

Efficient machine learning interatomic potentials robust for liquid and multiple solid polymorphs of NaF and KF

Zhao Fan ^{1,2} Michael L. Whittaker ^{1,3,*} and Mark Asta ^{1,2,†}

¹Materials Sciences Division, [Lawrence Berkeley National Laboratory](#), Berkeley, California 94720, USA

²Department of Materials Science and Engineering, [University of California, Berkeley](#), California 94720, USA

³Energy Geosciences Division, [Lawrence Berkeley National Laboratory](#), Berkeley, California 94720, USA



(Received 30 July 2025; accepted 10 September 2025; published 31 October 2025)

Achieving atomic-level understanding of crystallization of molten salts is of importance to a wide range of technological applications. Recent work [Fan *et al.*, [Proc. Natl. Acad. Sci. USA](#) **122**, e2425702122 (2025)] revealed that crystal nucleation in molten LiF salt is a multistage process according to the molecular-dynamics (MD) simulations based on an atomic cluster expansion (ACE) machine-learning interatomic potential (MLIP). In order to understand the influence of increasing cation size on nucleation pathways and nucleation rates of molten fluoride salts, here we develop two new ACE MLIPs for NaF and KF. The two ACE MLIPs feature DFT-SCAN-level accuracy for liquid and multiple solid polymorphs over a wide temperature (0–2000 K) and pressure (0–100 GPa) range, and also reproduce well a number of experimental data for solid and liquid equilibrium properties. The efficiency of the two ACE MLIPs enable million-atom-scale or microsecond-scale MD simulations. The two general-purpose ACE MLIPs are expected to be useful for atomistic simulations for different purposes, in addition to studying crystallization of molten NaF and KF salts.

DOI: [10.1103/xbfm-clgd](https://doi.org/10.1103/xbfm-clgd)

I. INTRODUCTION

Molten salts are a family of high-temperature ionic liquids technologically important for a wide range of applications. For example, molten salts can be utilized as a coolant or fuel salt for molten-salt reactors [1,2] and energy storage mediums for concentrated solar power [3,4], as well as serving as solvents for the synthesis of novel materials, such as colloidal III-V semiconductor nanocrystals [5] and cation-disordered rocksalt cathodes for Li-ion batteries [6]. In addition, crystallization of molten salts plays a critical role in geological processes [7,8] as well as industrial separation processes [9].

Molten salts relevant to these technological applications are generally multicomponent in nature. In order to accelerate the screening of molten salts suitable for different applications from the vast composition space, it is imperative to improve our understanding of liquid structure and its relation to thermochemical, thermophysical, and thermodynamic properties as well as crystallization behavior of molten salts on the molecular level. This is a challenging task for current experimental approaches due to the elevated temperature and high reactivity of molten salts. Molecular-dynamics (MD) simulations are effective and efficient alternatives to achieve this

long-standing goal. *Ab initio* molecular-dynamics (AIMD) simulations using density-functional theory (DFT) have been applied to investigate liquid structure and calculate certain thermophysical and thermodynamic properties of molten salts [10–13]. Although their accuracy is desirable, it is impractical to apply AIMD simulations to compute many practically relevant properties of molten salts, e.g., some transport properties, with acceptable uncertainties, let alone to study crystal nucleation of molten salts, because of the limited time and length scale of AIMD. Classical interatomic potentials, e.g., Tosi-Fumi (TF) or Born-Mayer-Huggins potential [14], enabled long-timescale and large-scale MD simulations of molten salts [8,15–18]; it is difficult for TF potentials to accurately describe structure of molten salts because polarization effect is ignored [19,20]. The polarizable-ion model (PIM) potential specifically took polarization effect into account and thus significantly improves the accuracy in simulating molten salts [19]. But MD simulations based on PIM potentials are usually computationally expensive [21].

Recently, machine-learning interatomic potentials (MLIP) have been demonstrated to possess accuracy near that of DFT and efficiency close to that of classical potentials simultaneously [22–26]. Recent MD simulations based on MLIPs made significant contributions in shedding new light on microscopic mechanisms of structure, phase transitions, and deformation of both semiconductors [27–29] and metals [30–32]. Over the past five years, there are also extensive efforts in developing MLIPs for molten salts [21,33–42]. These MLIPs were demonstrated to be highly accurate in predicting liquid structure as well as many thermochemical, thermophysical, and thermodynamic properties of molten salts, but their application to predicting properties of solid phases and crystallization

*Contact author: mwhittaker@lbl.gov

†Contact author: mdasta@lbl.gov

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](#) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

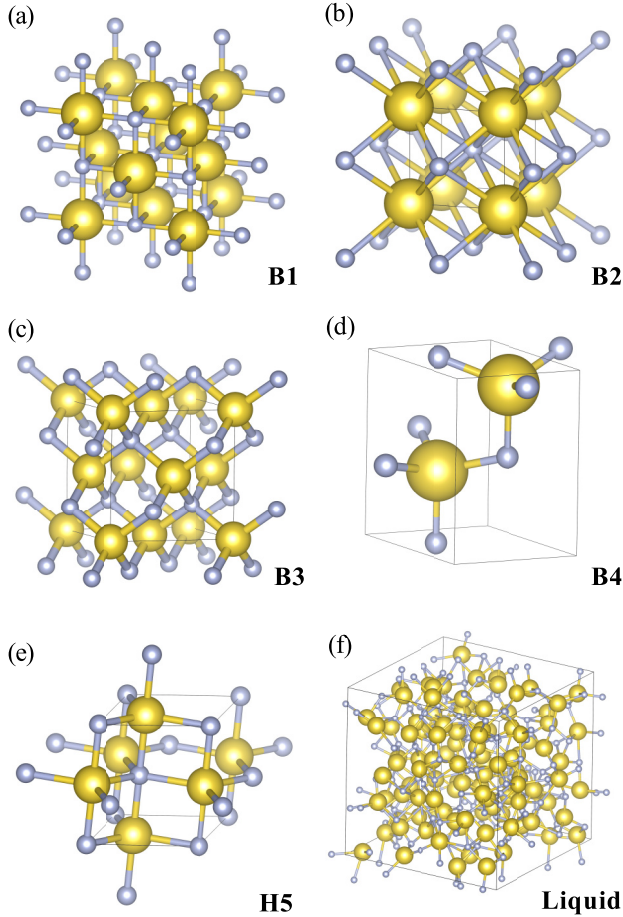


FIG. 1. The six phases considered in developing the MLIPs for both NaF and KF. (a)–(e) show the unit cells of the five solid polymorphs; (f) shows a representative liquid snapshot containing 216 atoms.

behavior has not been explored. For applications to crystal nucleation from molten salts, our prior work [43] has demonstrated the importance of configurations corresponding to metastable solid polymorphs, which are not typically considered in the development of MLIPs for salts. Furthermore, it is difficult to apply some of these existing MLIPs to conduct long-timescale or large-scale MD simulations, which are desirable especially for studying crystal nucleation and growth.

In recent work [43], MD simulations based on an MLIP revealed that the crystal nucleation in LiF melt is a multi-stage process and the structure evolution over time during crystal nucleation shows resemblance to the structure variation in space across crystal-melt interface. These results raise

interesting questions about crystallization of molten salts. For example, what is the variation of liquid structure as well as nucleation pathway and rate when increasing cation size in molten fluoride salts? Can crystal nucleation pathways of molten salts be inferred from equilibrated crystal-melt interface structure generally?

In the present work, in order to enable such questions to be addressed in MD simulations, we develop atomic cluster expansion (ACE) [26,44–47] MLIPs for NaF and KF, trained on the reference data from DFT calculations with the strongly constrained and appropriately normed (SCAN) [48] functional. The two ACE potentials are not only accurate in describing the DFT-SCAN energies and forces for liquid and multiple solid polymorph configurations representative of wide temperature (T) and pressure (P) ranges, but also can reproduce experimentally measured properties. Thanks to its efficiency, it is feasible to employ each of the two ACE potentials to conduct million-atom MD simulations for tens of nanoseconds or microsecond-scale MD simulations containing around 10 000 atoms. The two ACE potentials are anticipated to be applicable for MD simulations over wide T and P ranges for different purposes, such as establishing atomic-level structure-property relationships for molten salts and investigating the mechanisms underlying polymorphic transitions and melting of alkali halides [49–51], in addition to understanding crystallization of molten salts. Moreover, the training datasets presented in the current work for NaF and KF in conjunction with that for LiF in Ref. [43] can be expanded to train robust MLIPs for pseudobinary and -ternary fluoride salt systems.

II. METHODS

A. Generating training datasets

As we aim to develop general-purpose MLIPs, our training datasets include not only configurations representative of the liquid and stable rocksalt (B1) crystal structure, but also four metastable crystalline phases, i.e., B2, B3, B4, and H5 phases, over wide temperature and pressure ranges. Figure 1 displays the unit cells of the five solid polymorphs as well as a representative snapshot from the liquid. The structural characteristics of the five solid polymorphs are summarized in Table I. Especially, the H5 phase is a crystalline phase encountered in our prior work developing the MLIP for LiF [43]. The coordination polyhedron of the H5 phase is trigonal bipyramidal (fivefold coordination), and the sublattice features hexagonal symmetry similar to hexagonal close packing (HCP) but with a highly compressed c/a ratio of 1.06 (relative to the ideal close-packed value of 1.63). The sublattice is thus

TABLE I. The structural characteristics of the five solid polymorphs considered for both NaF and KF.

Index	Description	Strukturbericht designation	Prototype	Space group	Sublattice	Polyhedron
1	Rock salt	B1	NaCl	$Fm\bar{3}$	FCC	Octahedron
2	Caesium chloride	B2	CsCl	$Pm\bar{3}$	SC	Cube
3	Zincblende	B3	ZnS	$F\bar{4}3m$	FCC	Tetrahedron
4	Wurtzite	B4	ZnS	$P6_3mc$	HCP	Tetrahedron
5	H5	None	Unknown	$P6_3/mmc$	CHP	Trigonal bipyramidal

termed compressed hexagonal packing (CHP). It was shown that B1, B3, B4, and H5 phases are (meta)stable but B2 phase is unstable when $P < 100$ GPa for LiF simulated with the ACE-SCAN potential [43]. As will be shown later, all of the five solid polymorphs including B2 phase are (meta)stable over a certain temperature range even at zero pressure for both NaF and KF. Furthermore, in order to ensure that the developed MLIPs are valid for simulating crystallization of molten salts, our training datasets also include crystals formed from liquids, which are highly distorted or defective compared to bulk crystals.

We used an active learning strategy to gradually expand the training and test datasets for both NaF and KF. Supplemental Material (SM) [52] records the complete process of generating the training and test datasets and describes detailed information of atomic configurations included in the training and test datasets for both NaF and KF. The final training and test datasets for NaF contain 15 720 configurations (4 124 984 atoms) and 3961 configurations (1 041 926 atoms), respectively. And the final training and test datasets for KF contain 14 964 configurations (3 642 464 atoms) and 3765 configurations (915 946 atoms), respectively.

B. DFT calculations

We use the Vienna *ab initio* simulation package (VASP) [53–55] to conduct static DFT calculations to obtain energies and forces of atomic configurations in both training and test datasets, which are the target properties in fitting the MLIPs. Because we found that the MLIP trained on reference data from DFT calculations with strongly constrained and appropriately normed (SCAN) [48] functional and without including corrections for van der Waals dispersion reproduce data closest to experiments for LiF [43], here all DFT calculations use SCAN functional and do not include dispersion corrections for both NaF and KF. The same plane-wave energy cutoff of 600 eV is used for both NaF and KF. Brillouin zone integration is performed using the Γ point only, because the configurations on which static DFT calculations were conducted are relatively large (>200 atoms). The convergence criterion for electronic self-consistent calculations is 0.1 meV.

C. Fitting MLIPs

The MLIPs are trained based on the ACE formalism [26,44–47]. When fitting ACE potentials, a linear embedding function and 268 basis functions are used for both NaF and KF. Cutoff radii of 6.3 and 6.5 Å are used for NaF and KF, respectively. More basis functions and larger cutoff values were considered but we did not find significant improvement in the training result.

Starting from a small training dataset, we iteratively expanded the training datasets through adding new configurations with large extrapolation grade (γ , a parameter measuring the dissimilarity degree between configurations in MD simulations based on an ACE potential and those included in the training dataset for constructing the ACE potential [46]) or short bond lengths or new configurations selected at uniform time intervals from MD simulations based on the latest ACE potential and then retraining the ACE potentials until the maximum absolute error (MAX_AE) of energy is smaller than

5 meV/atom for all phases over the T range from 0 to 2000 K and the P range from 0 to 100 GPa for NaF and over the T range from 0 to 1600 K and the P range from 0 to 100 GPa for KF. Configurations selected from each MD simulations were always split into the training and test datasets as a ratio of 4:1. Please see SM [52] for the detailed processes of iteratively expanding the training and test datasets for both Na and KF. As shown in Fig. 2, the final training (test) root-mean-square error (RMSE) of energy are 2.05 (2.14) meV/atom and 1.95 (1.94) meV/atom for NaF and KF, respectively. And the final training (test) RMSE of force are 37.25 (35.77) meV/Å and 28.79 (28.86) meV/Å for NaF and KF, respectively. An additional validation will be presented later in Sec. III A to demonstrate that the final ACE potentials are robust for both perfect and defected/distorted configurations of different phases over wide T and P ranges.

D. MD simulations

All MD simulations are carried out using the large-scale atomic/molecular massively parallel simulator (LAMMPS) [56]. The time step is always 1 fs and periodic boundary conditions are applied in all three directions. Unless otherwise stated, all simulations are conducted in the isothermal-isobaric (NPT) ensemble under a Nosé–Hoover thermostat with zero pressure. The visualization of atomic configurations is realized using the OVITO [57] and VESTA [58] packages. The self-diffusion coefficient is derived from the slope of mean-square displacement (MSD) against time. The heat capacity C_p was computed using the enthalpy fluctuation formulation, i.e., $\langle \delta(U + PV)^2 \rangle = k_B T^2 C_p$, where $\langle \delta(U + PV)^2 \rangle = \langle (U + PV)^2 \rangle - \langle (U + PV) \rangle^2$, U is the sum of potential and kinetic energy, k_B is the Boltzmann constant, T and P are the specified temperature and pressure, respectively [59]. At each T , data points collected over 0.1 ns in the NPT ensemble after equilibration are used to compute $\langle \delta(U + PV)^2 \rangle$. Standard deviations are obtained from five independent time windows.

E. Estimating melting points

The melting point T_m of the NaF and KF models based on the optimized ACE potentials are determined in two steps. First, MD- NP_zAT simulations of solid (B1 crystal)-liquid systems are performed at various trial temperatures and the system volume evolution with time was monitored throughout the simulation. In the NP_zAT ensemble [60], the total number of atoms N and area parallel to the interface A are constant, while average temperature T and stress normal to the interface (P_z , set to zero in the simulations) are set using a Nose-Hoover thermostat and barostat. The solid-liquid system contains 32 000 atoms and has an initial dimension of $10a \times 10a \times 40a$, where a is the lattice constant of rocksalt crystal (B1 phase) of NaF or KF at the temperature used to equilibrate the system. When T is close to T_m , the system volume should be almost constant in the MD- NP_zAT simulation. From this step, the initial estimate of T_m for NaF and KF are 1325 and 1073 K, respectively. In the second step, in order to determine T_m as accurately as possible, the so-called coexistence approach [61–64] is used. Specifically, a series of microcanonical MD (fixed energy, volume, and atom number, i.e., NVE) simulations are performed on the solid-

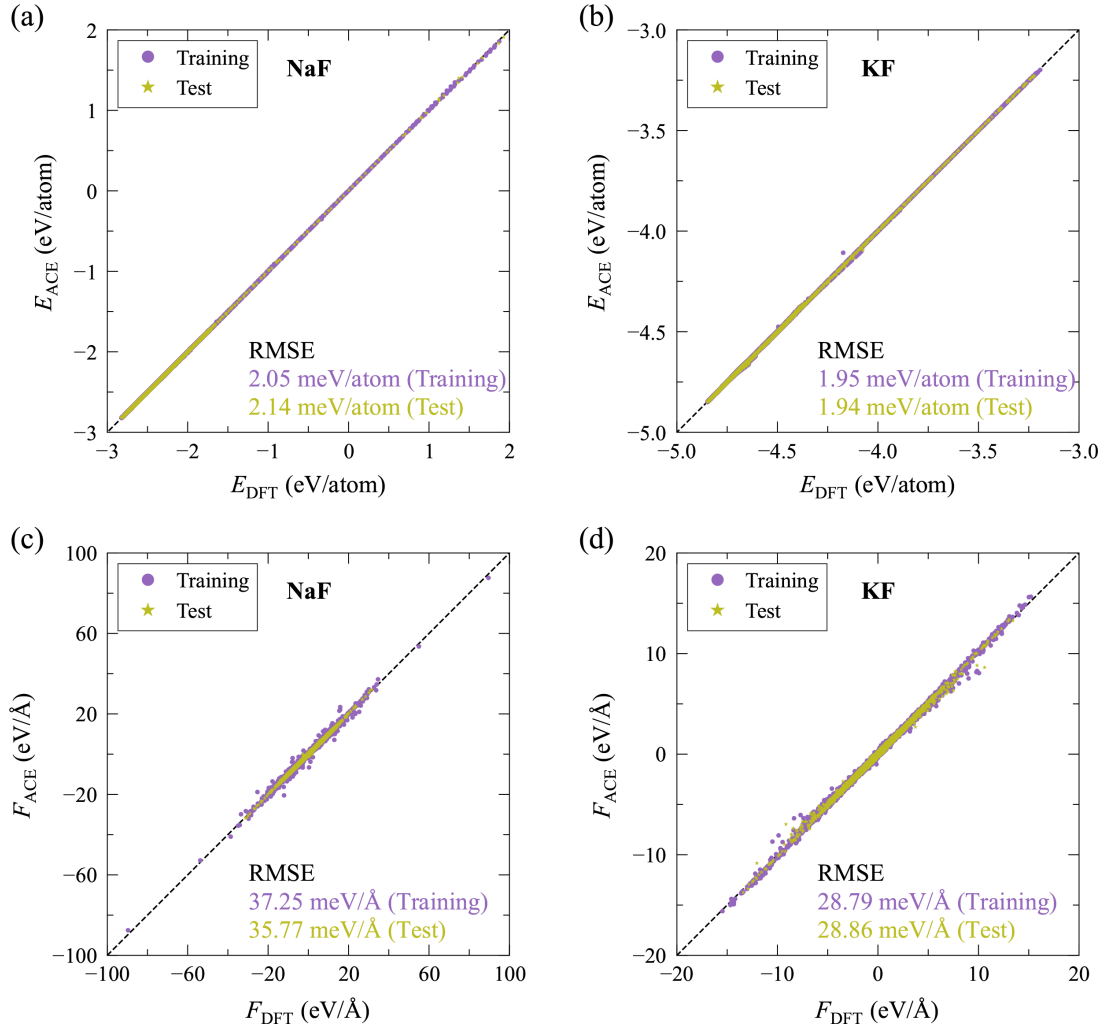


FIG. 2. The final training and test results. (a) and (b) show the energies predicted by the final ACE potential against those from DFT calculations in both training and test datasets for NaF and KF, respectively. (c) and (d) show the force components along three directions predicted by the final ACE potential against those from DFT calculations in both training and test datasets for NaF and KF, respectively. In (c) and (d), only 50 000 data points (force components) with the largest errors are shown for each dataset. The final root-mean-squared error (RMSE) for both energy and forces are denoted on each panel.

liquid system that are first equilibrated at the T_m initially estimated from the MD- NP_zAT simulation, after assigning slightly different initial T . In the NVE simulation, the temperature of the system, as well as the volume fraction of solid and liquid phases, will evolve until equilibrium between the solid and liquid phases is achieved. The accurate estimate of T_m at zero pressure is the zero-stress coexistence temperature, which is determined through fitting the data points of average temperature versus average pressure over the last 0.5 ns of these NVE simulations. We also used larger simulation boxes to determine T_m but did not find significant differences.

III. RESULTS AND DISCUSSION

In this section, we will first demonstrate the accuracy of the two ACE potentials relative to DFT data for configurations representing liquid and five solid polymorphs over wide T and P ranges. We further show that the ACE potentials can correctly capture the key characteristics of ionic bonding,

even though charge is not considered explicitly in the ACE implementation. Then we will show that various properties produced from MD simulations based on the ACE potentials are consistent with experimental measurements. Finally, the efficiency of the ACE potentials will be discussed.

A. DFT-level accuracy over wide T and P range

According to Figs. 3(a) and 3(b) showing the energy-volume curves at 0 K from DFT calculations for the five solid polymorphs of NaF and KF, respectively, the stable crystal at 0 K is B1 phase for both compounds with the SCAN functional, consistent with experimental observations. Further, B3, B4, and H5 are metastable phases with larger volume while B2 is a denser metastable phase. Figures 3(c) and 3(d) show the energy difference in these equations of states between ACE and DFT for NaF and KF, respectively, indicating that the ACE potentials can accurately reproduce the results for the five solid polymorphs over a wide volume range. The absolute-energy difference is typically only a few meV/atom,

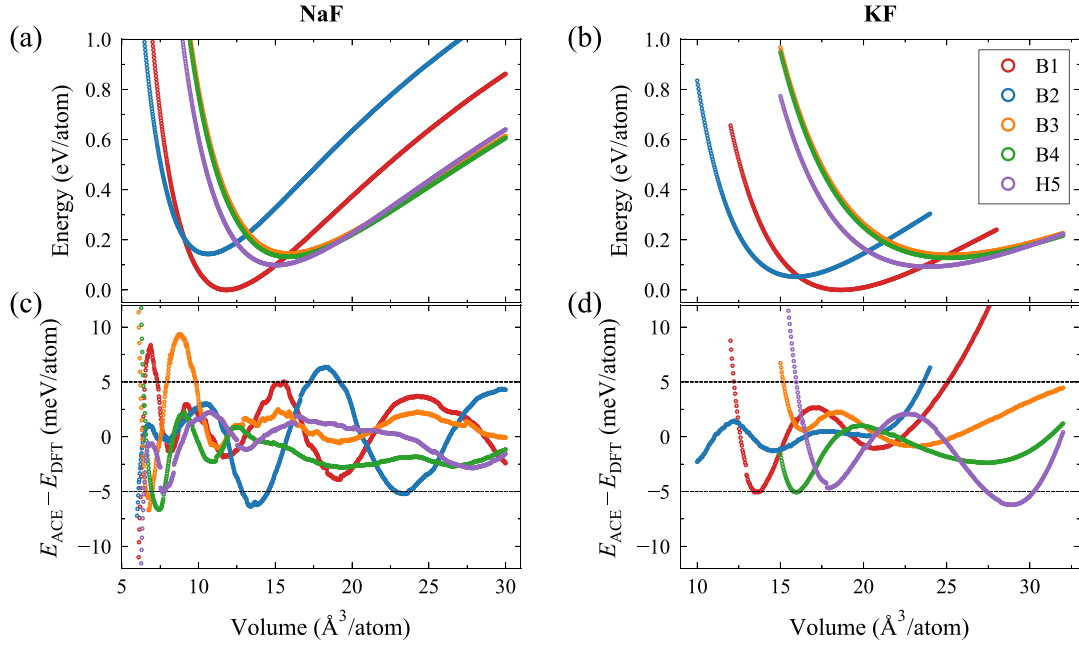


FIG. 3. (a) and (b) show the energy-volume curves at 0 K from DFT calculations for the five solid polymorphs of NaF and KF, respectively. (c) and (d) show the energy difference in these energy-volume curves between ACE and DFT for NaF and KF, respectively.

and takes values greater than 5 meV/atom only at volumes far from equilibrium values. These results demonstrate that the ACE potentials reproduce well DFT energies for the five solid polymorphs over a wide range of P at 0 K.

In what follows, we will show that the ACE potentials also possess high accuracy relative to DFT at finite T up to 2000 K and over a wide P range. Although we achieved very low error of both energy and force for the final training and test datasets (Fig. 2), the final training and test results derived through active learning may not clearly reflect the performance of the final ACE potentials because the final datasets include configurations generated using early-version potentials. Additionally, some existing MLIPs appearing to perform well in short timescale (several nanoseconds) simulations may fail when running long-timescale MD simulations. The failure cause is that there is more probability of sampling configurations far beyond the configuration space upon which one MLIP is trained during long-timescale MD simulations. The MLIP may predict nonphysical and extremely high forces for atoms in these much different configurations, leading to large displacements of the atoms. Once such situation occurs, the LAMMPS would terminate the MD simulation and report an error of losing atoms. Therefore, we used the final ACE potentials to run additional long-timescale MD simulations starting from different phases, from which we selected configurations to assess the error of both energy and force between ACE and DFT over different T and P ranges. These tests are made to ensure that the two ACE potentials are robust for long-timescale MD simulations.

Using the ACE potential of NaF, we conducted long-timescale (up to 100 ns) isothermal-isobaric MD simulations starting from both bulk B1 and liquid phases at $P = 0$ and different T . 50 snapshots were selected at uniform time intervals from each of these simulations to assess energy and

force errors between DFT and ACE. As shown in Fig. 4(a), the absolute energy difference is smaller than 3 meV/atom and the RMSE of force is smaller than 85 meV/ $\text{\AA}$ for all of these snapshots. In addition, in order to ensure that the ACE potential is also accurate for supercooled liquids and solids formed from liquids at low T , we conducted 10 independent short (1 ns) isothermal-isobaric MD simulations starting from the liquid at each $T \leq 1200$ K. After 1 ns, these liquids may remain supercooled liquids or transform into crystals with more or less defects or glasses, according to the wide distribution of volume of the final snapshots of these short simulations [the upper panel of Fig. 4(a)]. We confirmed that the energy difference and RMSE of force of these final snapshots are also smaller than 3 meV/atom and 85 meV/ $\text{\AA}$, respectively, see the middle and lower of Fig. 4(a).

We increased hydrostatic P from 0 to 100 GPa for the bulk B1 crystal at $T = 350$ K and liquid at $T = 2050$ K using a constant rate of 5 MPa/ps. Both bulk B1 and liquid directly transformed into B2 crystal when P is high enough. We also decreased P from 100 to 0 GPa for the bulk B2 at two different T (350 and 2050 K). Bulk B2 transformed into B1 and liquid at 350 and 2050 K, respectively, when P is low enough. The absolute-energy difference and RMSE of force for the snapshots selected at uniform time intervals from these simulations are smaller than 5 meV/atom and 75 meV/ $\text{\AA}$, respectively, see Fig. 4(b).

When heating bulk B2, B3, B4, and H5 phases from 50 K to high T at $P = 0$ and increasing hydrostatic P on B3, B4, and H5 phases from 0 to high P at 50 K, all these metastable phases transformed into a B1 crystal when T or P is high enough. And the absolute energy difference and RMSE of force of the snapshots selected at uniform time intervals from these simulations are smaller than 5 meV/atom and 35 meV/ $\text{\AA}$, respectively, see Figs. 4(c) and 4(d).

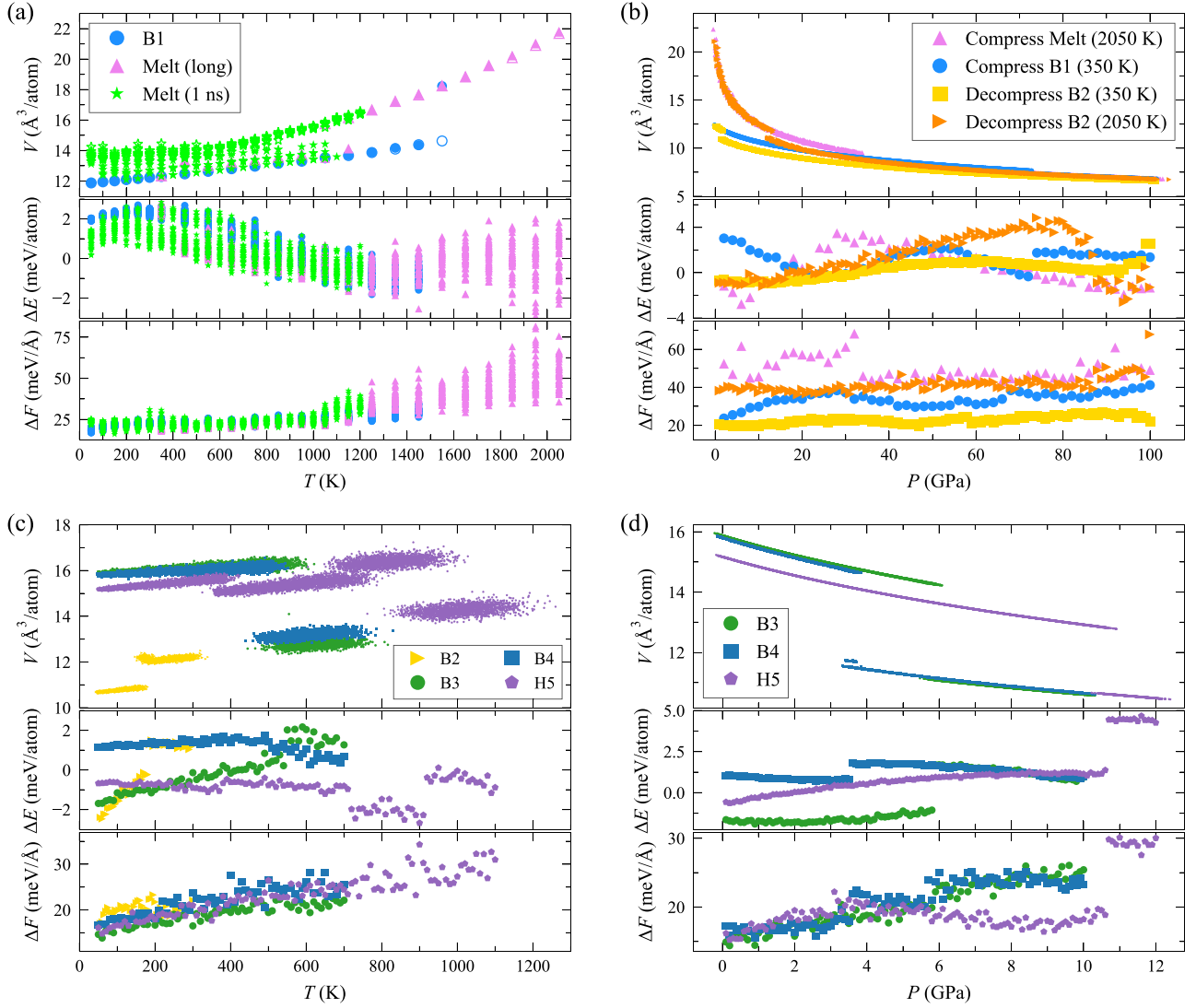


FIG. 4. Energy and force errors of the ACE potential for different phases of NaF at different T and P . The upper of (a) shows the initial (empty symbols) and final (solid symbols) volume of isothermal-isobaric MD simulations starting from the B1 and liquid phases at $P = 0$ and different T . At each T , a long-timescale (up to 100 ns) MD simulation starting from the B1 and liquid phases was conducted, and 10 independent short (1 ns) MD simulations starting from the liquid were conducted at each $T \leq 1200$ K. The middle and lower of (a) show the energy error ΔE and force error ΔF , respectively, of the 50 snapshots selected at uniform time intervals from each long MD simulation and the final snapshots of all short simulations. The upper of (b) shows the volume variation when increasing hydrostatic P on B1 and liquid phases and decreasing P on B2 phase at a rate of 5 MPa/ps. The upper of (c) shows the volume variation when heating B2, B3, B4, and H5 phases at $P = 0$ with a rate of 1×10^{10} K/s. The upper of (d) shows the volume variation when increasing hydrostatic P on B3, B4, and H5 phases at 50 K with a rate of 100 Pa/ps. The middle and lower of (b)–(d) show ΔE and ΔF , respectively, of the snapshots selected at uniform time intervals from the corresponding MD simulations. Here ΔE is the energy difference of each snapshot between DFT and ACE and ΔF is the root-mean-squared error of all force components in each snapshot between DFT and ACE. The sample size in the simulations starting from the liquid, B1, and B2 phases in (b) are 512, 512, and 432 atoms, respectively, and the sample size in the remaining simulations starting from the liquid, B1, B2, B3, B4, and H5 phases are 216, 216, 250, 216, 288, and 256 atoms, respectively.

For the ACE potential of KF, we also conducted similar isothermal-isobaric, heating, and compressing MD simulations starting from different phases and confirmed that the absolute-energy difference and RMSE of force of the snapshots from these simulations are smaller than 5 meV/atom and 100 meV/Å, respectively, when $T \leq 1600$ K and $P \leq 100$ GPa, as shown in Fig. 5.

In addition to the level of agreement documented above for both energies and forces between DFT and ACE, the liquid

structure [characterized using partial radial-density functions $g(r)$], liquid density, and lattice constant of B1 crystals are almost identical between DFT and ACE for both NaF and KF, as shown in Figs. 6(c) and 6(d) and Figs. 7(a) and 7(b).

B. Capturing the characteristic of ionic bonding

Since no explicit charges are included in the ACE models, one concern is whether the key characteristics of ionic

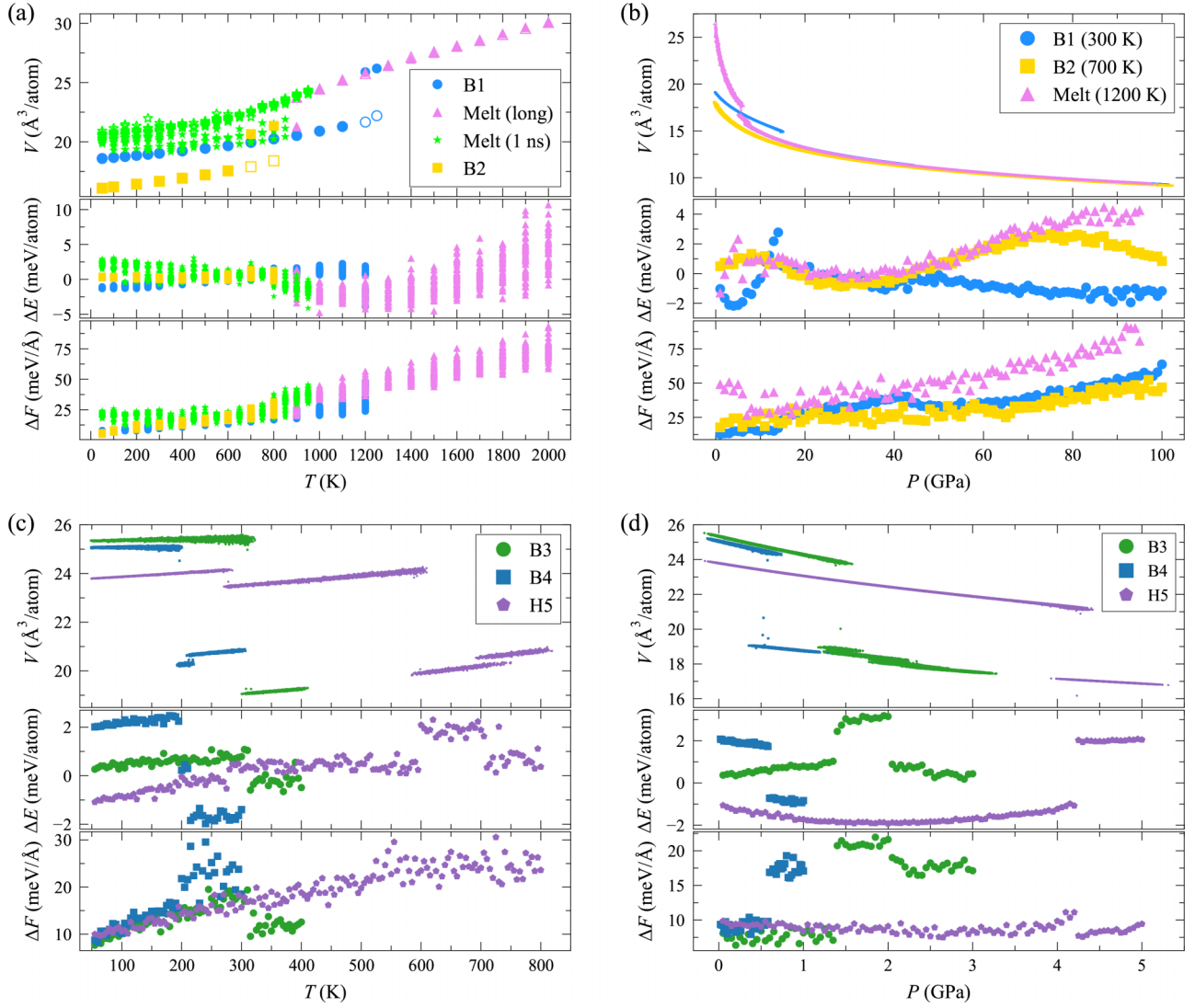


FIG. 5. Energy and force errors of the ACE potential for different phases of KF at different T and P . The upper of (a) shows the initial (empty symbols) and final (solid symbols) volume of isothermal-isobaric MD simulations starting from the B1, B2, and liquid phases at $P = 0$ and different T . At each T , a long-timescale (up to 100 ns) MD simulation starting from the B1, B2, and liquid phases was conducted, and 10 independent short (1 ns) MD simulations starting from the liquid were conducted at each $T \leq 950$ K. The middle and lower of (a) show the energy error ΔE and force error ΔF , respectively, of the 50 snapshots selected at uniform time intervals from each long MD simulation and the final snapshots of all short simulations. The upper of (b) shows the volume variation when increasing hydrostatic P on B1, B2, and liquid phases at a rate of 2.5 MPa/ps. The upper of (c) shows the volume variation when heating B3, B4, and H5 phases at $P = 0$ with a rate of 5×10^9 , 5×10^9 , and 1×10^{10} K/s, respectively. The upper of (d) shows the volume variation when increasing hydrostatic P on B3, B4 and H5 phases at 50 K with a rate of 50, 20 and 100 Pa/ps, respectively. The middle and lower of (b-d) show ΔE and ΔF , respectively, of the snapshots selected at uniform time intervals from the corresponding MD simulations. Here, ΔE is the energy difference of each snapshot between DFT and ACE and ΔF is the root-mean-squared error of all force components in each snapshot between DFT and ACE. The sample size in the simulations starting from the liquid, B1, and B2 phases in (b) are 512, 512, and 432 atoms, respectively, and the sample size in the remaining simulations starting from the liquid, B1, B2, B3, B4, and H5 phases are 216, 216, 250, 216, 288, and 256 atoms, respectively.

bonding of the two salts are captured. We visualized many atomic configurations of both solid and liquids of the two salts over wide T and P ranges and observed that the first-nearest-neighboring shell of cations (anions) are indeed occupied by anions (cations). As an example, Figs. 6(a) and 6(b) shows a liquid snapshot for NaF and KF, respectively. Furthermore, we verified that the location of the first peak of $g_{F-M}(r)$ (here M denotes alkali metal sodium or potassium) is much smaller than that of the first peak of $g_{F-F}(r)$ or $g_{M-M}(r)$ over wide

T and P ranges. Figures 6(c) and 6(d) show the $g(r)$ curves at the melting point for NaF and KF, respectively, as an example. These results corroborate that the key characteristic of ionic bonding is indeed captured by the ACE potentials.

C. Consistent well with experiments

As shown in Table II, the equilibrium melting point T_m of the ACE model for NaF and KF are higher by ~ 54 K and

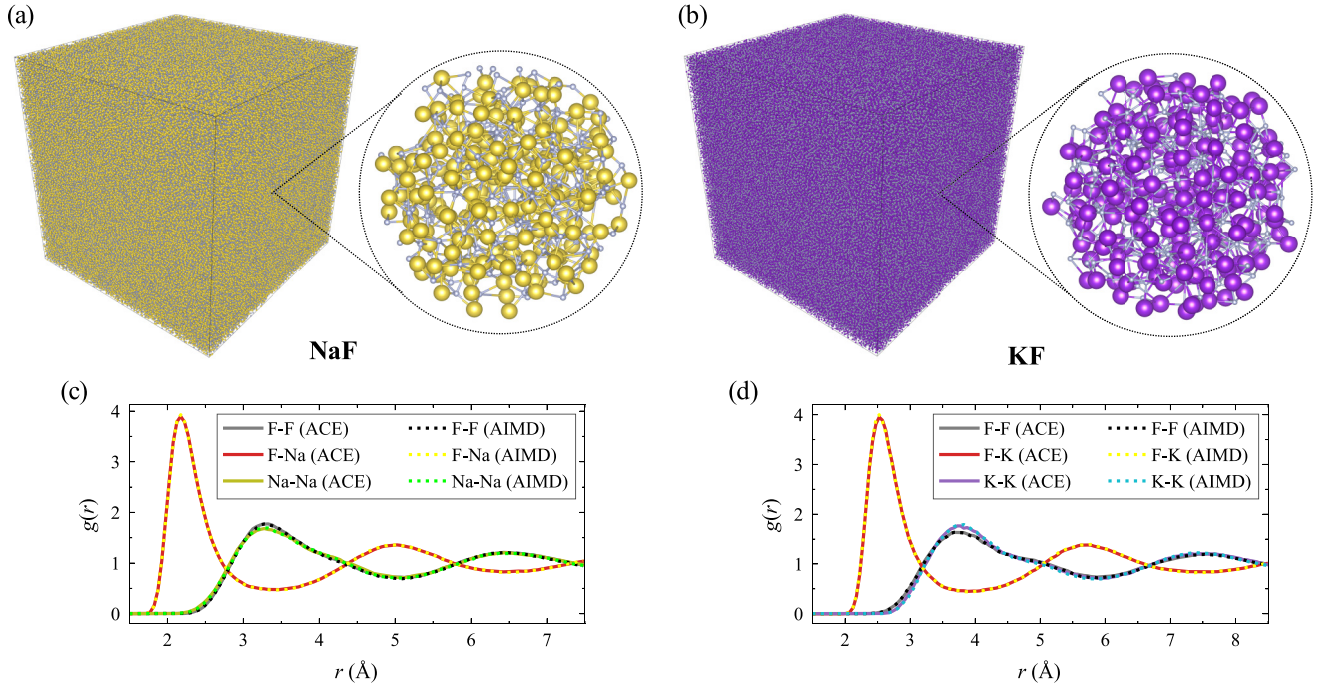


FIG. 6. (a) and (b) showcase a snapshot for NaF and KF melts, respectively, at their melting points. Each simulation box contains one million atoms. A closeup of a local configuration is shown for each salt to show that the first neighboring shell of cations (anions) are indeed occupied by anions (cations). (c) and (d) display the partial radial-distribution functions $g(r)$ of NaF and KF melts, respectively, at their melting points. $g(r)$ from both ACE-MD and AIMD are shown for comparison. In both ACE-MD and AIMD, the same sample size (216 atoms) and the same microcanonical (NVE) ensemble are used. The volume is set as the equilibrium volume determined from the ACE-MD simulations.

lower by ~ 58 K compared to the corresponding experimental values, respectively. The difference in T_m here is comparable to that of MLIPs for other materials [27,37,43] and may reflect limits in accuracy of the underlying DFT functional. Other equilibrium properties at the melting point, including the latent heat of melting (ΔH_m), entropy of melting (ΔS_m), and volume change upon melting ($\Delta V/V_s$) of the ACE potentials for NaF and KF are consistent to the corresponding experimental values, see Table II.

Figure 7(a) shows the temperature-dependent lattice constant of the B1 crystal at zero pressure for both NaF and KF obtained from MD simulations based on the ACE potentials, compared with experimental measurements. The largest deviation of these lattice constants compared to the available experimental data are less than 0.7% and 0.4% for NaF and KF, respectively. Figure 7(b) compares the temperature-dependent liquid density between MD simulations based on

the ACE potentials and experiments for both NaF and KF. The deviation in liquid density is $\sim 5\%$ for both NaF and KF. There is, however, scatter in the experimental data and for KF there is very good agreement with one set of experimental data. Also shown in Fig. 7(b) is a comparison between DFT and ACE potential predictions for liquid density, with the agreement being on the order of 97.5%.

Figure 7(c) compares the temperature-dependent self-diffusion coefficients of both molten NaF and KF salts from the MD simulations based on the ACE potentials and experimental data. The results show good correspondence within the uncertainty ranges. As shown in Fig. 7(d), the heat capacity C_p of B1 crystals of both NaF and KF from the MD simulations based on the ACE potentials agree well with the corresponding experimental data over the T range from around 600 K to T_m . At lower T , there is obvious deviation in C_p for crystals of both NaF and KF. This is because MD simulations cannot cap-

TABLE II. Comparison of equilibrium thermodynamic properties between MD simulations based on the ACE potentials and experiments for both NaF and KF.

	NaF		KF	
	Exp.	ACE	Exp.	ACE
T_m (K)	1268 ± 3 [65,66]	1322.1	1131 ± 3 [66,67]	1073.3
ΔH_m^a	8.03 [68], 7.78 [65]	8.11	6.75 [65], 7.05 [66]	6.35
ΔS_m^b	6.33, 6.14	6.13	5.97, 6.23	5.92
$\Delta V/V_s$	$24.0\% \pm 2.0\%$ [66]	21.6%	$17.2\% \pm 3.0\%$ [66]	18.1%

^aThe unit of ΔH_m is one kilocalorie per mole.

^bThe unit of ΔS_m is one calorie per kelvin per mole.

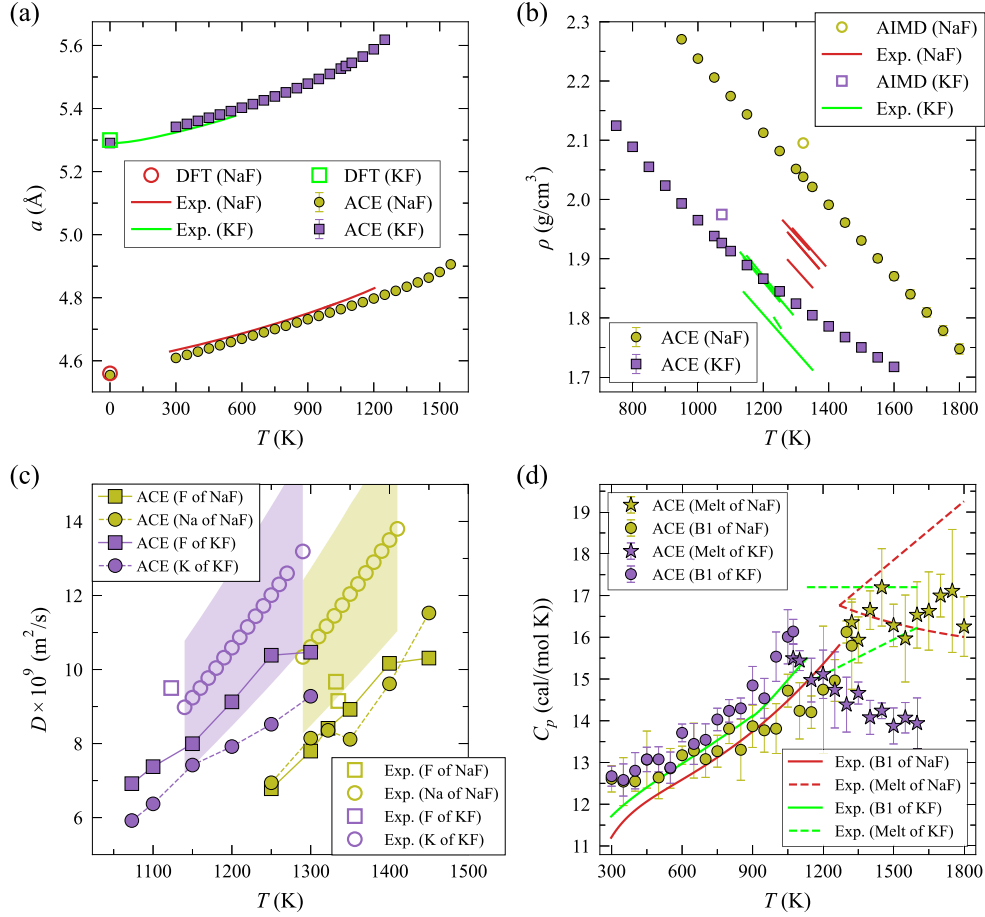


FIG. 7. Comparison of various properties between MD simulations based on the ACE potentials developed here and real experiments for both NaF and KF. (a) Temperature-dependent lattice constants a of rocksalt crystals from both MD simulations and experiments as well as lattice constants at 0 K determined from DFT calculations. The experimental data of NaF is from Ref. [69], and the experimental data of KF is determined using the lattice constant at 0 K [70] and thermal expansion coefficients over the T range from 70 to 570 K [71]. (b) Temperature-dependent liquid density ρ . For each of the two salts, five sets of experimental data collected in Refs. [66,72] as well as the values at T_m determined from AIMD simulations are presented. (c) Temperature-dependent self-diffusion coefficients D for the two species in each of the two molten salts. The single experimental data for F in KF melt is from Ref. [73] while the remaining experimental data is from Ref. [74]. The uncertainty of experimental data for both Na in NaF melt and K in KF melt is around $\pm 20\%$, indicated using the shadow areas. (d) Temperature-dependent heat capacity C_p for both rocksalt crystals (B1 phases) and melts. The experimental data for crystals are from NIST Chemistry WebBook [75,76], while for melts, an additional set of experimental data from Ref. [66] is presented for both NaF and KF, in addition to the data from NIST [75,76].

ture the quantum effect in the occupancy of the phonon modes below the Debye temperature. As for C_p of liquid phase, we found two sets of experimental data that differ from each other for both NaF and KF melts. The C_p of NaF melt from the MD simulations based on the ACE potential is closest to the experimental data reported on the NIST Chemistry WebBook [75], while the C_p of KF melt from ACE-MD simulations is different from the two sets of experimental data, see Fig. 7(d). Further work is warranted to determine whether these differences originate from inaccuracy of the DFT or uncertainties in the experimental measurements.

D. Efficiency of ACE potentials

Figure 8 shows the CPU time required to run 10 000 MD timesteps using a single CPU node (128 cores) on NERSC for both NaF and KF melts of different sample sizes up to 1

million atoms at their melting points with the ACE potentials. This means that it only takes around 97 and 70 hours when using a single CPU node to run 1 ns MD simulations containing 1 million atoms for NaF and KF, respectively. Thus, it is feasible to use multiple CPU nodes to conduct long-timescale and large-scale MD simulations with the two ACE potentials, which is desirable especially for studying crystal nucleation and growth of molten salts. The excellent efficiency of the two ACE potentials is due to the linear formalism of ACE models as well as the relatively short cutoff we have chosen in defining local atomic environments.

IV. SUMMARY

In summary, we report the development of ACE potentials for both NaF and KF. The two ACE potentials are demonstrated to possess DFT-SCAN-level accuracy for liquid and

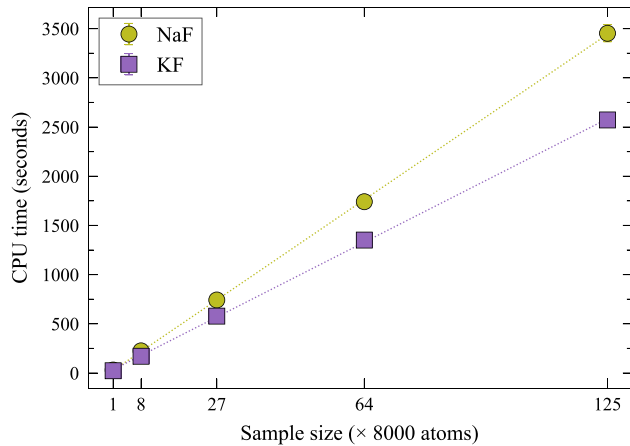


FIG. 8. The CPU time required to run 10 000 MD timesteps using a single CPU node (128 cores) on NERSC for both NaF and KF melts of different sample sizes at their melting points with the ACE potentials. For each sample size, five independent runs were conducted to determine the standard deviation, which is smaller than the symbol size.

five solid polymorphs over wide T and P ranges and correctly capture the key characteristic of ionic bonding, although charge is not included in the development of the two ACE potentials. More importantly, MD simulations based on the two ACE potentials can well reproduce experimental data

for multiple properties. Furthermore, it is feasible to apply the two ACE potentials to conduct long-timescale and large-scale MD simulations thanks to the excellent efficiency of the two potentials. The accurate and efficient ACE potentials for both NaF and KF are anticipated to find application in MD simulations for different purposes, such as establishing structure-property relationships for molten salts and studying polymorphic transitions and melting of alkali halides, in addition to understanding crystallization of molten salts.

ACKNOWLEDGMENTS

This work is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences through the Science of Direct MINeral to Energy Storage Synthesis (MINES) program at the Lawrence Berkeley National Laboratory (LBNL) under Contract No. DE-AC02-05-CH11231 (MINES project, FWP: FP00014914). This research used resources of the National Energy Research Scientific Computing Center (NERSC), a Department of Energy Office of Science User Facility using NERSC Awards No. BES-ERCAP0025752, No. BES-ERCAP0027659 and No. BES-ERCAP0032361.

DATA AVAILABILITY

The two ACE potentials and associated training datasets are openly available [77].

- [1] C. Le Brun, Molten salts and nuclear energy production, *J. Nucl. Mater.* **360**, 1 (2007).
- [2] J. Serp, M. Allibert, O. Beneš, S. Delpech, O. Feynberg, V. Ghetta, D. Heuer, D. Holcomb, V. Ignatiev, J. L. Kloosterman *et al.*, The molten salt reactor (MSR) in generation IV: Overview and perspectives, *Prog. Nucl. Energy* **77**, 308 (2014).
- [3] U. Pelay, L. Luo, Y. Fan, D. Stitou, and M. Rood, Thermal energy storage systems for concentrated solar power plants, *Renewable Sustainable Energy Rev.* **79**, 82 (2017).
- [4] T.-C. Ong, M. Sarvghad, K. Lippiatt, L. Griggs, H. Ryan, G. Will, and T. A. Steinberg, Review of the solubility, monitoring, and purification of impurities in molten salts for energy storage in concentrated solar power plants, *Renewable Sustainable Energy Rev.* **131**, 110006 (2020).
- [5] J. C. Ondry, Z. Zhou, K. Lin, A. Gupta, J. H. Chang, H. Wu, A. Jeong, B. F. Hammel, D. Wang, H. C. Fry *et al.*, Reductive pathways in molten inorganic salts enable colloidal synthesis of III-V semiconductor nanocrystals, *Science* **386**, 401 (2024).
- [6] Z. Lun, B. Ouyang, D.-H. Kwon, Y. Ha, E. E. Foley, T.-Y. Huang, Z. Cai, H. Kim, M. Balasubramanian, Y. Sun *et al.*, Cation-disordered rocksalt-type highentropy cathodes for Li-ion batteries, *Nat. Mater.* **20**, 214 (2021).
- [7] N. W. Dunbar, G. K. Jacobs, and M. T. Naney, Crystallization processes in an artificial magma: variations in crystal shape, growth rate and composition with melt cooling history, *Contrib. Mineral. Petrol.* **120**, 412 (1995).
- [8] A. Hawken, G. Sun, and P. Harrowell, Role of interfacial inherent structures in the fast crystal growth from molten salts and metals, *Phys. Rev. Mater.* **3**, 043401 (2019).
- [9] C. Valeriani, E. Sanz, and D. Frenkel, Rate of homogeneous crystal nucleation in molten NaCl, *J. Chem. Phys.* **122**, 194501 (2005).
- [10] B. A. Frandsen, S. D. Nickerson, A. D. Clark, A. Solano, R. Baral, J. Williams, J. Neufeind, and M. Memmott, The structure of molten flinak, *J. Nucl. Mater.* **537**, 152219 (2020).
- [11] N. Winner, H. Williams, R. O. Scarlat, and M. Asta, *Ab-initio* simulation studies of chromium solvation in molten fluoride salts, *J. Mol. Liq.* **335**, 116351 (2021).
- [12] D. Andersson and B. W. Beeler, *Ab initio* molecular dynamics (AIMD) simulations of NaCl, UCl₃ and NaCl-UCl₃ molten salts, *J. Nucl. Mater.* **568**, 153836 (2022).
- [13] K. Duemmler, M. Woods, T. Karlsson, R. Gakhar, and B. Beeler, First-principles-derived transport properties of molten chloride salts, *J. Nucl. Mater.* **585**, 154601 (2023).
- [14] M. Tosi and F. Fumi, Ionic sizes and born repulsive parameters in the NaCl-type alkali halides—XII: The generalized Huggins-Mayer form, *J. Phys. Chem. Solids* **25**, 45 (1964).
- [15] T. Koishi, K. Yasuoka, and T. Ebisuzaki, Large scale molecular dynamics simulation of nucleation in supercooled NaCl, *J. Chem. Phys.* **119**, 11298 (2003).
- [16] A. Galashev, Molecular dynamics study of ionic diffusion and the FLiNaK salt melt structure, *Nuclear Engineering and Technology* **55**, 1324 (2023).
- [17] Q. Yang, L. Sang, and J. Huang, Molecular dynamics simulation on thermal and transport properties of LiF-KF, *J. Mol. Liq.* **410**, 125627 (2024).

- [18] R. Wang, S. Mehdi, Z. Zou, and P. Tiwary, Is the local ion density sufficient to drive nacl nucleation from the melt and aqueous solution? *J. Phys. Chem. B* **128**, 1012 (2024).
- [19] M. Sangster and M. Dixon, Interionic potentials in alkali halides and their use in simulations of the molten salts, *Adv. Phys.* **25**, 247 (1976).
- [20] M. Wilson and P. Madden, Polarization effects in ionic systems from first principles, *J. Phys.: Condens. Matter* **5**, 2687 (1993).
- [21] S.-C. Lee, Y. Zhai, Z. Li, N. P. Walter, M. Rose, B. J. Heuser, and Y. Z., Comparative studies of the structural and transport properties of molten salt FLiNaK using the machine-learned neural network and reparametrized classical forcefields, *J. Phys. Chem. B* **125**, 10562 (2021).
- [22] J. Behler and M. Parrinello, Generalized neuralnetwork representation of high-dimensional potentialenergy surfaces, *Phys. Rev. Lett.* **98**, 146401 (2007).
- [23] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons, *Phys. Rev. Lett.* **104**, 136403 (2010).
- [24] A. V. Shapeev, Moment tensor potentials: A class of systematically improvable interatomic potentials, *Multiscale Modeling & Simulation* **14**, 1153 (2016).
- [25] L. Zhang, J. Han, H. Wang, R. Car, and W. E, Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics, *Phys. Rev. Lett.* **120**, 143001 (2018).
- [26] R. Drautz, Atomic cluster expansion for accurate and transferable interatomic potentials, *Phys. Rev. B* **99**, 014104 (2019).
- [27] L. Bonati and M. Parrinello, Silicon liquid structure and crystal nucleation from *ab initio* deep metadynamics, *Phys. Rev. Lett.* **121**, 265701 (2018).
- [28] Z. Fan and H. Tanaka, Microscopic mechanisms of pressure-induced amorphous-amorphous transitions and crystallisation in silicon, *Nat. Commun.* **15**, 368 (2024).
- [29] Y. Wang, J. Ding, Z. Fan, L. Tian, M. Li, H. Lu, Y. Zhang, E. Ma, J. Li, and Z. Shan, Tension-compression asymmetry in amorphous silicon, *Nat. Mater.* **20**, 1371 (2021).
- [30] S. Yin, Y. Zuo, A. Abu-Odeh, H. Zheng, X.-G. Li, J. Ding, S. P. Ong, M. Asta, and R. O. Ritchie, Atomistic simulations of dislocation mobility in refractory high-entropy alloys and the effect of chemical short-range order, *Nat. Commun.* **12**, 4873 (2021).
- [31] B. Zhang, E. Chen, and M. Asta, Oxygen grain-boundary segregation in HCP Ti—computational investigations using an atomic cluster expansion potential, *Comput. Mater. Sci.* **248**, 113577 (2025).
- [32] W. Wang, P. Kumar, D. H. Cook, F. Walsh, B. Zhang, P. P. Borges, D. Farkas, R. O. Ritchie, and M. Asta, Ductility mechanisms in complex concentrated refractory alloys from atomistic fracture simulations, *arXiv:2502.17750*.
- [33] S. Tovey, A. Narayanan Krishnamoorthy, G. Sivaraman, J. Guo, C. Benmore, A. Heuer, and C. Holm, DFT Accurate interatomic potential for molten NaCl from machine learning, *J. Phys. Chem. C* **124**, 25760 (2020).
- [34] A. Rodriguez, S. Lam, and M. Hu, Thermodynamic and transport properties of lif and flibe molten salts with deep learning potentials, *ACS Applied Materials & Interfaces* **13**, 55367 (2021).
- [35] S. T. Lam, Q.-J. Li, R. Ballinger, C. Forsberg, and J. Li, Modeling lif and flibe molten salts with robust neural network interatomic potential, *ACS Applied Materials & Interfaces* **13**, 24582 (2021).
- [36] G. Sivaraman, J. Guo, L. Ward, N. Hoyt, M. Williamson, I. Foster, C. Benmore, and N. Jackson, Automated development of molten salt machine learning potentials: Application to LiCl, *J. Phys. Chem. Lett.* **12**, 4278 (2021).
- [37] Q.-J. Li, E. Küçükbenli, S. Lam, B. Khaykovich, E. Kaxiras, and J. Li, Development of robust neural-network interatomic potential for molten salt, *Cell Reports Physical Science* **2**, 100359 (2021).
- [38] J. Guo, L. Ward, Y. Babuji, N. Hoyt, M. Williamson, I. Foster, N. Jackson, C. Benmore, and G. Sivaraman, Compositiontransferable machine learning potential for LiCl-KCl molten salts validated by high-energy x-ray diffraction, *Phys. Rev. B* **106**, 014209 (2022).
- [39] S. Attarian, D. Morgan, and I. Szlufarska, Studies of Ni-Cr complexation in FLiBe molten salt using machine learning interatomic potentials, *J. Mol. Liq.* **400**, 124521 (2024).
- [40] N. Rybin, D. Maksimov, Y. Zaikov, and A. Shapeev, Thermophysical properties of molten FLiNaK: A moment tensor potential approach, *J. Mol. Liq.* **410**, 125402 (2024).
- [41] S. Attarian, C. Shen, D. Morgan, and I. Szlufarska, Best practices for fitting machine learning interatomic potentials for molten salts: A case study using NaCl-MgCl₂, *Comput. Mater. Sci.* **246**, 113409 (2025).
- [42] C. Shen, S. Attarian, Y. Zhang, H. Zhang, M. Asta, I. Szlufarska, and D. Morgan, Supersalt: equivariant neural network force fields for multicomponent molten salts system, *Nat. Commun.* **16**, 7280 (2025).
- [43] Z. Fan, D. Rawat, P. Zarzycki, M. L. Whittaker, and M. Asta, Multistage nucleation pathway in lif molten salt mirrors the crystal-melt interface structure, *Proc. Natl. Acad. Sci. USA* **122**, e2425702122 (2025).
- [44] R. Drautz, Atomic cluster expansion of scalar, vectorial, and tensorial properties including magnetism and charge transfer, *Phys. Rev. B* **102**, 024104 (2020).
- [45] A. Bochkarev, Y. Lysogorskiy, S. Menon, M. Qamar, M. Mrovec, and R. Drautz, Efficient parametrization of the atomic cluster expansion, *Phys. Rev. Mater.* **6**, 013804 (2022).
- [46] Y. Lysogorskiy, A. Bochkarev, M. Mrovec, and R. Drautz, Active learning strategies for atomic cluster expansion models, *Phys. Rev. Mater.* **7**, 043801 (2023).
- [47] Y. Lysogorskiy, C. van der Oord, A. Bochkarev, S. Menon, M. Rinaldi, T. Hammerschmidt, M. Mrovec, A. Thompson, G. Csányi, C. Ortner *et al.*, Performant implementation of the atomic cluster expansion (pace) and application to copper and silicon, *npj Comput. Mater.* **7**, 97 (2021).
- [48] J. Sun, A. Ruzsinszky, and J. P. Perdew, Strongly constrained and appropriately normed semilocal density functional, *Phys. Rev. Lett.* **115**, 036402 (2015).
- [49] Y. Sato-Sorensen, Phase transitions and equations of state for the sodium halides: NaF NaCl, NaBr, and NaI, *J. Geophys. Res. Solid Earth* **88**, 3543 (1983).
- [50] I. Ruff, A. Baranyai, E. Spohr, and K. Heinzinger, Isothermal-isobaric molecular dynamics simulation of polymorphic phase transitions in alkali halides, *J. Chem. Phys.* **91**, 3148 (1989).
- [51] T. Yagi, T. Suzuki, and S.-I. Akimoto, New high-pressure polymorphs in sodium halides, *J. Phys. Chem. Solids* **44**, 135 (1983).

- [52] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/xbfm-clgd> for the details of generating training and test datasets.
- [53] G. Kresse, J. Furthmüller, and J. Hafner, Theory of the crystal structures of selenium and tellurium: the effect of generalized-gradient corrections to the local-density approximation, *Phys. Rev. B* **50**, 13181 (1994).
- [54] G. Kresse and J. Furthmüller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set, *Phys. Rev. B* **54**, 11169 (1996).
- [55] G. Kresse and J. Furthmüller, Efficiency of *ab initio* total energy calculations for metals and semiconductors using a plane-wave basis set, *Comput. Mater. Sci.* **6**, 15 (1996).
- [56] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. In't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen *et al.*, LAMMPS-A flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, *Comput. Phys. Commun.* **271**, 108171 (2022).
- [57] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO—the open visualization tool, *Modell. Simul. Mater. Sci. Eng.* **18**, 015012 (2010).
- [58] K. Momma and F. Izumi, Vesta: a three-dimensional visualization system for electronic and structural analysis, *J. Appl. Crystallogr.* **41**, 653 (2008).
- [59] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, 2017).
- [60] Y. Yang, H. Humadi, D. Buta, B. B. Laird, D. Sun, J. J. Hoyt, and M. Asta, Atomistic simulations of nonequilibrium crystal-growth kinetics from alloy melts, *Phys. Rev. Lett.* **107**, 025505 (2011).
- [61] J. R. Morris, C. Wang, K. Ho, and C. T. Chan, Melting line of aluminum from simulations of coexisting phases, *Phys. Rev. B* **49**, 3109 (1994).
- [62] D. Alfè, M. Gillan, and G. Price, Complementary approaches to the *ab initio* calculation of melting properties, *J. Chem. Phys.* **116**, 6170 (2002).
- [63] H. Ramalingam, M. Asta, A. Van de Walle, and J. Hoyt, Atomic-scale simulation study of equilibrium solute adsorption at alloy solid-liquid interfaces, *Interface Sci.* **10**, 149 (2002).
- [64] J. J. Hoyt, J. W. Garvin, E. B. Webb III, and M. Asta, An embedded atom method interatomic potential for the Cu–Pb system, *Modell. Simul. Mater. Sci. Eng.* **11**, 287 (2003).
- [65] A. Dworkin and M. Bredig, The heat of fusion of the alkali metal halides, *J. Phys. Chem.* **64**, 269 (1960).
- [66] G. J. Janz, C. B. Allen, N. Bansal, R. Murphy, and R. Tomkins, Physical properties data compilations relevant to energy storage, II. Molten salts: Data on single and multi-component salt systems, Nasa Sti/recon Technical Report N **80**, 10643 (1979).
- [67] J. Johnson and M. Bredig, Miscibility of metals with salts in the molten state. III. The potassium-potassium halide systems, *J. Phys. Chem.* **62**, 604 (1958).
- [68] C. O'Brien and K. Kelley, High temperature heat contents of cryolite, anhydrous aluminum fluoride and sodium fluoride, *J. Am. Chem. Soc.* **79**, 5616 (1957).
- [69] P. Pathak, J. Trivedi, and N. Vasavada, Thermal expansion of NaF, KBr and RbBr and temperature variation of the frequency spectrum of NaF, *Foundations of Crystallography* **29**, 477 (1973).
- [70] M. Prencipe, A. Zupan, R. Dovesi, E. Apra, and V. Saunders, *Ab initio* study of the structural properties of LiF, NaF, KF, LiCl, NaCl, and KCl, *Phys. Rev. B* **51**, 3391 (1995).
- [71] J. E. Rapp and H. Merchant, Thermal expansion of alkali halides from 70 to 570 K, *J. Appl. Phys.* **44**, 3919 (1973).
- [72] G. J. Janz, Thermodynamic and transport properties for molten salts: correlation equations for critically evaluated density, surface tension, electrical conductance, and viscosity data, *J. Phys. Chem. Ref. Data* **17** (1988).
- [73] V. Sarou-Kanian, A.-L. Rollet, M. Salanne, C. Simon, C. Bessada, and P. A. Madden, Diffusion coefficients and local structure in basic molten fluorides: *In situ* NMR measurements and molecular dynamics simulations, *Phys. Chem. Chem. Phys.* **11**, 11501 (2009).
- [74] G. J. Janz and N. Bansal, Molten salts data: diffusion coefficients in single and multi-component salt systems, *J. Phys. Chem. Ref. Data* **11**, 505 (1982).
- [75] <https://webbook.nist.gov/cgi/cbook.cgi?ID=C7681494&Mask=6F>.
- [76] <https://webbook.nist.gov/cgi/cbook.cgi?ID=C7789233&Mask=1E9F>.
- [77] Z. Fan, Figshare (2025), <https://doi.org/10.6084/m9.figshare.29695496>.