as new training data becomes available. Users may also choose to download the trained model to perform predictions on their own computers. To illustrate how tools developed based on this work can help, we assume a user would like to perform an experiment at the SEQUOIA direct geometry chopper spectrometer instrument. Before the measurements, important instrument settings such as the incident neutron energy and chopper frequency need to be decided. Knowing approximately what the spectra would look like will allow users to obtain the optimal setup efficiently well before arriving at the instrument. This software could even be used in beamtime proposals in order to illustrate ranges of energy transfer that will be explored if beamtime is granted to the applicant. During the experiment, the collected data will be examined to determine whether the experiment is successfully measuring the expected signals from the sample, how strong is the contribution from the instrument background and sample container, and if there are any unusual observations that require more focused measurements. All these can be greatly facilitated with the predicted spectra at hand, and the capability to perform on-demand predictions with additional/modified structure models. Rapid prediction of $S(Q,E)$ will also allow users to disentangle phonons and magnons in the measured total spectra in real-time. This has been a long-standing challenge of signal separation, given the dual neutron-matter interactions (nuclear interactions and spin interactions). It is worth mentioning that after the training stage (which the users do not have to do as the trained models will be provided on the data analysis servers), all these predictions can be achieved on a personal computer almost instantaneously by a generic user even with no prior modeling expertise. This is in stark contrast to the conventional approach of running DFT simulations on a computer cluster or supercomputer, which could take significant computational cost and expertise to perform. The prediction accuracy is close to *ab initio* results within acceptable error, and will be further improved as we include more data in the training database.

The workflow and methodology are not limited to neutron spectra prediction but have a few immediate and broader implications. First, calculating the phonon contribution of the specific heat is a natural next step in using the predicted PDOS. This enables the selection of materials with low specific-heat for thermal switch and high specific-heat for thermal storage materials. Additionally, several other straightforward extensions such as inelastic x-ray scattering, Raman scattering, infrared spectroscopy, and electron energy loss spectroscopy can be performed in a similar fashion, aiming to connect crystal structure with multi-dimensional experimental data or properties. Third, since vibrational spectroscopy like INS is very sensitive to local vibrational modes associated with defects, the work, which allows fast prediction and analysis of the spectra for a large variety of defective structure models, also offers an alternative avenue to

perform non-contact, non-destructive defect characterization and quantification, particularly for point defects that can directly create new local vibrational nodes.

4. Methods

4.1. Generation of INS database

The phonon database at Kyoto university, created by Atsushi Togo, contains DFT calculated force constants for over 10 000 inorganic crystals. The 2018-04-17 version [8] database was used to calculate the INS spectra in this work. Phonopy [11] was used to calculate the eigenfrequencies and polarization vectors on a Γ -centered mesh with spacing smaller than 0.1 \AA^{-1} . The $T = 0 \text{ K}$ INS spectra were then calculated with OCLIMAX [12]. The 1D $S(E)$ spectra were calculated for the VISION spectrometer using its instrument parameters [13]. The 2D $S(Q,E)$ spectra were calculated in the energy range of 0–150 meV (bin size 0.5 meV), Q range of 0–15 \AA^{-1} (bin size 0.05 \AA^{-1}), using an energy resolution of 1.5 meV. The calculations were performed on a computer cluster with 1600 CPU cores. It took about a day of computing time in total (this implies that regenerating this database for a different instrument, sample environment, E/Q range, or temperature will not be an issue).

4.2. Euclidean neural network

A Euclidean neural network very similar to that used by Chen *et al* [5] was trained to perform the prediction of 1D $S(E)$ spectra and the latent space vector, respectively. In the former case, the 1D spectra calculated from OCLIMAX were processed to have an energy range of 10–1200 cm^{-1} ($1 \text{ cm}^{-1} = 0.124 \text{ meV}$), with a bin size of 10 cm^{-1} (i.e. 120 bins to represent the spectra). In the latter case, the 50-dimensional latent space vectors are the features to predict. A relatively large cutoff radius (6 \AA) was used here to include more neighbors around each center atom. This is to capture potential coherent effects in the $S(Q,E)$ which only becomes apparent when the unit cell size is small, the symmetry is high, and all elements in the crystal are strong coherent scatterers. The dataset was categorized as 80% for training, 10% for validation, and 10% for testing.

4.3. Autoencoder

A fully connected deep feed-forward neural network was used as the autoencoder. The seven hidden layers contain 90 000, 2500, 500, 50, 500, 2500, 90 000 nodes, respectively. Since the original $S(Q,E)$ has strong intensities in a small fraction of the image and a large dark area in the rest, we applied this cubic function $g(x) = 1 - (1 - x)^3$ for brightening. This function was chosen because it is invertible, maps $[0, 1] \rightarrow [0, 1]$, and is monotonous. The latent space vectors are normalized so that the maximum number in the vector is unity. The dataset was categorized as 5% for validation, and 1% for testing, and 94% for training.

4.4. DFT calculation of $\text{Ba}_3\text{ZnRu}_2\text{O}_9$

Spin-polarized DFT calculations of $\text{Ba}_3\text{ZnRu}_2\text{O}_9$ (interlayer antiferromagnetic configuration) were performed using the Vienna *Ab initio* Simulation Package (VASP) [18]. The calculation used projector augmented wave [19, 20] method to describe the effects of core electrons, and Perdew–Burke–Ernzerhof [21] implementation of the generalized gradient approximation for the exchange–correlation functional. Energy cutoff was 800 eV for the plane-wave basis of the valence electrons. The lattice parameters and atomic coordinates from literature [22] were used as the initial structure. The electronic structure was calculated on a $6 \times 6 \times 2$ Γ -centered mesh. The total energy tolerance for electronic energy minimization was 10^{-8} eV, and for structure optimization it was 10^{-7} eV. The maximum interatomic force after relaxation was below $0.001 \text{ eV \AA}^{-1}$. A (U – J) of 2.3 eV was applied to account for the localized 3d orbitals of Ru. A $3 \times 3 \times 1$ supercell was created for phonon calculations. The interatomic force constants were calculated by density functional perturbation theory as implemented in VASP, and the vibrational eigen-frequencies and modes were then calculated using Phonopy [11]. The OCLIMAX software [12] was used to convert the DFT-calculated phonon results to the simulated INS spectra.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI [23]: www.nature.com/articles/s41597-022-01926-x.

Acknowledgments

A portion of this research used resources at the Spallation Neutron Source, a DOE Office of Science User Facility operated by the Oak Ridge National Laboratory (ORNL). The computing resources for INS

simulations were made available through the VirtuES and the ICE-MAN projects, funded by Laboratory Directed Research and Development (LDRD) Program and Compute and Data Environment for Science (CADES) at ORNL. The machine learning research is sponsored by the Artificial Intelligence Initiative as part of the LDRD program of ORNL, managed by UT-Battelle, LLC, for the US Department of Energy under Contract DE-AC05-00OR22725. G W thanks support by the U S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internships Program (SULI). M Li acknowledges support from U S. Department of Energy, Office of Science, Basic Energy Sciences, Award No. DE-SC0021940.

ORCID iDs

Yongqiang Cheng  <https://orcid.org/0000-0002-3263-4812>

Matthew B Stone  <https://orcid.org/0000-0001-7884-9715>

Andrei T Savici  <https://orcid.org/0000-0001-5127-8967>

Mingda Li  <https://orcid.org/0000-0002-7055-6368>

Anibal J Ramirez-Cuesta  <https://orcid.org/0000-0003-1231-0068>

References

- [1] Doucet M, Samarakoon A M, Do C, Heller W T, Archibald R, Alan Tennant D, Proffen T and Granroth G E 2020 Machine learning for neutron scattering at ORNL *Mach. Learn.: Sci. Technol.* **2** 023001
- [2] Chen Z et al 2021 Machine learning on neutron and x-ray scattering and spectroscopies *Chem. Phys. Rev.* **2** 031301
- [3] Chang M-C, Tung C-H, Chang S-Y, Carrillo J M, Wang Y, Sumpter B G, Huang G-R, Do C and Chen W-R 2022 A machine learning inversion scheme for determining interaction from scattering *Commun. Phys.* **5** 46
- [4] Samarakoon A M et al 2020 Machine-learning-assisted insight into spin ice $\text{Dy}_2\text{Ti}_2\text{O}_7$ *Nat. Commun.* **11** 892
- [5] Chen Z, Andrejevic N, Smidt T, Ding Z, Xu Q, Chi Y-T, Nguyen Q T, Alatas A, Kong J and Li M 2021 Direct prediction of phonon density of states with Euclidean neural networks *Adv. Sci.* **8** 2004214
- [6] Thomas N, Smidt T, Kearnes S, Yang L, Li L, Kohlhoff K and Riley P 2018 Tensor field networks: rotation-and translation-equivariant neural networks for 3D point clouds (arXiv:1802.08219)
- [7] Geiger M and Smidt T 2022 e3nn: Euclidean neural networks (arXiv:2207.09453)
- [8] Togo A 2018 Phonon database at Kyoto University vol 2022 (available at: <http://phonondb.mtl.kyoto-u.ac.jp/>)
- [9] Jain A et al 2013 Commentary: the materials project: a materials genome approach to accelerating materials innovation *APL Mater.* **1** 011002
- [10] Ong S P, Richards W D, Jain A, Hautier G, Kocher M, Cholia S, Gunter D, Chevrier V L, Persson K A and Ceder G 2013 Python materials genomics (pymatgen): a robust, open-source python library for materials analysis *Comput. Mater. Sci.* **68** 314
- [11] Togo A and Tanaka I 2015 First principles phonon calculations in materials science *Scr. Mater.* **108** 1
- [12] Cheng Y Q, Daemen L L, Kolesnikov A I and Ramirez-Cuesta A J 2019 Simulation of inelastic neutron scattering spectra using OCLIMAX J. *Chem. Theory Comput.* **15** 1974
- [13] Seeger P A, Daemen L L and Larese J Z 2009 Resolution of VISION, a crystal-analyzer spectrometer *Nucl. Instrum. Methods Phys. Res. A* **604** 719
- [14] Samarakoon A M, Laurell P, Balz C, Banerjee A, Lampen-Kelley P, Mandrus D, Nagler S E, Okamoto S and Tennant D A 2022 Extraction of interaction parameters for α - RuCl_3 from neutron data using machine learning *Phys. Rev. Res.* **4** L022061
- [15] Samarakoon A, Tennant D A, Ye F, Zhang Q and Grigera S 2021 Integration of machine learning with neutron scattering: Hamiltonian tuning in spin ice with pressure (arXiv:2110.15817)
- [16] Stone M B et al 2014 A comparison of four direct geometry time-of-flight spectrometers at the spallation neutron source *Rev. Sci. Instrum.* **85** 045113
- [17] Cai L, Shi Y, Hrdina K, Moore L, Wu J, Daemen L L and Cheng Y 2018 Low-temperature vibrational dynamics of fused silica and binary silicate glasses *Phys. Rev. B* **97** 054311
- [18] Kresse G and Furthmüller J 1996 Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set *Phys. Rev. B* **54** 11169
- [19] Blöchl P E 1994 Projector augmented-wave method *Phys. Rev. B* **50** 17953
- [20] Kresse G and Joubert D 1999 From ultrasoft pseudopotentials to the projector augmented-wave method *Phys. Rev. B* **59** 1758
- [21] Perdew J P, Burke K and Ernzerhof M 1996 Generalized gradient approximation made simple *Phys. Rev. Lett.* **77** 3865
- [22] Beran P, Ivanov S A, Nordblad P, Middey S, Nag A, Sarma D D, Ray S and Mathieu R 2015 Neutron powder diffraction study of $\text{Ba}_3\text{ZnRu}_{2-x}\text{Ir}_x\text{O}_9$ ($x = 0, 1, 2$) with 6H-type perovskite structure *Solid State Sci.* **50** 58
- [23] Cheng Y, Stone M B and Ramirez-Cuesta A J 2023 A database of synthetic inelastic neutron scattering spectra from molecules and crystals *Sci. Data* **10** 54