

# Predicting neutron experiments from first principles: A workflow powered by machine learning

Eric Lindgren,<sup>1</sup> Adam J. Jackson,<sup>2</sup> Erik Fransson,<sup>1</sup> Esmée Berger,<sup>1</sup> Goran Škoro,<sup>3</sup>  
Svemir Rudić,<sup>3</sup> Rastislav Turanyi,<sup>2</sup> Sanghamitra Mukhopadhyay,<sup>3</sup> and Paul Erhart<sup>1,\*</sup>

<sup>1</sup>*Department of Physics, Chalmers University of Technology, SE-41296, Gothenburg, Sweden*

<sup>2</sup>*Scientific Computing Department, STFC Rutherford Appleton Laboratory, Didcot OX11 0QX, UK*

<sup>3</sup>*ISIS Neutron and Muon Source, STFC Rutherford Appleton Laboratory, Didcot OX11 0QX, UK*

Machine learning has emerged as a powerful tool in materials discovery, enabling the rapid design of novel materials with tailored properties for countless applications, including in the context of energy and sustainability. To ensure the reliability of these methods, however, rigorous validation against experimental data is essential. Scattering techniques—using neutrons, X-rays, or electrons—offer a direct way to probe atomic-scale structure and dynamics, making them ideal for this purpose. In this work, we describe a computational workflow that bridges machine learning-based simulations with experimental validation. The workflow combines density functional theory, machine-learned interatomic potentials, molecular dynamics, and autocorrelation function analysis to simulate experimental signatures, with a focus on inelastic neutron scattering. We demonstrate the approach on three representative systems: crystalline silicon, crystalline benzene, and hydrogenated scandium-doped BaTiO<sub>3</sub>, comparing the simulated spectra to measurements from four different neutron spectrometers. While our primary focus is inelastic neutron scattering, the workflow is readily extendable to other modalities, including diffraction and quasi-elastic scattering of neutrons, X-rays, and electrons. The good agreement between simulated and experimental results highlights the potential of this approach for guiding and interpreting experiments, while also pointing out areas for further improvement.

## I. INTRODUCTION

Advancements in materials science are pivotal for technological progress, driving innovations in energy storage, electronics, and catalysis. Computational methodologies, particularly density functional theory (DFT), have become essential tools in materials discovery by predicting materials properties and guiding experimental efforts [1–5]. The integration of machine learning (ML) with these computational techniques has further accelerated the discovery of novel materials, e.g., by enabling rapid screening of vast chemical spaces [3, 4, 6–9]. ML has also facilitated the development of machine-learned interatomic potentials (MLIPs), which allow accurate and efficient atomic-scale simulations, bridging the gap between empirical potentials and first-principles methods [10–12].

However, the predictive power of these computational approaches necessitates rigorous experimental validation. Scattering experiments, such as neutron, X-ray, and electron scattering, provide critical insights into the structure and dynamics of materials but require precise simulations to interpret the data accurately [13]. Bridging the gap between computational predictions and experimental observations remains a significant challenge in the field. Predictive simulations could also significantly enhance experimental planning and execution by ensuring that data acquisition is optimized for maximum information gain while reducing the likelihood of inconclusive

or ambiguous results [14–16]. Furthermore, such simulations can support the preparation of beamline proposals, providing quantitative justifications for instrument time requests by demonstrating expected signal strengths and resolving power. As experimental facilities increasingly integrate computational tools into their workflows, predictive capabilities are poised to play a crucial role in streamlining the experimental process, improving the overall efficiency of materials characterization, and ultimately accelerating scientific discoveries.

In response to these challenges, we here describe a comprehensive workflow that integrates DFT calculations, MLIPs in the neuroevolution potential (NEP) format, molecular dynamics (MD) simulations using GPUMD, [17] the computation of autocorrelation functions via DYNASOR [18, 19], and their convolution with atomic form factors, instrument resolution functions and kinematic constraints. This enables instrument-specific predictions of scattering data from first-principles, allowing direct comparisons between simulations and experimental measurements.

We demonstrate the efficacy of this workflow by applying it to three example systems, including elemental Si, crystalline benzene, and hydrogenated Sc-doped BaTiO<sub>3</sub>, showcasing both its potential for guiding experimental design and accelerating the discovery of new materials as well as its current limitation. We focus specifically on simulating inelastic neutron scattering (INS) experiments, but the general workflow can be easily used to simulate other experimental modalities, including diffraction as well as quasi-elastic and inelastic scattering of neutrons, X-rays, and electrons.

---

\* erhart@chalmers.se

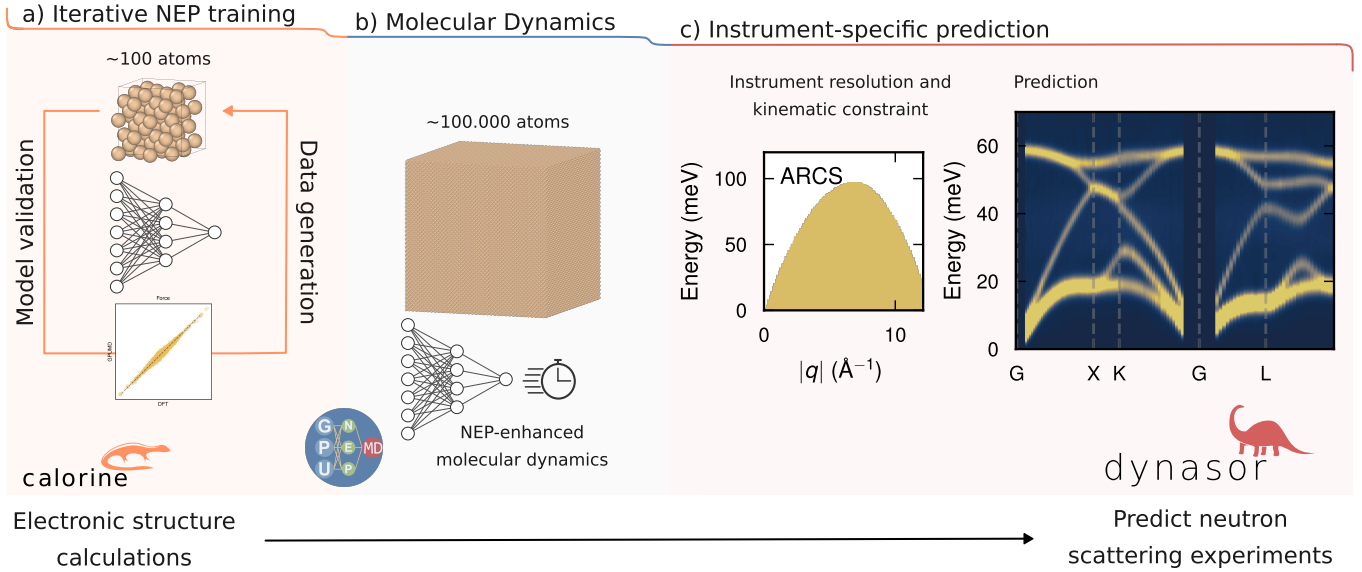


FIG. 1. Workflow for simulating neutron scattering experiments from first principles. (a) The first step of the workflow comprises constructing machine-learned interatomic potentials (MLIPs) using an iterative cycle combining both data generation and model validation. Here, the training is facilitated by the GPUMD and CALORINE packages. (b) The final MLIP is used in the second step to run large-scale MD simulations using the GPUMD package. (c) The dynamic structure factor is computed from the MD trajectories using the DYNASOR package. The dynamic structure factor is weighted by species-dependent scattering lengths, and broadened with an instrument-specific resolution function in order to predict the outcome of a particular neutron scattering experiment.

## II. METHODS

The workflow that we demonstrate for predicting neutron scattering experiments from first principles consists of three steps. The first step is the construction of MLIPs based on the NEP framework (Fig. 1a; Sect. II A). These accurate and efficient MLIPs enable the second step of the workflow, which are large-scale MD simulations (Fig. 1b; Sect. II B). The MD trajectories that result from the second step are then used in the third step of the workflow, in which we compute the dynamic structure factor that is then weighted by neutron scattering lengths as well as the instrument resolution function and kinematic constraint (Fig. 1c; Sect. II C). The weighted dynamic structure factor can then be compared directly to experimental data.

### A. Construction of the machine-learned potentials

The workflow starts with training of a MLIP, or using an already trained model. Here, we used three different MLIPs based on the NEP framework [17, 20, 21], one for each of the systems studied in this work. For Si, the model published in [17] was used, while we constructed new models for crystalline benzene and hydrogenated Sc-doped  $\text{BaTiO}_3$  ( $\text{BaTi}_{1-x}\text{Sc}_x\text{O}_3\text{H}_x$ ) using the iterative procedure described in Ref. 22 utilizing the GPUMD [17] and CALORINE packages [23] (Fig. 1a).

The training set was initially composed of strained and scaled structures, based on ideal structures using reference data from DFT calculations (Sect. II D). In the case of benzene, the initial dataset also included dimer configurations to ensure that intermolecular interactions are correctly captured. An initial model was trained on all available data. The dataset was then augmented with structures from several iterations of active learning. To this end, we trained an ensemble of five models by randomly splitting the data into training and validation sets, which was subsequently used to estimate the model uncertainty. MD simulations were then carried out between 10 and 200 K and at pressures ranging from 0 to 10 GPa for benzene, and from 300 to 2000 K and at pressures ranging from  $-1$  to 10 GPa for Sc-doped  $\text{BaTiO}_3$  using the respective current generation of NEP models. The ensemble was used to select structures with a high prediction uncertainty, quantified by a range of predictions over the ensemble, for which we computed reference energies, forces, and stresses via DFT. These configurations were subsequently included when training the next-generation NEP model. In total the training set for benzene consisted of 798 structures, corresponding to a total of 94 470 atoms. For Sc-doped  $\text{BaTiO}_3$ , a total of 2280 structures, corresponding to a total of 138 438 atoms. Structures were generated and manipulated using the ASE [24] and HIPHIVE packages [25].

We obtained the final NEP models after 13 iterations for benzene and 6 iterations for Sc-doped  $\text{BaTiO}_3$ . The

final models were trained on all available data. For the benzene model, the resulting average root mean square errors (RMSEs) over the ensemble are 1.1(8) meV/atom for the energies and 63(30) meV/Å for the forces, and 8.4(16) meV/atom for the virials. The corresponding average coefficients of determination on the same folds are  $R^2 = 0.9997(3)$ ,  $R^2 = 0.9949(51)$ , and  $R^2 = 0.9988(6)$  for energies, forces, and virials, respectively. The RMSEs and  $R^2$  scores for the final benzene model were 8.510 meV/atom and  $R^2 = 0.9998$  for the energies, 59 meV/Å and  $R^2 = 0.9962$  for the forces, and 9.3 meV/atom and  $R^2 = 0.9987$  for the virials (Sect. S1). For the Sc-doped BaTiO<sub>3</sub> model, the ensemble RMSEs were 6.6(13) meV/atom for the energies, 186(34) meV/Å for the forces, and 31(5) meV/atom for the virials. The respective coefficients of correlation ( $R^2$ ) were  $R^2 = 0.99998(1)$ ,  $R^2 = 0.9765(58)$ , and  $R^2 = 0.9964(9)$  for energies, forces, and virials, respectively. The RMSEs and  $R^2$  scores for the final Sc-doped BaTiO<sub>3</sub> model were 6.2 meV/atom and  $R^2 = 0.9999$  for the energies, 172 meV/Å and  $R^2 = 0.9792$  for the forces, and 30 meV/atom and  $R^2 = 0.9963$  for the virials (Sect. S2).

The resulting NEP models along with the reference data used for training are available via zenodo as specified in the Data Availability statement.

## B. Molecular dynamics

In the second step of the workflow we perform MD simulations using the MLIPs from the first step for large supercells (Fig. 1b). The resulting MD trajectories are later used to compute the dynamic structure factor as detailed in the next section.

For Si, a supercell comprising  $38 \times 38 \times 38$  primitive cells for a total of 438 976 atoms was simulated at 300 K, 900 K, 1200 K, and 1500 K, with equilibration of the system in the NPT ensemble and production for 1 ns in the NVE ensemble, with a timestep of 2 fs. The atomic positions were written to file every 14 fs in order to accurately resolve the fastest vibrations in the system when computing the dynamic structure factor.

Crystalline benzene was simulated in a supercell containing a total of 57 024 atoms. The benzene system was equilibrated in the path-integral molecular dynamics (PIMD) ensemble [26, 27] to avoid the significant underestimation of the cell volume in the classical NPT ensemble at low temperatures. The MD simulations were conducted at 127 K to strike a balance between computational cost and the number of PIMD beads (see Sect. S3 and Sect. S4 in the Supplementary Information). Production runs were then performed for 1 ns in the NVE ensemble. A timestep of 0.5 fs was used, and the positions were written every 3 fs. Ten independent MD runs were performed to improve the statistics of the computed dynamic structure factor.

Hydrogenated supercells of Sc-doped BaTiO<sub>3</sub> were constructed for various Sc concentrations in the range

16% to 70% in both the cubic and hexagonal phase. The supercell contained  $\simeq 40\,000$  atoms. Equilibration was performed in the PIMD ensemble at a temperature of 15 K, and production was carried out for 350 ps in the thermostated ring-polymer MD ensemble [28] with a timestep of 0.5 fs. This approach captures nuclear quantum effects on the frequencies [27], but it should be noted that the phonon occupation statistics are still classical. Both equilibration and production runs used 32 PIMD beads, for an effective system size of  $\simeq 1\,300\,000$  atoms, limiting the length of the production run compared to Si and benzene because of the increased computational cost.

## C. Auto-correlation functions and instrument-specific kinematic constraints

The central quantity analyzed in the third step of the workflow is the dynamic structure factor,  $S(\mathbf{q}, \omega)$ .  $S(\mathbf{q}, \omega)$  is directly proportional to the intensity measured in scattering experiments, and can be readily extracted from MD simulations. While the procedure has been described in detail in [18, 19], we briefly summarize it here for completeness. Let  $n(\mathbf{r}, t)$  denote the particle density defined as

$$n(\mathbf{r}, t) = \sum_i^N \delta(\mathbf{r} - \mathbf{r}_i(t)). \quad (1)$$

$\mathbf{r}_i(t)$  is the position of particle  $i$  at time  $t$ , and  $N$  is the total number of particles. The particle density can now be Fourier transformed in space,

$$n(\mathbf{q}, t) = \int_{-\infty}^{\infty} \sum_i^N \delta(\mathbf{r} - \mathbf{r}_i(t)) e^{i\mathbf{q} \cdot \mathbf{r}} d\mathbf{r} = \sum_i^N e^{i\mathbf{q} \cdot \mathbf{r}_i(t)}, \quad (2)$$

with the autocorrelation function of  $n(\mathbf{q}, t)$  yielding the intermediate scattering function  $F(\mathbf{q}, t)$ ,

$$F(\mathbf{q}, t) = \frac{1}{N} \langle n(\mathbf{q}, t) n(-\mathbf{q}, 0) \rangle, \quad (3)$$

where the brackets denote an ensemble average. The intermediate scattering function can then be Fourier transformed in time to yield the dynamic structure factor,

$$S(\mathbf{q}, \omega) = \int_{-\infty}^{\infty} F(\mathbf{q}, t) e^{-i\omega t} dt. \quad (4)$$

The dynamic structure factor in Eq. (4) can be further generalized for multi-component systems. Different atomic nuclei scatter neutrons, X-rays, and electrons with varying intensity, which can be taken into account by weighting the partial dynamic structure factor for species  $\alpha$  and  $\beta$  accordingly. In the case of neutrons,

the partial dynamic structure factor should be weighted by the scattering lengths,  $b_\alpha$  and  $b_\beta$ ,

$$S(\mathbf{q}, \omega) = \sum_{\alpha} \sum_{\beta} b_{\alpha} b_{\beta} S_{\alpha\beta}(\mathbf{q}, \omega). \quad (5)$$

The dynamic structure factor in Eq. (5) was computed from the MD trajectories using the DYNASOR package [18, 19] in the third step of the workflow (Fig. 1c).  $\mathbf{q}$ -points and time lags were selected to match the accessible range of the simulated neutron scattering instruments. Specifically, for Si a Brillouin zone path was sampled connecting the high-symmetry points  $\Gamma$ , X, K, and L. The path was sampled in 52 different Brillouin zones, randomly selected from the first zone up to  $|\mathbf{q}| = 12 \text{ \AA}^{-1}$  for a total of 6136  $\mathbf{q}$ -points. Randomly selected  $\mathbf{q}$ -points up to a magnitude  $|\mathbf{q}| = 14 \text{ \AA}^{-1}$  and  $|\mathbf{q}| = 18 \text{ \AA}^{-1}$  were sampled for benzene and Sc-doped  $\text{BaTiO}_3$ , respectively, yielding 2116 and 2601  $\mathbf{q}$ -points, respectively. Gaussian broadening with a width of  $0.01 \text{ \AA}^{-1}$  was then applied to each  $\mathbf{q}$ -point, followed by averaging over spherical shells in  $|\mathbf{q}|$  to produce  $S(q, \omega)$ .

Instrument-specific resolution functions and kinematic constraints were applied to the calculated spectra using the EUPHONIC package [29] with the resolution functions defined in the RESINS package [30]. The resolution functions used here are Gaussians with energy-dependent width; the functions for TOSCA and Lagrange are based on implementations in ABINS, and the functions for MAPS and ARCS are based on PYCHOP [31]. (The instrument functions for both ABINS and PYCHOP are distributed in MANTID [32].) Note that the true resolution functions are four-dimensional and non-Gaussian, but these 1-D approximations are used routinely in INS simulations. The kinematic constraints have their origin in the instrument geometry and transformation from time-of-flight measurements to  $(q, \omega)$  space. In the simulations they are applied as a mask to data computed directly in the  $(q, \omega)$  space.

Finally, a quantum correction factor was applied to all dynamic structure factors, in order to correct for the classical phonon statistics generated by the MD simulations. Specifically, we applied the following correction factor based on first-order Stokes-Raman scattering [33, 34],

$$S(q, \omega)_{\text{corrected}} = S(q, \omega) \frac{\beta \hbar \omega}{1 - \exp(-\beta \hbar \omega)}. \quad (6)$$

#### D. Density functional theory calculations

To generate reference data for the construction of the MLIPs (Sect. II A) we performed non-spin polarized DFT calculations using the projector augmented wave method [35, 36] as implemented in the Vienna ab-initio simulation package [37–39] with a plane wave energy cutoff of 520 eV using the vdW-DF-cx exchange correlation functional [40] for benzene and the r2SCAN functional [41] for Sc-doped  $\text{BaTiO}_3$ . The Brillouin zone was sampled with

automatically generated  $\Gamma$ -centered  $\mathbf{k}$ -point grids with an approximate spacing of  $0.25/\text{\AA}$  and the partial occupancies in each orbital were set using Gaussian smearing with a width of 0.1 eV. The DFT data are available via zenodo as specified in the Data Availability statement.

#### E. Inelastic neutron scattering experiments on crystalline benzene

For validation of the predictions for crystalline benzene, inelastic neutron scattering experiments were performed at the TOSCA neutron spectrometer [42, 43] at the ISIS Neutron and Muon Source. The liquid sample was placed in a 1 mm thick standard flat TOSCA aluminum cell which was then briefly submerged into liquid nitrogen. As soon as the sample solidified it was quickly transferred into the TOSCA closed cycle refrigerator and allowed to further cool to the cryostat base temperature below 10 K. The short INS measurements (approximately 8 minutes per spectrum, i.e., total exposure of  $20 \mu\text{A h}$ ) were performed as part of a cooling run at a rate of 3 K/min, with the initial spectrum taken at a starting temperature of 127 K and followed by other measurements at a starting temperature of 103 K, 75 K, 46 K, and 24 K. The longer INS measurement (approximately 2 hours, i.e., total exposure of  $285 \mu\text{A h}$ ) was performed at the base temperature of 10 K, giving a superior spectral signal-to-noise ratio. The raw data, i.e., time-of-flight events, were reduced using MANTID [32].

#### F. Post-processing and plotting

The NEP models and calculated correlation functions were post-processed and analyzed using PYTHON scripts, utilizing the NUMPY [44], PANDAS [45, 46], and SCIPY [47] packages. Plots were generated using MATPLOTLIB [48], with color maps from PERFECT-CMAPS [49]. Atomic structures were visualized and analyzed using OVITO [50].

### III. RESULTS

#### A. Anharmonicity in Si

We begin by applying the workflow outlined in the methodology section to simulate an INS experiment on crystalline Si at 300 K reported in Ref. 51, that was carried out at the ARCS wide range angular spectrometer (BL-18) at the spallation neutron source (SNS) at Oak Ridge National Laboratory (Fig. 2a). The simulation is made instrument-specific by applying the resolution function and kinematic constraint for ARCS to the simulated dynamic structure factor. The  $q$  and energy range measured by ARCS is relatively broad (Fig. 3), and thus

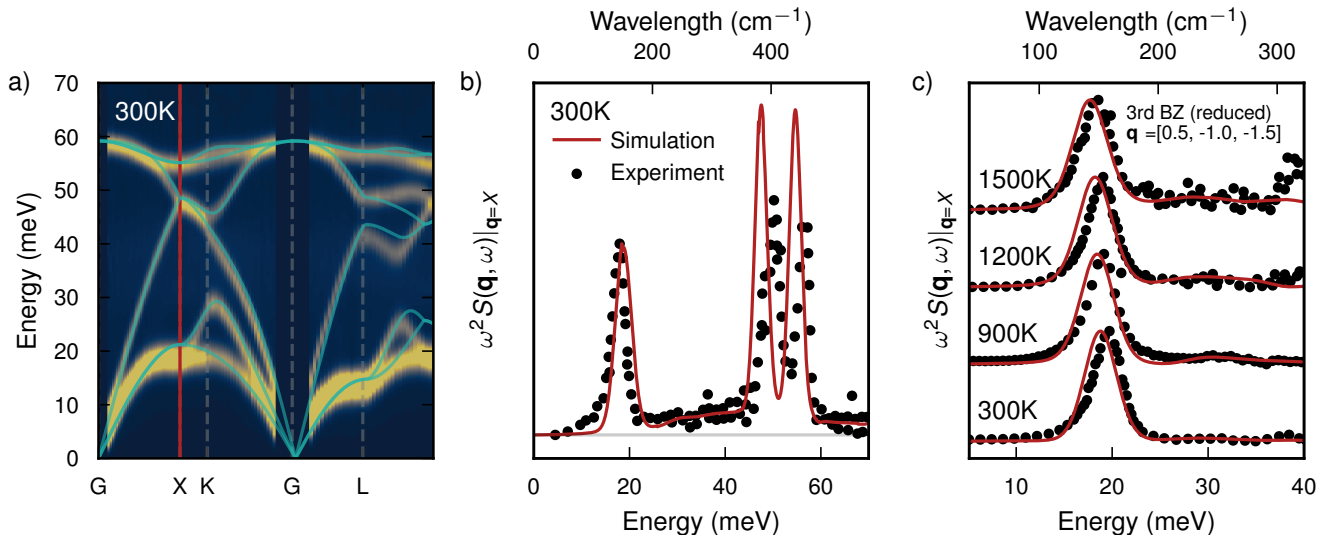


FIG. 2. (a) Simulated INS dispersion of Si from MD for the ARCS spectrometer at spallation neutron source, with the harmonic phonon dispersion calculated using the NEP model overlaid (turquoise lines). (b) The simulated intensity at the X-point, as well as (c) the first peak in the intensity at the X-point in the third Brillouin zone for 300 K, 900 K, 1200 K, and 1500 K (red lines). Note that the dispersion in (a) and intensity in (b) is aggregated over 52 different Brillouin zones, while in (c) the results for only a single Brillouin zone are shown. The experimental data (black circles) is from [51]. The MD simulation captures both the anharmonicity and multi-phonon effects present in the experimental data, as well as the mode softening as the temperature is increased. The multi-phonon effects manifest as non-zero intensity between the acoustic and optical branches at the X-point.

we computed the dynamic structure in multiple different Brillouin zones to accurately sample the full range allowed by the instrument. In total, 52 Brillouin zones were sampled. The dynamic structure factors  $S(\mathbf{q}, \omega)$  were then weighted by a factor  $\propto 1/|\mathbf{q}|^2$  since to first order the scattered intensity grows as  $|\mathbf{q}|^2$ . Furthermore, the Debye-Waller factor,  $\exp(-q^2 U/3)$ , was corrected for in each of the Brillouin zones before the zones were averaged together. This expression for the Debye-Waller factor assumes an isotropic displacement of the atoms in all Cartesian directions with  $U = \langle u^2 \rangle$  being the mean squared displacement in the system [52].  $U$  was estimated to be  $0.013 \text{ \AA}^2$  from a 100 ps MD simulation of the Si system at 300 K, otherwise following the same protocol as the other simulations of Si in this work.

We validate our results by comparing them to the experimental data measured on the ARCS spectrometer in Ref. 51. Specifically, we consider the intensity at the high-symmetry X-point (Fig. 2b), where the simulated intensity has been multiplied by an extra factor of  $\omega^2$ . Our results are in quantitative agreement with experiments, with the centroids of the phonon mode peaks agreeing well. The relative intensity between the different phonon mode peaks in the experiment are not entirely reproduced in the simulation, which could be due to a missing correction factor or experimental variability.

We note, in particular, the nonzero intensity measured in the experiment and captured by the simulation in the region 30 meV to 50 meV. This scattered

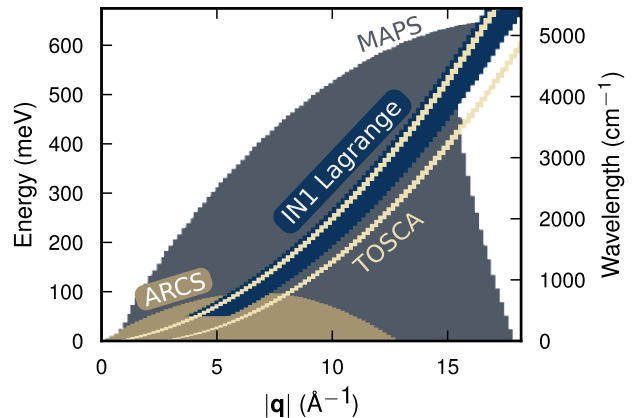


FIG. 3. Kinematic constraints for the four neutron instruments simulated in this work. ARCS (BL-18) is a wide-range spectrometer at the spallation neutron source (SNS) at the Oak Ridge National Laboratory (USA). TOSCA and MAPS are spectrometers at the ISIS Muon and Neutron Source (UK). Finally, IN1 Lagrange is a spectrometer at the *Institute Laue-Langevin* in Grenoble (France). Note that all spectrometers differ in  $q$  and energy range and resolution, owing to their respective kinematic constraints and resolution functions.

intensity corresponds to multi-phonon effects, which are inherently captured by MD simulations. Furthermore, the effect of thermal expansion as the temperature is

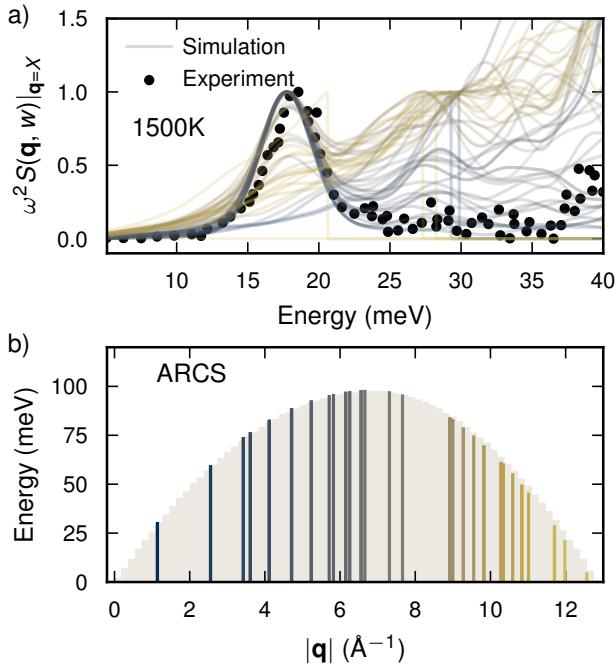


FIG. 4. (a) Simulated intensity at the X-point at 1500 K in crystalline Si for each of the 52 Brillouin zones accessible by the ARCS spectrometer. Note that the intensity of the multi-phonon shoulder at 30 meV varies greatly with the Brillouin zone. (b) The kinematic constraint for the ARCS spectrometer at SNS, which limits the range of non-zero intensities in (a) depending on the  $|q|$  for the X-point in each Brillouin zone.

varied is also directly included by the MD simulations, where specifically the low-energy mode at the X-point is softened as temperature is increased (Fig. 2c). One can observe that some of the predicted mode energies are slightly shifted compared to experiments, by approximately 1 meV. Given that the MLIP accurately reproduces the harmonic phonon dispersion from DFT, it is most likely due to the underlying exchange-correlation functional.

At 1500 K it moreover appears that in the region around 40 meV the simulated and experimental *intensities* differ. This discrepancy could be due to the Brillouin zone in which the simulations have been conducted. Here, we show the X-point in the third Brillouin zone, i.e.,  $\mathbf{q} = [0.5, -1.0, -1.5]$ , while in the experimental reference [51] the exact  $\mathbf{q}$ -point is not specified. In fact, the intensity at the X-point varies substantially depending on the Brillouin zone, especially the intensity of the multi-phonon shoulder around 30 meV (Fig. 4). For the comparison shown in (Fig. 2c), we selected the Brillouin zone for which the simulated spectrum best reproduces the experimental data, based on the mean-squared error calculated over the spectrum.

In MD, the dynamics of the system described by the potential model is captured at the classical level, in-

cluding high-order phonon effects, thermal expansion, and full anharmonicity. Efforts have been made in recent years to include the effects of anharmonicity on top of harmonic models, including but not limited to using higher-order force constants [25], temperature-dependent effective potentials [53], anharmonic lattice models [54], and the self-consistent harmonic approximation [55]. However, a harmonic model is inherently limited in describing such complicated dynamic events.

## B. Corrections in crystalline benzene

We now turn to simulating an INS measurement of crystalline benzene at 127 K at the TOSCA spectrometer at the ISIS Neutron and Muon Source (UK) in order to study the effects of the resolution function and quantum correction in more detail (Fig. 5a).

The scattered intensity from benzene is dominated by incoherent scattering from hydrogen, owing to the exceptionally large incoherent scattering length of hydrogen. We thus study the INS spectrum directly. The dynamic structure factor can be integrated over  $|q|$  in order to obtain  $S(\omega) = \int S(q, \omega) dq$ . Comparing the raw simulated spectrum with the experimental data, we find that the simulated spectrum captures the peaks corresponding to different modes but the relative intensity between them is not reproduced. Furthermore, the low-energy peak at 10 meV is not captured. The reason for these discrepancies is to a large extent due to the kinematic constraint and resolution function of the TOSCA spectrometer (Fig. 3). The two detector banks of TOSCA map out two lines in  $q$ - $\omega$  space, where high (low) frequencies correspond to large (low)  $q$ . By sampling along these  $q$ - $\omega$  lines and convoluting the resulting spectrum with the resolution function of the instrument, the agreement improves notably.

However, the ratio in intensity between the high and low-energy regions is still not reproduced. The main reason for this discrepancy is the classical statistics of MD simulations, which we correct for with the quantum correction factor according to Eq. (6). Applying both kinematic constraint and quantum correction yields a simulated spectrum that is in near-quantitative agreement with experiments. The remaining difference to experiments is a redshift of the simulated spectrum by approximately 25 meV. We attribute this redshift to the DFT functional used to train the NEP model. A more detailed discussion comparing experiments and first-principles calculations to the predictions from the NEP model can be found in the Supplementary Information (Sect. S5).

We can further elucidate the simulated INS spectrum by comparing it to the phonon dispersion according to the underlying MLIP (Fig. 5b). The simulated INS spectrum differs notably from the phonon density of states in terms of intensity, owing to the kinematic constraint of the TOSCA spectrometer, and the quantum correction.



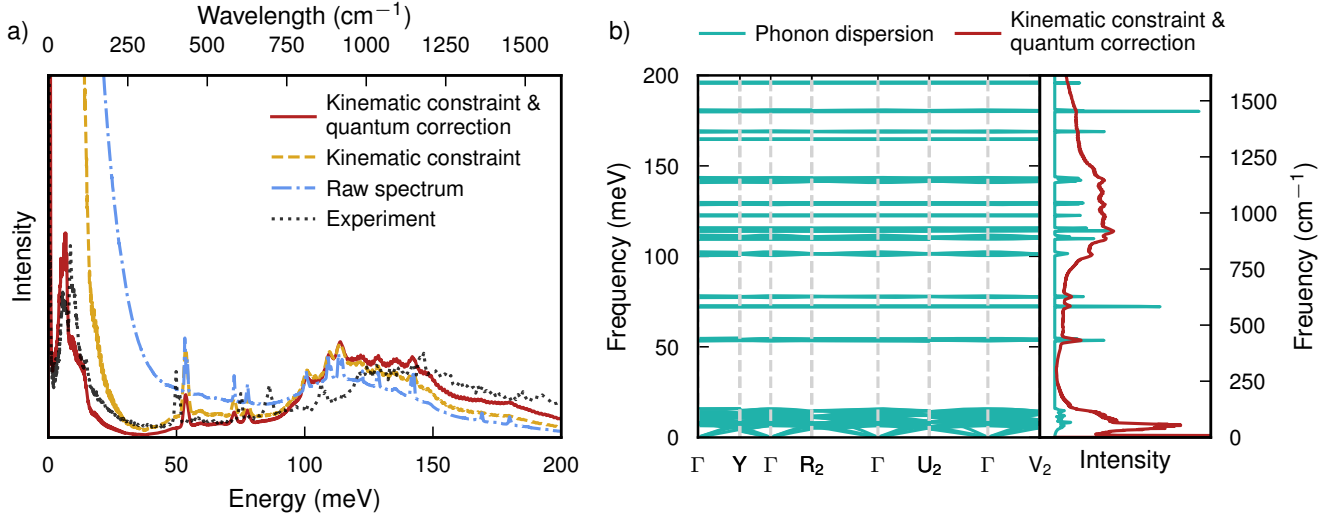


FIG. 5. (a) Simulated INS spectra for crystalline benzene, at increasing levels of refinement, compared to an experimental spectrum obtained at 127 K at the TOSCA spectrometer at the ISIS Neutron and Muon Source (UK). The first level of accuracy is the raw simulated spectrum from MD with only scattering lengths applied (raw spectrum). Correcting for the resolution function and kinematic constraint of the TOSCA spectrometer yields a marked increase in accuracy, with further improvement when additionally applying a quantum correction factor to compensate for the classical statistics in MD simulations, especially for low energies around 10 meV. The spectra have been individually scaled to match the experimental spectrum as closely as possible above 50 meV. (b) Phonon dispersion and density of states compared to the simulated INS spectrum at 127 K, corrected for the kinematic constraint of the TOSCA spectrometer and with quantum statistics.

Furthermore, the full anharmonicity included in the MD simulation in combination with the TOSCA resolution function yields a broadening of the peaks in the simulated INS spectrum.

In summary, this study of crystalline benzene highlights the importance of considering the resolution and kinematic constraints of the specific instrument, as well as correcting the statistics from classical MD simulations, when aiming for quantitative predictions of neutron scattering experiments.

### C. Hydrogen dynamics in hydrogenated Sc-doped BaTiO<sub>3</sub>

Finally, we turn to a more complicated system, in the form of hydrogenated Sc-doped BaTiO<sub>3</sub> (BaTi<sub>1-x</sub>Sc<sub>x</sub>O<sub>3</sub>H<sub>x</sub>) where  $x$  is the doping fraction of the tetravalent site (Ti, Sc). Perrichon *et al.* have performed a detailed INS study of the hydrogen dynamics in this system at three different spectrometers: the TOSCA and MAPS spectrometers at the ISIS Neutron and Muon source as well as IN1 Lagrange at the *Institut Laue-Langevin* [56]. The experiments were carried out at temperatures below 20 K. INS spectra were then obtained by averaging the dynamic structure factor  $S(q, \omega)$  up to a magnitude of  $|q| = 12/\text{\AA}$ . Our simulations presented in this section were averaged over  $q$  up to the limit of the kinematic constraint for MAPS,  $|q| = 18/\text{\AA}$ , in order to obtain better statistics.

Sc-doped BaTiO<sub>3</sub> undergoes a phase transition from a hexagonal structure to a cubic perovskite structure as the Sc concentration increases. On MD time scales, both structures are, however, at least metastable over the entire composition range, which (in contrast to experiment) allows us to sample structure and composition independently (Fig. 6).

The simulated spectra using our workflow and the experimental spectra agree well in the full energy range 0 to 500 meV. The peak at 125 meV corresponds to O–H vibrations, and is best described by the hexagonal phase for low Sc concentrations and by the cubic phase for high Sc-concentrations (Fig. 6a). However, the overtone peak at 250 meV is underestimated in both simulated phases. This discrepancy could be due to the quantum correction factor in Eq. (6), which is only valid for first-order scattering. This is further supported by the simulated spectra obtained using ABINS, which accurately capture the intensity of the 250 meV overtone peak. The latter method handles multi-phonon effects perturbatively and includes quantum effects but does not account for anharmonicity, which explains the sharper first-order features compared to the MD-based simulations.

The results from TOSCA highlight the 125 meV feature further (Fig. 6b). For low Sc concentrations the simulated spectrum using our workflow for the hexagonal structure agrees well with experiments, although the simulated spectrum is redshifted by approximately 25 meV. The simulated spectrum for the cubic structure agrees better with experiments as the Sc concentration

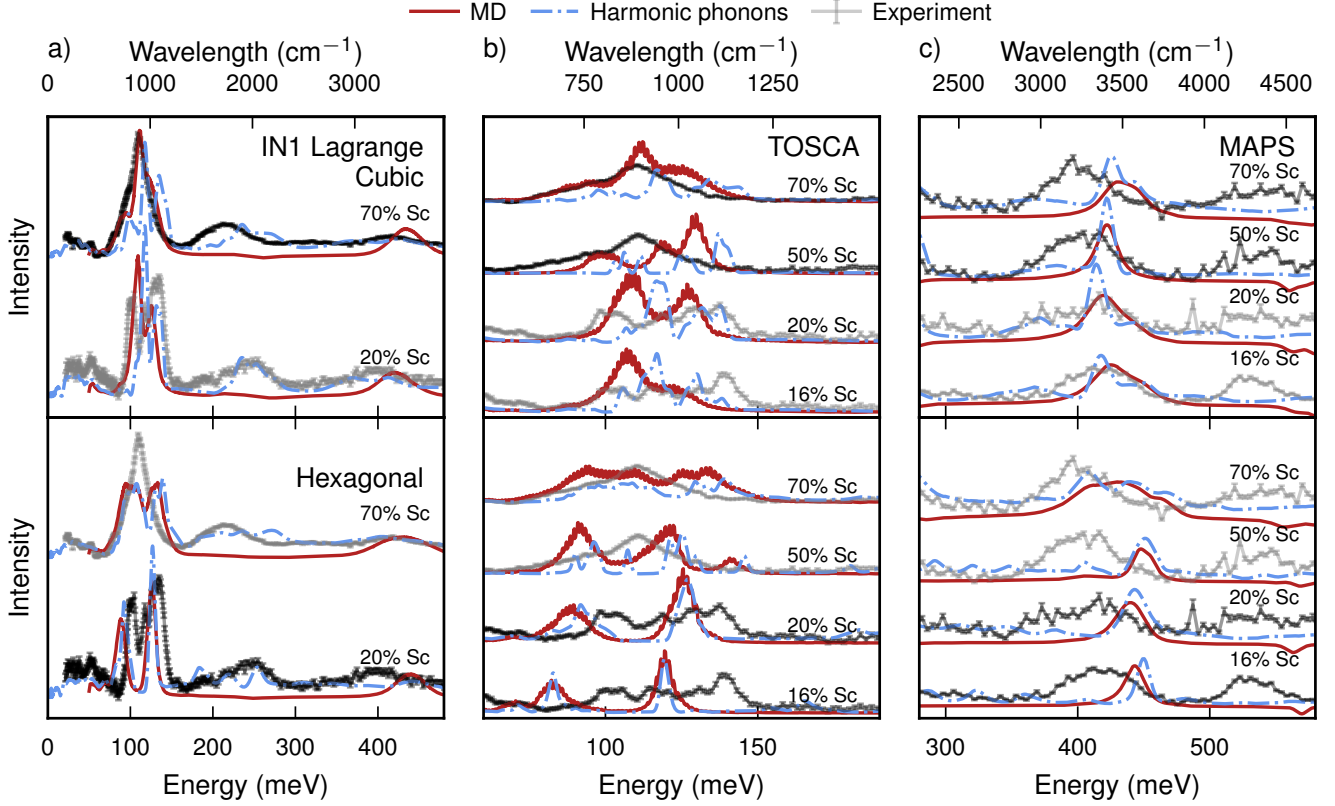


FIG. 6. Hydrogen dynamics in hydrogenated Sc-doped  $\text{BaTiO}_3$  ( $\text{BaTi}_{1-x}\text{Sc}_x\text{O}_3\text{H}_x$ ) for various concentrations of dopants, compared with experimental results measured at three different neutron spectrometers: IN1 Lagrange, TOSCA, and MAPS. Experimental data are from [56]. (a) IN1 Lagrange is a wide-range spectrometer that probes the dynamics in the region from 0 to 500 meV. (b) TOSCA and (c) MAPS on the other hand can be used to study the region around 100 meV and 500 meV, respectively, utilizing the higher energy resolution they offer. Simulated spectra using our workflow (labeled MD) corrected for quantum statistics and the kinematic constraint of the specific instrument are shown for all concentrations of dopants considered experimentally, with the simulated structure both in the cubic and hexagonal structure. Experimentally, only one of these phases is stable for a given concentration of Sc, but the energy of the two phases are sufficiently close that both phases are stable on the timescales of the MD simulations. The experimental data has thus been duplicated in the upper and lower rows of plots, where the compositions that are stable for each phase are indicated by the black lines. Additionally, simulated spectra based on harmonic phonons obtained via ABINS are included for comparison.

is increased, which is in line with the hexagonal to cubic phase transition with increasing Sc concentration. We can thus clearly distinguish the spectra for the two phases of Sc-doped  $\text{BaTiO}_3$ , as the Sc-doping is varied.

Finally, the MAPS spectrometer probes the high-energy region between 300 meV to 600 meV. The feature in the experimental spectra at 450 meV corresponds to stretching of the O–H bond according to Perrichon *et al.*, with the peak at 550 meV assigned as a combination mode of the O–H wag mode at 120 meV and the O–H stretch mode at 450 meV. The fundamental vibrational peak at 450 meV is captured by the simulations, although with a slight blueshift of 10 meV. However, the intensity for the combination mode is not reproduced by the simulation, neither using the MD-based workflow nor ABINS. In this case, we can further elucidate the nature of the combination modes at 550 meV

using ABINS (Fig. 7). In these harmonic incoherent-approximation simulations, the intensity in that region is mainly composed of fourth-order scattering events and above. Such high-order phonons require a higher-order correction factor in order for the statistics to come out correctly using the MD-based workflow. However, applying a higher-order correction factor is not straightforward, as one would have to know *a priori* in which region of the spectrum to apply the correction, and the order of the higher-order scattering process.

#### D. Discussion

The demonstrated workflows enable predictions of neutron scattering experiments, here in the form of INS spectra, from first principles. Starting from an atomic



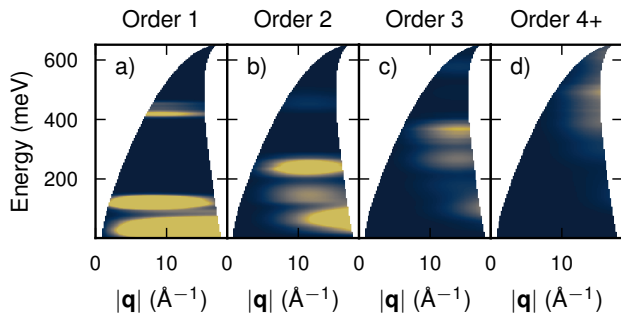


FIG. 7. Contribution to the total dynamic structure factor from different scattering orders for hydrogenated Sc-doped BaTiO<sub>3</sub> with 16% Sc in the cubic phase, simulated for the MAPS spectrometer using ABINS. (a) Fundamental modes with higher orders corresponding to overtones and combination modes in panels (b–d). The color scale is the same for all subpanels, which leads to clipping of the intensity in panel (a). Specifically note that the feature at 550 meV originates from fourth-order scattering processes and above.

representation of a material, we develop MLIPs using the NEP framework for performing accurate MD simulations for systems comprising at least tens of thousands of atoms over at least a few nanoseconds, from which the dynamic structure factor can be computed using the DYNASOR package. We have focused on crystalline materials in the examples above, but the workflow is directly applicable to disordered systems as well, including liquids as well as amorphous or biological materials. The predictions are made instrument-specific by applying resolution functions and kinematic constraints, and are additionally corrected to account for the classical statistics inherent to the MD simulations.

The predictions show remarkable quantitative agreement with experiments. Almost all experimental features in the form of vibrational peaks are faithfully reproduced in the predicted spectra, including their relative intensities after applying correction factors for the quantum statistics. Remaining differences between the spectra, such as systematic blue- or redshifts of certain peaks in the case of crystalline benzene, can be most likely attributed to the DFT functional used for training the MLIP. However, the intensity for some higher-order phonon processes, such as in Sc-doped BaTiO<sub>3</sub>, are not reproduced faithfully compared to experiments at the present level. In principle these discrepancies can, however, be corrected for by applying higher-order correction factors to recover the correct statistics for overtones and combination modes.

Such a correction would not represent a general-purpose *ab initio* approach as it is only possible in regions where these features can clearly be identified and separated, such that a single correction can be applied. For an unknown system, identifying such modes is problematic, although some information can be gained from harmonic calculations, e.g., using the ABINS package. However,

these corrections only affect the relative intensity of these peaks, the positions are directly obtained from the MD simulations. Applying higher-order quantum corrections is thus not strictly necessary in order to give a reasonable prediction of a neutron scattering experiment. In general, we suggest applying the lowest order correction for the whole spectrum to be sufficient for the purpose of guiding neutron scattering experiments.

The workflow as presented in this work relies on MLIPs in order to run accurate and efficient MD simulations. Classical force fields could be used but the results might be of limited accuracy, especially in systems involving both bonded and non-bonded interactions. However, training a MLIP or selecting an appropriate force field for a system of interest constitutes a bottleneck in the workflow, requiring domain knowledge and effort. Foundational models trained on large parts of the periodic table, such as MACE-MP-0 or CHGNet among others [57–61], offer an appealing alternative to creating bespoke MLIPs or using existing force fields. These foundational models can either be used out-of-the-box, or fine tuned with a small number of structures from DFT to yield an accurate model with comparatively low effort. It should, however, be noted that these models are computationally much more demanding than either NEP models or classical force fields.

#### IV. CONCLUSION

In this study, we have presented a workflow that enables predictions of neutron scattering experiments from first principles, by combining DFT calculations, MLIPs, autocorrelation functions from MD simulations as well as instrument resolution functions and kinematic constraints. We envision this workflow to be of great use in the context of materials discovery, offering an avenue for generating simulated experimental signatures for novel materials that can be directly compared to neutron, X-ray, and electron scattering experiments. ML in the form of MLIPs plays a central role, as the latter enable the accurate MD simulations and the extensive sampling that are the foundation of the present workflow. By integrating these components into a cohesive pipeline, our approach bridges the gap between theory and experiment, facilitating a more efficient feedback loop in the design and characterization of new materials. Ultimately, this workflow stands to accelerate materials analysis and discovery processes by providing high-fidelity, simulation-based insights that are directly aligned with experimental observables.

#### V. CONFLICTS OF INTEREST

There are no conflicts to declare.

## VI. DATA AVAILABILITY

The MLIP models, training data and databases of DFT calculations for benzene and hydrogenated Sc-doped BaTiO<sub>3</sub>, as well as the reduced INS data for crystalline benzene are available on zenodo at <https://doi.org/10.5281/zenodo.15283533>. The development of GPUMD package is hosted at <https://github.com/brucefan1983/GPUMD> and its documentation can be found at <https://gpumd.org>. The CALORINE package is hosted at <https://gitlab.com/materials-modeling/calorine>, its documentation is provided at <https://calorine.materialsmodeling.org>, and releases are available at <https://doi.org/10.5281/zenodo.7919206>. The DYNASOR package is hosted at <https://gitlab.com/materials-modeling/dynasor>, its documentation is provided at <https://dynasor.org>.

materialsmodeling.org, and releases are available at <https://doi.org/10.5281/zenodo.10012241>.

## VII. ACKNOWLEDGMENTS

We gratefully acknowledge funding from the Swedish Foundation for Strategic Research via the SwedNESS graduate school (GSn15-0008) and the Swedish Research Council (Nos. 2020-04935 and 2021-05072) as well as computational resources provided by the National Academic Infrastructure for Supercomputing in Sweden at NSC, PDC, and C3SE partially funded by the Swedish Research Council through grant agreement No. 2022-06725, as well as the Berzelius resource provided by the Knut and Alice Wallenberg Foundation at NSC. We are grateful to the Science and Technology Facilities Council (STFC), and to the ISIS Neutron and Muon Source in particular, for the provision of beamtime via TOSCA Xpress Access route (RB1990290).

- 
- [1] K. Shahzad, A. I. Mardare, and A. W. Hassel, Accelerating materials discovery: Combinatorial synthesis, high-throughput characterization, and computational advances, *Science and Technology of Advanced Materials: Methods* **4**, 2292486 (2024).
  - [2] C. Li, L. Bao, Y. Ji, Z. Tian, M. Cui, Y. Shi, Z. Zhao, and X. Wang, Combining machine learning and metal-organic frameworks research: Novel modeling, performance prediction, and materials discovery, *Coordination Chemistry Reviews* **514**, 215888 (2024).
  - [3] C. Chen, D. T. Nguyen, S. J. Lee, N. A. Baker, A. S. Karakoti, L. Lauw, C. Owen, K. T. Mueller, B. A. Bilodeau, V. Murugesan, and M. Troyer, Accelerating Computational Materials Discovery with Machine Learning and Cloud High-Performance Computing: From Large-Scale Screening to Experimental Validation, *Journal of the American Chemical Society* **146**, 20009 (2024).
  - [4] S. Fujii, J. Hyodo, K. Shitara, A. Kuwabara, S. Kasamatsu, and Y. Yamazaki, Emerging computational and machine learning methodologies for proton-conducting oxides: Materials discovery and fundamental understanding, *Science and Technology of Advanced Materials* **25**, 2416383 (2024).
  - [5] G. Hautier, C. C. Fischer, A. Jain, T. Mueller, and G. Ceder, Finding Nature’s Missing Ternary Oxide Compounds Using Machine Learning and Density Functional Theory, *Chemistry of Materials* **22**, 3762 (2010).
  - [6] G. Huang, F. Huang, and W. Dong, Machine learning in energy storage material discovery and performance prediction, *Chemical Engineering Journal* **492**, 152294 (2024).
  - [7] K. Guo, Z. Yang, C.-H. Yu, and M. J. Buehler, Artificial intelligence and machine learning in design of mechanical materials, *Materials Horizons* **8**, 1153 (2021).
  - [8] T. Lookman, P. V. Balachandran, D. Xue, and R. Yuan, Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design, *npj Computational Materials* **5**, 1 (2019).
  - [9] S. Wu, Y. Kondo, M.-a. Kakimoto, B. Yang, H. Yamada, I. Kuwajima, G. Lambard, K. Hongo, Y. Xu, J. Shiomi, C. Schick, J. Morikawa, and R. Yoshida, Machine-learning-assisted discovery of polymers with high thermal conductivity using a molecular design algorithm, *npj Computational Materials* **5**, 1 (2019).
  - [10] M. H. Müser, S. V. Sukhomlinov, and L. Pastewka, Interatomic potentials: Achievements and challenges, *Advances in Physics: X* **8**, 2093129 (2023).
  - [11] O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko, and K.-R. Müller, Machine Learning Force Fields, *Chemical Reviews* **121**, 10142 (2021).
  - [12] J. Behler, Four Generations of High-Dimensional Neural Network Potentials, *Chemical Reviews* **121**, 10037 (2021).
  - [13] Z. Chen, N. Andrejevic, N. C. Drucker, T. Nguyen, R. P. Xian, T. Smidt, Y. Wang, R. Ernstorfer, D. A. Tennant, M. Chan, and M. Li, Machine learning on neutron and x-ray scattering and spectroscopies, *Chemical Physics Reviews* **2**, 031301 (2021).
  - [14] G. Ehlers, M. L. Crow, Y. Diawara, F. X. Gallmeier, X. Geng, G. E. Granroth, R. D. Gregory, F. F. Islam, R. O. Knudson, F. Li, M. S. Loyd, and B. Vacaliuc, Modern Trends in Neutron Scattering Instrument Technologies, *Instruments* **6**, 22 (2022).
  - [15] A. Borgschulte, J. Terreni, E. Billeter, L. Daemen, Y. Cheng, A. Pandey, Z. Łodziana, R. J. Hemley, and A. J. Ramirez-Cuesta, Inelastic neutron scattering evidence for anomalous H–H distances in metal hydrides, *Proceedings of the National Academy of Sciences* **117**, 4021 (2020).
  - [16] Y. Q. Cheng, A. I. Kolesnikov, and A. J. Ramirez-Cuesta, Simulation of Inelastic Neutron Scattering Spectra Directly from Molecular Dynamics Trajectories, *Journal of Chemical Theory and Computation* **16**, 7702 (2020).
  - [17] Z. Fan, Y. Wang, P. Ying, K. Song, J. Wang, Y. Wang, Z. Zeng, K. Xu, E. Lindgren, J. M. Rahm, A. J.

- Gabourie, J. Liu, H. Dong, J. Wu, Y. Chen, Z. Zhong, J. Sun, P. Erhart, Y. Su, and T. Ala-Nissila, GPUMD: A package for constructing accurate machine-learned potentials and performing highly efficient atomistic simulations, *The Journal of Chemical Physics* **157**, 114801 (2022).
- [18] E. Fransson, M. Slabanja, P. Erhart, and G. Wahnström, Dynasor—A Tool for Extracting Dynamical Structure Factors and Current Correlation Functions from Molecular Dynamics Simulations, *Advanced Theory and Simulations* **4**, 2000240 (2021).
- [19] E. Berger, E. Fransson, F. Eriksson, E. Lindgren, G. Wahnström, T. H. Rod, and P. Erhart, Dynasor 2: From Simulation to Experiment Through Correlation Functions (2025), arXiv:2503.21957 [cond-mat].
- [20] Z. Fan, Z. Zeng, C. Zhang, Y. Wang, K. Song, H. Dong, Y. Chen, and T. Ala-Nissila, Neuroevolution machine learning potentials: Combining high accuracy and low cost in atomistic simulations and application to heat transport, *Physical Review B* **104**, 104309 (2021).
- [21] Z. Fan, Improving the accuracy of the neuroevolution machine learning potential for multi-component systems, *Journal of Physics: Condensed Matter* **34**, 125902 (2022).
- [22] E. Fransson, J. Wiktor, and P. Erhart, Phase Transitions in Inorganic Halide Perovskites from Machine-Learned Potentials, *The Journal of Physical Chemistry C* **127**, 13773 (2023).
- [23] E. Lindgren, M. Rahm, E. Fransson, F. Eriksson, N. Österbacka, Z. Fan, and P. Erhart, Calorine: A Python package for constructing and sampling neuroevolution potential models, *Journal of Open Source Software* **9**, 6264 (2024).
- [24] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, The atomic simulation environment—a Python library for working with atoms, *Journal of Physics: Condensed Matter* **29**, 273002 (2017).
- [25] F. Eriksson, E. Fransson, and P. Erhart, The Hiphive Package for the Extraction of High-Order Force Constants by Machine Learning, *Advanced Theory and Simulations* **2**, 1800184 (2019).
- [26] M. Parrinello and A. Rahman, Study of an F center in molten KCl, *The Journal of Chemical Physics* **80**, 860 (1984).
- [27] P. Ying, W. Zhou, L. Svensson, E. Berger, E. Fransson, F. Eriksson, K. Xu, T. Liang, J. Xu, B. Song, S. Chen, P. Erhart, and Z. Fan, Highly efficient path-integral molecular dynamics simulations with GPUMD using neuroevolution potentials: Case studies on thermal properties of materials, *The Journal of Chemical Physics* **162**, 064109 (2025).
- [28] M. Rossi, M. Ceriotti, and D. E. Manolopoulos, How to remove the spurious resonances from ring polymer molecular dynamics, *The Journal of Chemical Physics* **140**, 234116 (2014).
- [29] R. Fair, A. Jackson, D. Voneshen, D. Jochym, D. Le, K. Refson, and T. Perring, Euphonic: Inelastic neutron scattering simulations from force constants and visualization tools for phonon properties, *Journal of Applied Crystallography* **55**, 1689 (2022).
- [30] R. Turanyi, A. Jackson, and J. Wilkins, Pace-neutrons/resins: Python library for resolution functions of inelastic neutron scattering instruments (2025), (accessed 2025-04-10).
- [31] K. Dymkowski, S. F. Parker, F. Fernandez-Alonso, and S. Mukhopadhyay, AbINS: The modern software for INS interpretation, *Physica B: Condensed Matter The 11th International Conference on Neutron Scattering (ICNS 2017)*, **551**, 443 (2018).
- [32] O. Arnold, J. C. Bilheux, J. M. Borreguero, A. Buts, S. I. Campbell, L. Chapon, M. Doucet, N. Draper, R. F. Leal, M. A. Gigg, V. E. Lynch, A. Markvardsen, D. J. Mikkelsen, R. L. Mikkelsen, R. Miller, K. Palmen, P. Parker, G. Passos, T. G. Perring, P. F. Peterson, S. Ren, M. A. Reuter, A. T. Savici, J. W. Taylor, R. J. Taylor, R. Tolchenov, W. Zhou, and J. Zikovsky, Mantid—Data analysis and visualization package for neutron scattering and  $\mu$ SR experiments, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **764**, 156 (2014).
- [33] P. Rosander, E. Fransson, N. Österbacka, P. Erhart, and G. Wahnström, Untangling the Raman spectra of cubic and tetragonal  $\text{BaZrO}_3$ , *Physical Review B* **111**, 064107 (2025).
- [34] M. Cardona and G. Güntherodt, eds., *Light Scattering in Solids II*, Topics in Applied Physics, Vol. 50 (Springer, Berlin, Heidelberg, 1982).
- [35] P. E. Blöchl, Projector augmented-wave method, *Physical Review B* **50**, 17953 (1994).
- [36] G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, *Physical Review B* **59**, 1758 (1999).
- [37] G. Kresse and J. Hafner, Ab initio molecular dynamics for liquid metals, *Physical Review B* **47**, 558 (1993).
- [38] G. Kresse and J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, *Computational Materials Science* **6**, 15 (1996).
- [39] G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Physical Review B* **54**, 11169 (1996).
- [40] K. Berland and P. Hyldgaard, Exchange functional that tests the robustness of the plasmon description of the van der Waals density functional, *Physical Review B* **89**, 035412 (2014).
- [41] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, Accurate and Numerically Efficient r2SCAN Meta-Generalized Gradient Approximation, *The Journal of Physical Chemistry Letters* **11**, 8208 (2020).
- [42] S. F. Parker, F. Fernandez-Alonso, A. J. Ramirez-Cuesta, J. Tomkinson, S. Rudic, R. S. Pinna, G. Gorini, and J. Fernández Castañón, Recent and future developments on TOSCA at ISIS, *Journal of Physics: Conference Series* **554**, 012003 (2014).
- [43] R. S. Pinna, S. Rudić, S. F. Parker, J. Armstrong, M. Zanetti, G. Škoro, S. P. Waller, D. Zacek, C. A. Smith, M. J. Capstick, D. J. McPhail, D. E. Pooley, G. D. Howells, G. Gorini, and F. Fernandez-Alonso, The neutron guide upgrade of the TOSCA spectrometer, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated*

- Equipment **896**, 68 (2018).
- [44] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, Array programming with NumPy, *Nature* **585**, 357 (2020).
  - [45] Wes McKinney, Data Structures for Statistical Computing in Python, in *Proceedings of the 9th Python in Science Conference*, edited by Stéfan van der Walt and Jarrod Millman (2010) pp. 56 – 61.
  - [46] The pandas development team, Pandas-dev/pandas: Pandas, Zenodo (2024).
  - [47] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, and P. van Mulbregt, SciPy 1.0: Fundamental algorithms for scientific computing in Python, *Nature Methods* **17**, 261 (2020).
  - [48] J. D. Hunter, Matplotlib: A 2D Graphics Environment, *Computing in Science & Engineering* **9**, 90 (2007).
  - [49] M. Ulmestrand, Perfect-cmaps (2025), (accessed 2025-04-10).
  - [50] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO—the Open Visualization Tool, *Modelling and Simulation in Materials Science and Engineering* **18**, 015012 (2009).
  - [51] D. S. Kim, O. Hellman, J. Herriman, H. L. Smith, J. Y. Y. Lin, N. Shulumba, J. L. Niedziela, C. W. Li, D. L. Abernathy, and B. Fultz, Nuclear quantum effect with pure anharmonicity and the anomalous thermal expansion of silicon, *Proceedings of the National Academy of Sciences* **115**, 1992 (2018).
  - [52] G. L. Squires, *Introduction to the Theory of Thermal Neutron Scattering*, 3rd ed. (Cambridge University Press, Cambridge, 2012).
  - [53] O. Hellman and I. A. Abrikosov, Temperature-dependent effective third-order interatomic force constants from first principles, *Physical Review B* **88**, 144301 (2013).
  - [54] T. Tadano, Y. Gohda, and S. Tsuneyuki, Anharmonic force constants extracted from first-principles molecular dynamics: Applications to heat transfer simulations, *Journal of Physics: Condensed Matter* **26**, 225402 (2014).
  - [55] L. Monacelli, R. Bianco, M. Cherubini, M. Calandra, I. Errea, and F. Mauri, The stochastic self-consistent harmonic approximation: Calculating vibrational properties of materials with full quantum and anharmonic effects, *Journal of Physics: Condensed Matter* **33**, 363001 (2021).
  - [56] A. Perrichon, N. Torino, E. Jedvik Granhed, Y.-C. Lin, S. F. Parker, M. Jiménez-Ruiz, M. Karlsson, and P. F. Henry, Local Coordination Environments and Vibrational Dynamics of Protons in Hexagonal and Cubic Sc-Doped BaTiO<sub>3</sub> Proton-Conducting Oxides, *The Journal of Physical Chemistry C* **124**, 8643 (2020).
  - [57] C. Chen and S. P. Ong, A universal graph deep learning interatomic potential for the periodic table, *Nature Computational Science* **2**, 718 (2022).
  - [58] B. Deng, P. Zhong, K. Jun, J. Riebesell, K. Han, C. J. Bartel, and G. Ceder, CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling, *Nature Machine Intelligence* **5**, 1031 (2023).
  - [59] A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon, and E. D. Cubuk, Scaling deep learning for materials discovery, *Nature* **624**, 80 (2023).
  - [60] F. Xie, T. Lu, S. Meng, and M. Liu, GPTFF: A high-accuracy out-of-the-box universal AI force field for arbitrary inorganic materials, *Science Bulletin* **69**, 3525 (2024).
  - [61] I. Batatia, P. Benner, Y. Chiang, A. M. Elena, D. P. Kovács, J. Riebesell, X. R. Advincula, M. Asta, M. Avaylon, W. J. Baldwin, F. Berger, N. Bernstein, A. Bhowmik, S. M. Blau, V. Cărare, J. P. Darby, S. De, F. D. Pia, V. L. Deringer, R. Elijošius, Z. El-Machachi, F. Falcioni, E. Fako, A. C. Ferrari, A. Genreith-Schriever, J. George, R. E. A. Goodall, C. P. Grey, P. Grigorev, S. Han, W. Handley, H. H. Heenen, K. Hermansson, C. Holm, J. Jaafar, S. Hofmann, K. S. Jakob, H. Jung, V. Kapil, A. D. Kaplan, N. Karimitari, J. R. Kermode, N. Kroupa, J. Kullgren, M. C. Kuner, D. Kuryla, G. Liepuoniute, J. T. Margraf, I.-B. Magdău, A. Michaelides, J. H. Moore, A. A. Naik, S. P. Niblett, S. W. Norwood, N. O'Neill, C. Ortner, K. A. Persson, K. Reuter, A. S. Rosen, L. L. Schaaf, C. Schran, B. X. Shi, E. Sivonxay, T. K. Stenczel, V. Svahn, C. Sutton, T. D. Swinburne, J. Tilly, C. van der Oord, E. Varga-Umbrich, T. Vegge, M. Vondrák, Y. Wang, W. C. Witt, F. Zills, and G. Csányi, A foundation model for atomistic materials chemistry (2024), arXiv:2401.00096 [physics].

# Supporting Information

## **Predicting neutron experiments from first principles: A workflow powered by machine learning**

Eric Lindgren<sup>1</sup>, Adam Jackson<sup>2</sup>, Erik Fransson<sup>1</sup>, Esmée Berger<sup>1</sup>, Goran Škoro<sup>1</sup>,  
Svemir Rudić<sup>1</sup>, Rastislav Turanyi<sup>2</sup>, Sanghamitra Mukhopadhyay<sup>3</sup>, and Paul  
Erhart<sup>1,\*</sup>

<sup>1</sup> *Department of Physics, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*

<sup>2</sup> *Scientific Computing Department, STFC Rutherford Appleton Laboratory, Didcot OX11 0QX, UK*

<sup>3</sup> *ISIS Neutron and Muon Source, STFC Rutherford Appleton Laboratory, Didcot OX11 0QX, UK*  
\**erhart@chalmers.se*

## **Contents**

<b>S1 NEP model for benzene: Training curves and parity plots</b>	<b>S2</b>
<b>S2 NEP model for hydrogenated Sc-doped BaTiO<sub>3</sub>: Training curves and parity plots</b>	<b>S4</b>
<b>S3 Path-integral molecular dynamics convergence study for crystalline benzene and hydrogenated Sc-doped BaTiO<sub>3</sub></b>	<b>S5</b>
<b>S4 Temperature dependence of Inelastic Neutron Scattering spectrum for crystalline benzene</b>	<b>S6</b>
<b>S5 Red-shift for vibrational spectrum for benzene</b>	<b>S7</b>
<b>References</b>	<b>S9</b>



# S1 NEP model for benzene: Training curves and parity plots

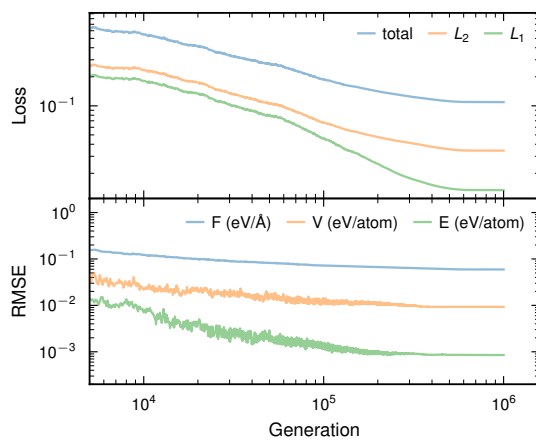


Figure S1: Training curves for the final model for benzene, trained on all available data. This model was used for production runs.

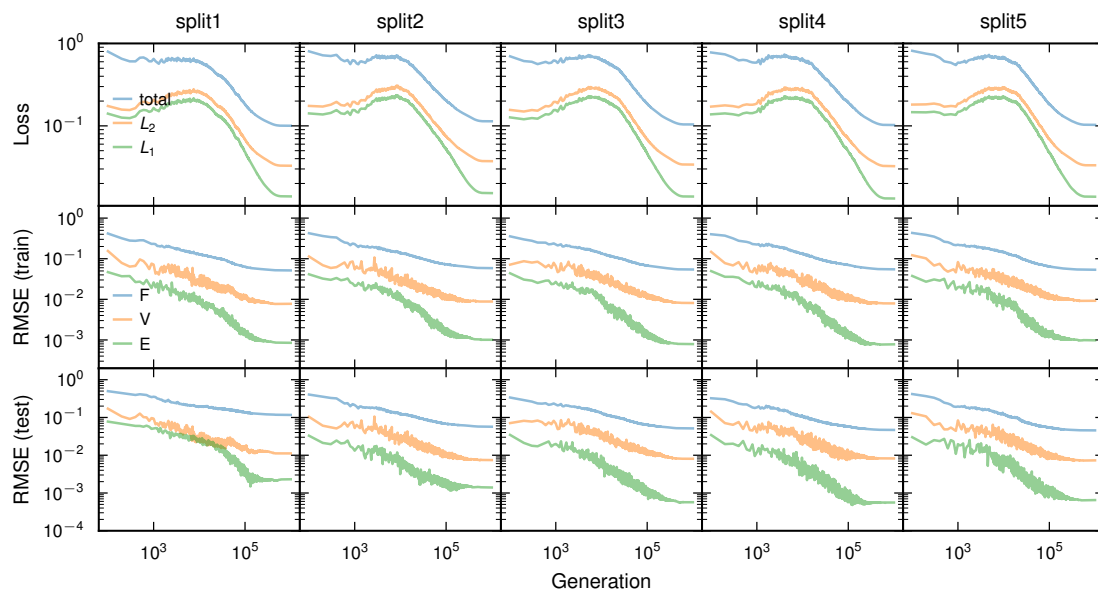


Figure S2: Training curves for the models in the ensemble trained on various K-fold cross validation splits of the training data.

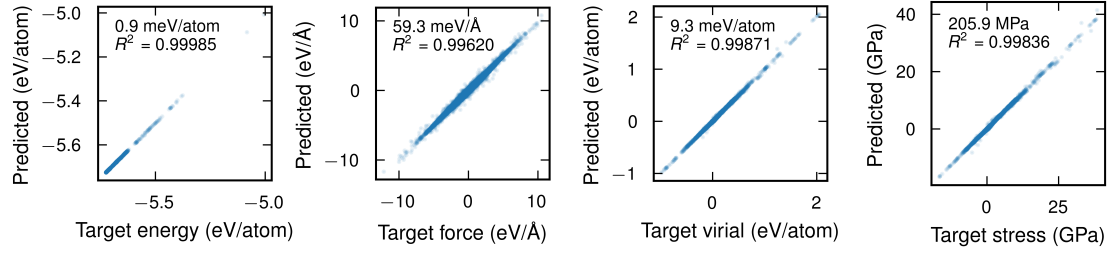


Figure S3: Parity plots for the final benzene model, evaluated on all structures in the data set.

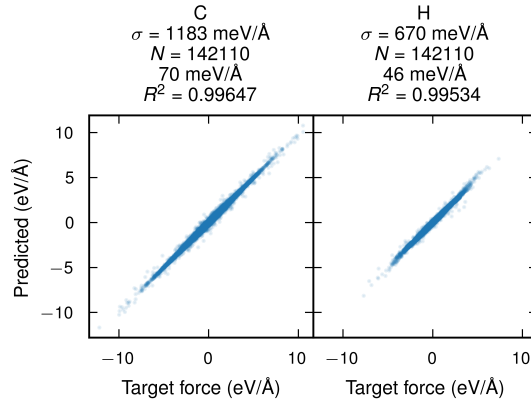


Figure S4: Parity plots for the final benzene model, evaluated on all structures in the data set, split by species. 'C' denotes carbon, and 'H' denotes hydrogen.

## S2 NEP model for hydrogenated Sc-doped BaTiO<sub>3</sub>: Training curves and parity plots

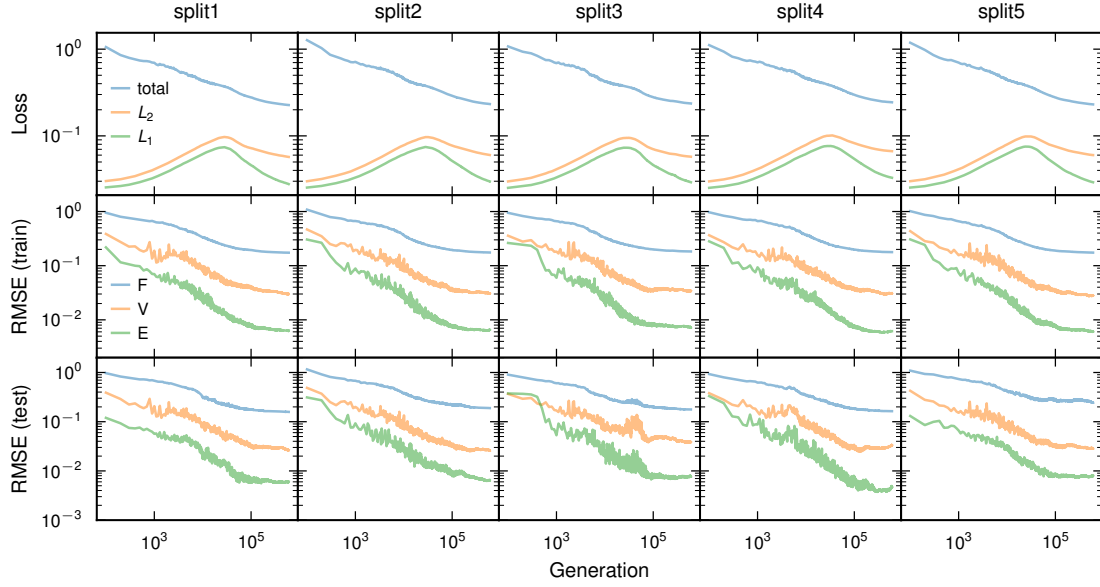


Figure S5: Training curves for the models in the ensemble trained on various K-fold cross validation splits of the training data, for hydrogenated Sc-doped BaTiO<sub>3</sub>. The 'split1' model was used for production runs.

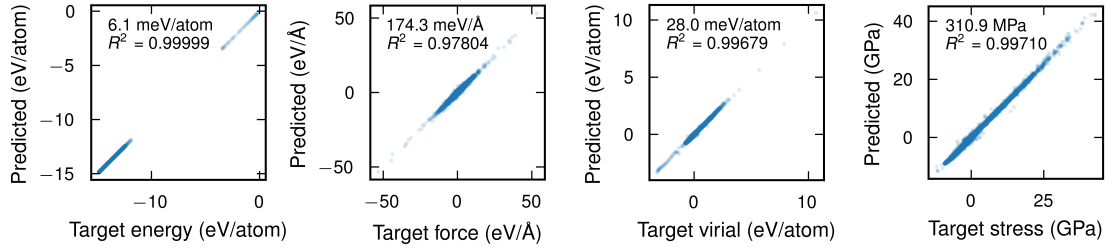


Figure S6: Parity plots for the final hydrogenated Sc-doped BaTiO<sub>3</sub> model, referred to as 'split1' in Fig. S5, evaluated on all structures in the data set.

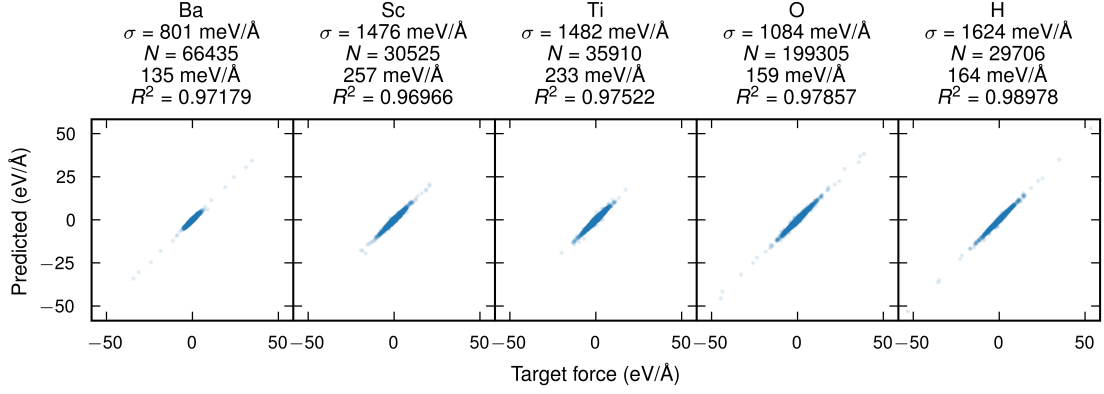


Figure S7: Parity plots for the final hydrogenated Sc-doped BaTiO<sub>3</sub> model, referred to as 'split1' in Fig. S5, evaluated on all structures in the data set, split by species.

### S3 Path-integral molecular dynamics convergence study for crystalline benzene and hydrogenated Sc-doped BaTiO<sub>3</sub>

Here we vary the number of PIMD beads and study the convergence for various properties for crystalline benzene and hydrogenated Sc-doped BaTiO<sub>3</sub>. For all simulations in this work, we used 32 PIMD beads. The computational cost of a PIMD simulation scales linearly with the number of beads. For benzene, we thus focused our simulations on a temperature of 127 K, since we had available experimental data at that temperature, and the computational cost of performing PIMD with more than 32 beads was deemed too excessive (Fig. S8). For completeness, we include simulations at a range of other temperatures for crystalline benzene compared to experimental data, but note that the number of beads is not necessarily sufficient in all cases (Sect. S4).

For hydrogenated Sc-doped BaTiO<sub>3</sub> we performed the convergence study in the hexagonal structure with a Sc doping fraction of 0.16 (Fig. S9). The simulations presented in the main work are performed at 15 K. 32 beads was again used to strike a balance between accuracy and computational cost, but do note that the potential energy and volume of the systems are not fully converged for that number of beads.

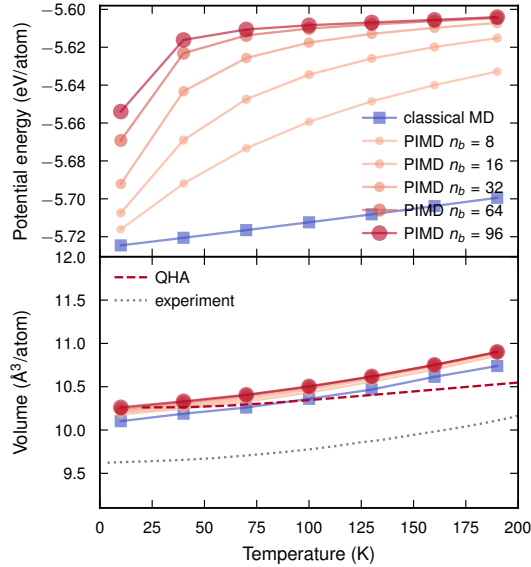


Figure S8: Potential energy and volume from molecular dynamics simulations run in the PIMD ensemble for crystalline benzene.

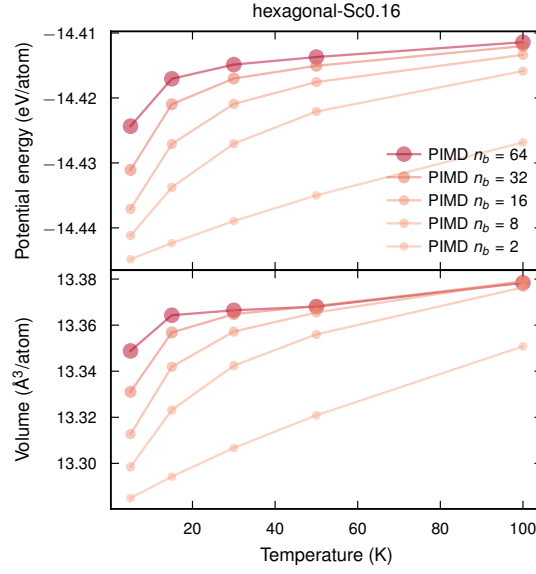


Figure S9: Potential energy and volume from molecular dynamics simulations run in the PIMD ensemble for hydrogenated Sc-doped BaTiO<sub>3</sub> in the hexagonal phase with a doping fraction of 0.16.

## S4 Temperature dependence of Inelastic Neutron Scattering spectrum for crystalline benzene

Experimental data compared with simulated spectra at 127 K, 103 K, 75 K, and 24 K. Note that the number of PIMD beads used to perform equilibration is not necessarily enough, see Sect. S3



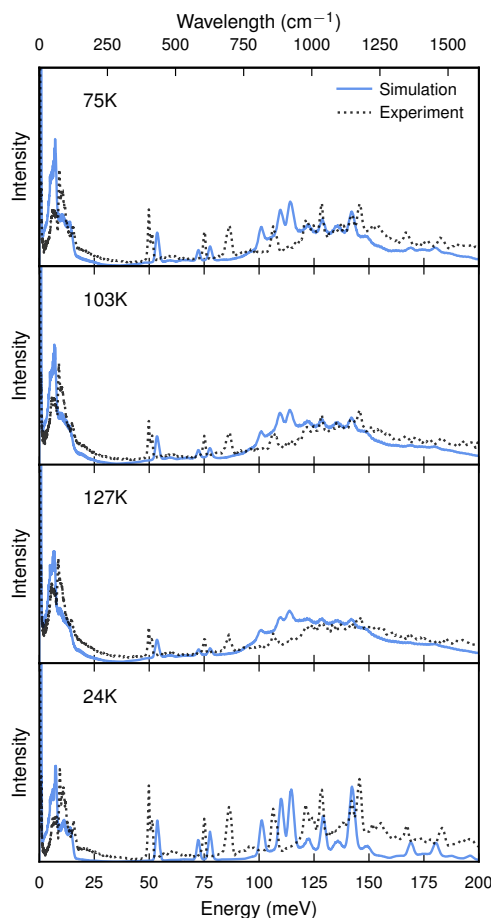


Figure S10: Predicted INS spectrum for crystalline benzene at the TOSCA spectrometer, at 127 K, 103 K, 75 K, and 24 K.

## S5 Red-shift for vibrational spectrum for benzene

The simulated INS spectrum for benzene is red-shifted compared to the experimental spectrum measured at TOSCA. We compare the harmonic density of states (DOS) calculated using our production NEP model for benzene to calculations and experiments performed at low temperature Fig. S11. We focus on the harmonic DOS to decouple the predicted mode energies from the temperature in our MD simulations, as well as the ensemble in which the simulation was conducted in (in this case, NVE). Focusing on the region around 100 meV, we observe that the DOS for our NEP model trained on the vdW-CX functional is red-shifted by approximately 25 meV compared to the collected Raman and IR spectroscopy experiments, as well as the simulated values based on Møller-Plesset perturbation theory (MP2) and DFTMD using a LDA functional. All collected experiments and simulations agree well with the experimental INS spectrum measured at TOSCA. The NEP model used in this work is well-converged with regards to the target energies, forces, and virials calculated using the vdW-CX functional, achieving errors of 8.510 meV atom<sup>-1</sup> for the energies, 59.29 meV Å<sup>-1</sup> for the forces, and 9.295 meV atom<sup>-1</sup> for the virials (Sect. S1). We thus attribute a systematic red-shift of the predicted DOS for our NEP model to be a feature of the vdW-CX functional.

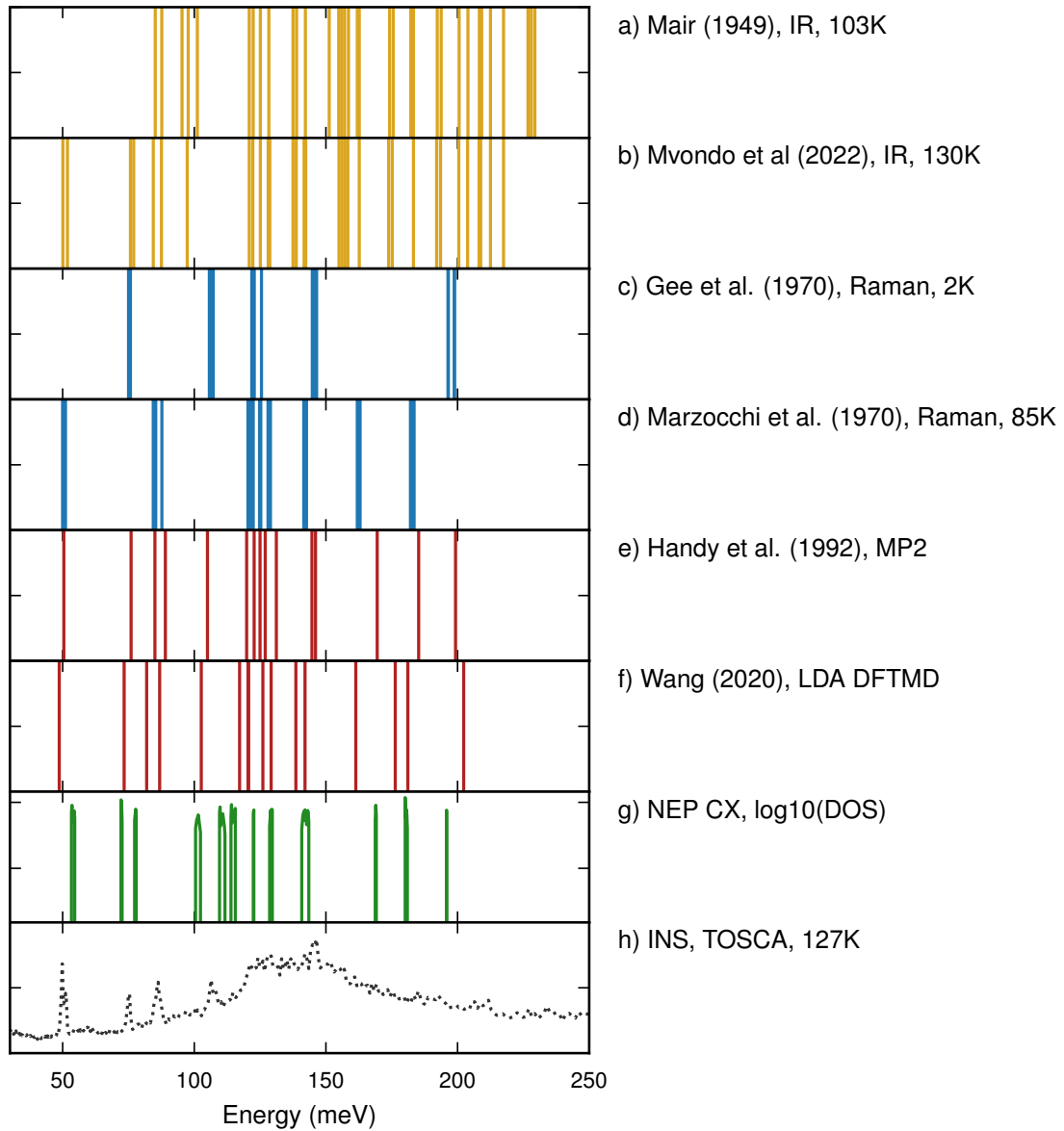


Figure S11: Compilation of density of vibrational frequencies for benzene from the literature. a) IR spectrum measured by Mair et al.<sup>1</sup> b) IR spectrum measured by Mvondo et al.<sup>2</sup> c) Raman spectrum measured by Gee et al.<sup>3</sup> d) Raman spectrum measured by Marzocchi et al.<sup>4</sup> e) Møller-Plesset perturbation theory (MP2) calculations by Handy et al.<sup>5</sup> f) Vibrational modes from DFTMD by Wang et al.<sup>6</sup> g) Harmonic density of states (DOS) computed using the NEP model for benzene developed in the course of the present work. h) Experimental INS spectrum for crystalline benzene measured at TOSCA at 127 K, measured in this work. The experimental results and simulations from the literature generally agree well with our measured INS spectrum in the region around 100 meV, indicating that the red-shift we observe for the our NEP model (g) is related to the DFT-functional used to train the model.

## References

- [1] Robert D. Mair and Donald F. Hornig. The Vibrational Spectra of Molecules and Complex Ions in Crystals. II. Benzene. *The Journal of Chemical Physics*, 17(12):1236–1247, December 1949. ISSN 0021-9606. doi: 10.1063/1.1747149. URL <https://doi.org/10.1063/1.1747149>.
- [2] Delphine Nna-Mvondo and Carrie M. Anderson. Infrared Spectra, Optical Constants, and Temperature Dependences of Amorphous and Crystalline Benzene Ices Relevant to Titan. *The Astrophysical Journal*, 925(2):123, February 2022. ISSN 0004-637X. doi: 10.3847/1538-4357/ac350c. URL <https://dx.doi.org/10.3847/1538-4357/ac350c>.
- [3] A. Roger Gee and G. Wilse Robinson. Raman Spectrum of Crystalline Benzene. *The Journal of Chemical Physics*, 46(12):4847–4853, June 1967. ISSN 0021-9606. doi: 10.1063/1.1840646. URL <https://doi.org/10.1063/1.1840646>.
- [4] M. P. Marzocchi, H. Bonadeo, and G. Taddei. Infrared Spectra in Polarized Light of Crystalline Benzene and Benzene-d6. *The Journal of Chemical Physics*, 53(3):867–875, August 1970. ISSN 0021-9606. doi: 10.1063/1.1674151. URL <https://doi.org/10.1063/1.1674151>.
- [5] Nicholas C. Handy, Paul E. Maslen, Roger D. Amos, Jamie S. Andrews, Christopher W. Murray, and Gregory J. Laming. The harmonic frequencies of benzene. *Chemical Physics Letters*, 197(4):506–515, September 1992. ISSN 0009-2614. doi: 10.1016/0009-2614(92)85808-N. URL <https://www.sciencedirect.com/science/article/pii/000926149285808N>.
- [6] Shaoqing Wang. Intrinsic molecular vibration and rigorous vibrational assignment of benzene by first-principles molecular dynamics. *Scientific Reports*, 10(1):17875, October 2020. ISSN 2045-2322. doi: 10.1038/s41598-020-74872-6. URL <https://www.nature.com/articles/s41598-020-74872-6>.