

# Exploring phononic properties of two-dimensional materials using machine learning interatomic potentials

Bohayra Mortazavi<sup>a,b,\*</sup>, Ivan S. Novikov<sup>c,d</sup>, Evgeny V. Podryabinkin<sup>c</sup>, Stephan Roche<sup>e,f</sup>, Timon Rabczuk<sup>g</sup>, Alexander V. Shapeev<sup>c,\*</sup>, Xiaoying Zhuang<sup>a,g</sup>

<sup>a</sup> Chair of Computational Science and Simulation Technology, Department of Mathematics and Physics, Leibniz Universität Hannover, Appelstraße 11, 30157 Hannover, Germany

<sup>b</sup> Cluster of Excellence PhoenixD (Photonics, Optics, and Engineering–Innovation Across Disciplines), Gottfried Wilhelm Leibniz Universität Hannover, Hannover, Germany

<sup>c</sup> Skolkovo Institute of Science and Technology, Skolkovo Innovation Center, Nobel St. 3, Moscow, 143026, Russia

<sup>d</sup> Institute of Materials Science, University of Stuttgart, Pfaffenwaldring 55, Stuttgart, 70569, Germany

<sup>e</sup> Catalan Institute of Nanoscience and Nanotechnology (ICN2), CSIC and BIST, Campus UAB, Bellaterra, Barcelona, 08193, Spain

<sup>f</sup> ICREA Institució Catalana de Recerca i Estudis Avançats, Barcelona, 08010, Spain

<sup>g</sup> College of Civil Engineering, Department of Geotechnical Engineering, Tongji University, Shanghai, China

## ARTICLE INFO

### Article history:

Received 6 March 2020

Revised 29 April 2020

Accepted 1 May 2020

### Keywords:

Machine-learning  
Interatomic potentials  
Phononic properties  
2D materials

## ABSTRACT

Phononic properties are commonly studied by calculating force constants using the density functional theory (DFT) simulations. Although DFT simulations offer accurate estimations of phonon dispersion relations or thermal properties, but for low-symmetry and nanoporous structures the computational cost quickly becomes very demanding. Moreover, the computational setups may yield nonphysical imaginary frequencies in the phonon dispersion curves, impeding the assessment of phononic properties and the dynamical stability of the considered system. Here, we compute phonon dispersion relations and examine the dynamical stability of a large ensemble of novel materials and compositions. We propose a fast and convenient alternative to DFT simulations which derived from machine-learning interatomic potentials passively trained over computationally efficient ab-initio molecular dynamics trajectories. Our results for diverse two-dimensional (2D) nanomaterials confirm that the proposed computational strategy can reproduce fundamental thermal properties in close agreement with those obtained via the DFT approach. The presented method offers a stable, efficient, and convenient solution for the examination of dynamical stability and exploring the phononic properties of low-symmetry and porous 2D materials.

© 2020 Elsevier Ltd. All rights reserved.

## 1. Introduction

Phonon dispersion relations (PDRs) are key components to study the lattice dynamics and atomic vibrations in a crystal [1,2]. They also provide useful information regarding the transport properties such as the thermal conductivity. In particular, PDRs are extensively employed to examine the dynamical stability of various compositions and structures. To acquire PDRs, the most popular theoretical approach is to calculate the force constants using the density functional theory (DFT) simulations. To assess the thermal properties, DFT calculations are commonly executed using

supercell structures. DFT is a computationally efficient approach for the majority of highly symmetrical lattices. In recent years, two-dimensional (2D) materials [3,4] are gaining remarkable attentions because of their unique properties, suitable to address critical challenges in various advanced technologies like nanoelectronics, energy harvesting devices or biosensors. Since the isolation of graphene [1,2], 2D materials family has been continuously and quickly extending. This large family of new materials includes highly symmetric lattices, such as the graphene-like structures [5,6] and 2H transition metal dichalcogenides [7,8], as well as many structures with low-symmetry or nanoporous lattices like graphdiyne [9] and 1T' transition metal dichalcogenides [10]. To date, high-symmetry and isotropic lattices have been excessively studied. Recently, however, anisotropic and nanoporous 2D lattices are attracting considerable attention, because they offer new possibilities to achieve angle-dependent properties and design more efficient energy storage/conversion systems.

\* Corresponding author at: Chair of Computational Science and Simulation Technology, Department of Mathematics and Physics, Leibniz Universität Hannover, Appelstraße 11, 30157 Hannover, Germany.

E-mail addresses: [bohayra.mortazavi@gmail.com](mailto:bohayra.mortazavi@gmail.com) (B. Mortazavi), [a.shapeev@skoltech.ru](mailto:a.shapeev@skoltech.ru) (A.V. Shapeev), [zhuang@hot.uni-hannover.de](mailto:zhuang@hot.uni-hannover.de) (X. Zhuang).

In order to theoretically predict novel 2D systems, it is thus essential to carefully examine the dynamical stability and explore the vibrational properties on the basis of PDRs. For low-symmetry and nanoporous 2D structures, currently employed DFT simulations exhibit limited flexibility due to severe computational issues. In these cases, to bypass the computational limitations, one usually treats smaller supercells and lower k-point grids and plane-wave cutoff energy within the DFT simulations. But such simplifications may actually result in poorer accuracy of the acquired PDRs. Moreover, in some cases nonphysical imaginary frequencies may appear in the PDRs thus impeding the examination of dynamical stability or thermal properties. Machine-learning based methods are currently among the most attractive solutions to address various challenges, also in materials science [11–17]. In this regard, machine-learning interatomic potentials (MLIP) [18] have shown outstanding efficiency in many applications of computational material science, such as predicting novel materials [19,20] lattice dynamics [21] and estimating the thermal conductivity [22,23], etc. The main advantage of MLIPs is that they enable the efficient use of classical molecular dynamics simulations to evaluate the forces and energies with the DFT level of accuracy. The employment of MLIPs is particularly promising to study the large systems, for those DFT-based methods become infeasible due to the known computational limitations.

In this work we propose a computationally efficient and accurate methodology to acquire PDRs and explore other critical phononic properties on the basis of passively fitted moment tensor potentials (MTPs) [24]. MTP belongs to the family of MLIPs [25–27] by which we understand potentials possessing flexible functional form that allows for systematically increasing of the accuracy with an increase in the number of parameters and the size of the training [28]. Similar to other MLIPs, MTP is able to approximate any model of interatomic interaction, offering a compromise choice between accuracy and computational efficiency. The proposed approach only requires inexpensive and short ab-initio molecular dynamics trajectories, without the need in active learning or additional DFT calculations. We employed such an approach to evaluate the phononic properties of a wide variety of 2D materials, for which close agreements with density functional perturbation theory (DFPT) results are observed. We describe our method on the basis of passively-trained MTPs, which can be conveniently employed to explore the phononic properties of low-symmetry and nanoporous 2D materials, with enhanced stability and computational efficiency.

## 2. Computational methods

First-principles DFT calculations in this work were carried out using the *Vienna Ab-initio Simulation Package* (VASP) [29–31]. The generalized gradient approximation (GGA), namely the Perdew–Burke–Ernzerhof (PBE) [32] functional was employed in the calculations. We assumed the plane-wave cutoff energies of 600 eV and 400 eV for the carbon-based and carbon-free systems, respectively. For geometry optimization, the convergence tolerances for the energy and forces were set to  $10^{-5}$  eV and 0.001 eV/Å, respectively. The DFT simulations were performed over supercell samples using a  $3 \times 3 \times 1$  Monkhorst-Pack [33] k-point grid. The plane-wave cutoff energy of the DFPT simulations was set as the default value by VASP. The PHONOPY code [34] was utilized to create the optimal sets of atomic position for DFPT calculations and also to acquire phonon dispersions and group velocities with the DFPT results as inputs. Ab-initio molecular dynamics (AIMD) simulations were performed with a time step of 1 fs using a  $3 \times 3 \times 1$  k-point grid.

A class of machine-learning interatomic potentials–moment tensor potentials [24] were used to describe interatomic interac-

tions. Similar to classical potentials, MTPs include parameters that are optimized on a set of training configurations. In this work AIMD simulations were used to create the training sets. MTP was first proposed for single-component systems [24] and has been recently generalized to multiple component systems [19,35]. This potential is local, i.e., the total energy  $E$  of the system containing  $N$  atoms is partitioned into contributions  $V$  of neighborhoods  $u_i$  of each  $i$ -th atom:  $E \equiv E^{MTP} = \sum_{i=1}^N V(u_i)$ . We refer to the  $j$ -th atom as a neighbor of the  $i$ -th (central) atom if the distance between them is less than a predefined cut-off distance  $R_{cut}$ . The neighborhood is then expressed as a collection,

$$u_i = (\{r_{i1}, z_i, z_1\} \dots, \{r_{ij}, z_i, z_j\} \dots, \{r_{iN_{nei}}, z_i, z_{N_{nei}}\}), \quad (1)$$

where  $r_{ij}$  are the relative atomic positions (interatomic vector),  $z_i$  and  $z_j$  are the types of the central and neighboring atoms, respectively, and  $N_{nei}$  is the number of atoms in neighborhood. Each contribution to the total energy has the following form:  $V(u_i) = \sum_{\alpha} \xi_{\alpha} B_{\alpha}(u_i)$ , where  $\xi_{\alpha}$  are the free parameters of the potential to be optimized, and  $B_{\alpha}$  are the basis functions. We construct the basis functions as all possible contractions of the moment tensor descriptors:

$$M_{\mu, \nu}(r_i) = \sum_{j=1}^{N_{nei}} f_{\mu}(|r_{ij}|, z_i, z_j) r_{ij}^{\otimes \nu} \quad (2)$$

yielding a scalar (see Ref. [24] for details). The first factor  $f_{\mu}(|r_{ij}|, z_i, z_j)$  in the aforementioned summation is the radial part depending only on the distance between atoms  $i$  and  $j$  and their types. We expand the radial part through a set of radial basis functions  $\varphi_{\beta}(|r_{ij}|)$  multiplied by a factor  $(R_{cut} - |r_{ij}|)^2$  for smoothing near the cut-off radius.

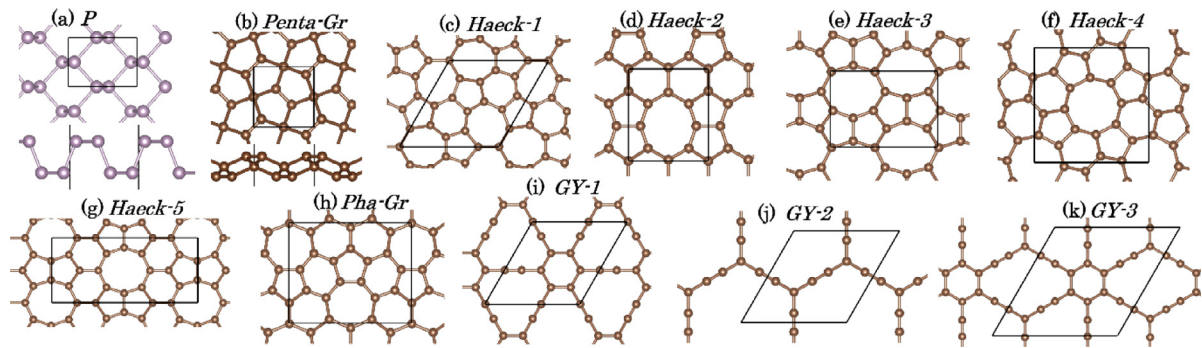
$$f_{\mu}(|r_{ij}|, z_i, z_j) = c_{\mu, z_i, z_j}^{(\beta)} \varphi_{\beta}(|r_{ij}|) (R_{cut} - |r_{ij}|)^2, \quad (3)$$

where  $c_{\mu, z_i, z_j}^{(\beta)}$  are the radial coefficients. We denote by “ $\otimes$ ” the outer product and refer to the second factor in Eq. 2 as the angular part. In order to optimize the  $\xi_{\alpha}$  and  $c_{\mu, z_i, z_j}^{(\beta)}$  parameters of an MTP, one needs to solve the following minimization problem (training of MTP):

$$\sum_{k=1}^K \left[ w_e (E_k^{AIMD} - E_k^{MTP})^2 + w_f \sum_i^N |f_{k,i}^{AIMD} - f_{k,i}^{MTP}|^2 + w_s \sum_{i,j=1}^3 |\sigma_{k,i,j}^{AIMD} - \sigma_{k,i,j}^{MTP}|^2 \right] \rightarrow \min, \quad (4)$$

where  $E_k^{AIMD}$ ,  $f_{k,i}^{AIMD}$  and  $\sigma_{k,i,j}^{AIMD}$  are the energy, atomic forces and stresses in the training set, respectively, and  $E_k^{MTP}$ ,  $f_{k,i}^{MTP}$  and  $\sigma_{k,i,j}^{MTP}$  are the corresponding values calculated with the MTP,  $K$  is the number of the configurations in the training set, and  $w_e$ ,  $w_f$  and  $w_s$  are the non-negative weights that express the importance of energies and forces and stresses in the optimization problem, respectively, which in our study were set to 1, 0.1 and 0.001, respectively. We note that the weights for the energy and force are the default values.

Here we used the PHONOPY code [34] to evaluate the phononic properties, in which MTP replaces VASP in the force calculation step [36]. To facilitate future studies, in the data availability section we present a guide to create the training set (VASP AIMD inputs), as well as the MTP training procedure and the integration code to calculate the MTP forces of the PHONOPY input structures. In particular, all the examples of this work are included. We hope that such information will facilitate the use of MTP for accurate evaluation of the phononic properties in structural systems of versatile complexity.



**Fig. 1.** Top views of atomic structures of the considered monoelemental systems. Side views are shown only for non-planar lattices. Black lines illustrate the primitive unit cell.

### 3. Results and discussions

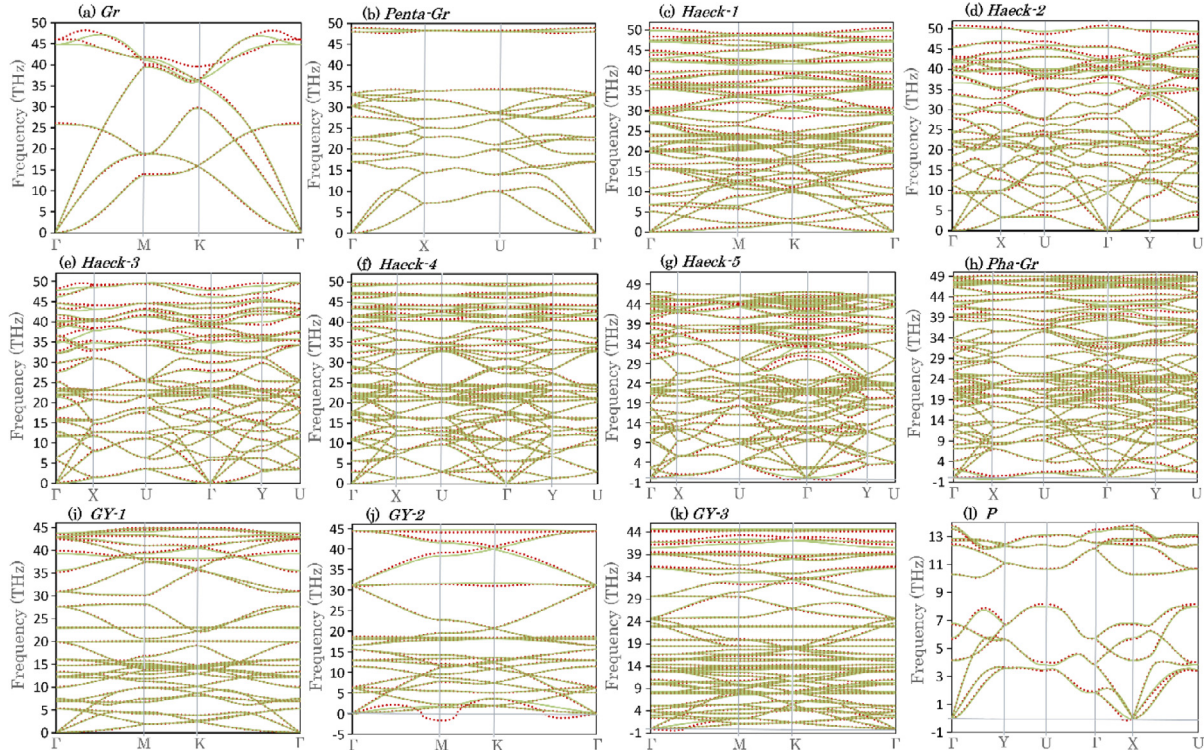
The main objective of this study is to replace DFPT with MTP in the force-constant calculations for the evaluation of phononic properties. Although our work is focused on the 2D materials, the proposed approach is equally valid for 3D or 1D lattices. First, we study the monoelemental 2D structures. Since carbon shows the unique ability to form diverse stable 2D atomic lattices, we mostly consider carbon-based structures. In Fig. 1, the studied monoelemental 2D lattices are illustrated. Apart from the graphene (Gr), we consider penta-graphene [37], various haeckelite [38] (Haeck) lattices, phagraphene [39] (Pha-Gr) and different graphyne [40] (GY) structures. We note that the haeckelite and phagraphene lattices include non-hexagonal carbon rings and can thus be considered as defective graphene lattices but with high degree of periodicities and close densities to that of the pristine graphene. On the other side, graphyne lattices are highly porous full carbon materials, which recently gained remarkable attentions because of experimental advances and their bright application prospect for the energy storage/conversion systems. Besides full-carbon structures, we also consider the single-layer black phosphorene (P) which has a rather complicated buckled lattice.

When using the DFPT method, the phononic properties are acquired on the basis of applying small displacements in supercell lattices. Therefore, the preliminary training data sets were prepared by conducting the AIMD simulations at a temperature of 50 K, with an overall simulation time of less than 1 ps (1000 simulation time steps). In this case, the AIMD trajectories include configurations with small lattice deformations from the equilibrium condition, consistent with our employed DFPT procedure. Since the MTPs are trained with a cutoff distance of 4 or 5 Å, the AIMD simulations were conducted over supercells with periodic box sizes over 10 Å. Larger supercells can be useful to describe the local large deflections better. Nonetheless, supercells with the minimum sizes closer to 10 Å are more computationally efficient because the costs of AIMD simulations increase exponentially with the number of atoms. To facilitate future studies, all created training sets in this work are provided in the data availability section. Using the AIMD results, we trained MTPs with 901 parameters for the monoelemental systems. The cutoff distance for graphene, other carbon allotropes and phosphorene were assumed to be 3, 4 and 5 Å, respectively. In Fig. 2 we compare the PDRs predicted by MTPs with those by the DFPT. Remarkably close agreement between the MTP and DFT results are obtained. This is a highly promising outcome since the developed MTPs are first attempts without any long sequence of optimization. Noticeably for the penta-graphene, haeckelite lattices and phosphorene, acoustic and optical branches are reproduced with a high level of accuracy. The same observation is also valid for graphyne lattices. Notably, in some cases (Haeck-5,

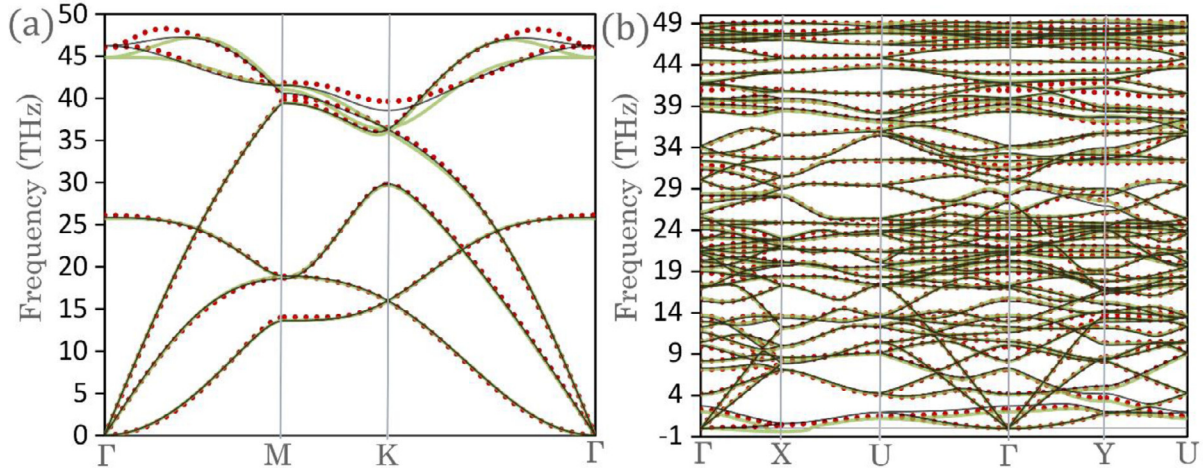
P, GY-2 and GY-3) slight imaginary frequencies appear in the PDRs calculated with DFPT, which are not observed in the MTP-based results. These imaginary branches in the DFPT results originate from computational artifacts, and are not representative of dynamical instabilities [41–43] and might be removed via more precise set-ups for the calculations, such as increasing the supercell size, increasing plane-wave cutoff energy or improving the resolution of k-point grids. These modifications can be affordably examined for highly symmetrical lattices like graphene, but for less symmetrical and nanoporous lattices may substantially increase the computational costs and lead to complexity of the problem.

Despite close agreement between the predicted PDRs for the most of monoelemental structures, two exceptions exist in which the MTP-based results show some inaccuracies. Among all the studied samples, it is clear that the high-frequency optical modes are well reproduced by the MTP. Unexpectedly for the highest symmetrical lattice of graphene, optical modes with the highest frequencies are not well reproduced by the MTP (Fig. 2a). On the other side for the phagraphene (Fig. 2h), a slight soft mode appears in the MTP result which is absent in the DFPT result. These disagreements between the DFPT and MTP results reveal that some degree of extrapolation may have occurred and accordingly suggest that for these samples the created passive training sets may not have been as accurate as those for other samples. We remind that when conducting the AIMD simulations at the low temperature of 50 K for a very short simulation time, the occurring configurations remain highly correlated and some critical configurations may never get explored in the training set. To address this issue, a simple effective solution is to expand the training set by including additional uncorrelated AIMD trajectories at higher temperatures. To that end, additional 1 ps AIMD trajectories were included by conducting the simulations at 200, 300, 500, 700 and 900 K (each one for 0.2 ps). This way the total AIMD trajectories include 2 ps, and the new training sets were created accordingly. In Fig. 3, we include the predicted PDRs for single-layer graphene and phagraphene by the improved training sets and compare the new results with the original results. It is clear that by incorporating higher temperature trajectories, the imaginary frequencies in the phagraphene's PDR vanishes, while the close agreement with DFPT results at higher frequencies is kept intact. Similarly, for the case of graphene, the accuracy of high-frequency optical modes is considerably improved without affecting the lower frequency optical and acoustic modes. The acquired results reveal that the incorporation of high-temperature trajectories can further enhance the accuracy and stability of MTPs. It is worth noting that in prior works, MLIPs have been successfully employed to accurately reproduce the phononic properties of graphene [44,45]. In comparison with these earlier works, the preparation of training sets in our approach is however more computationally efficient. In addi-





**Fig. 2.** Phonon dispersion relations of monoelemental 2D lattices acquired by the DFT (red-dotted lines) method and first-attempt MTPs (continuous green lines). The atomic lattices are shown in Fig. 1.



**Fig. 3.** Phonon dispersion relations of (a) graphene and (b) phagraphene acquired by the DFPT (red-dotted lines) and MTPs passively trained over AIMD trajectories at 50 K (continuous green lines) and 50 to 900 K (continuous green lines).

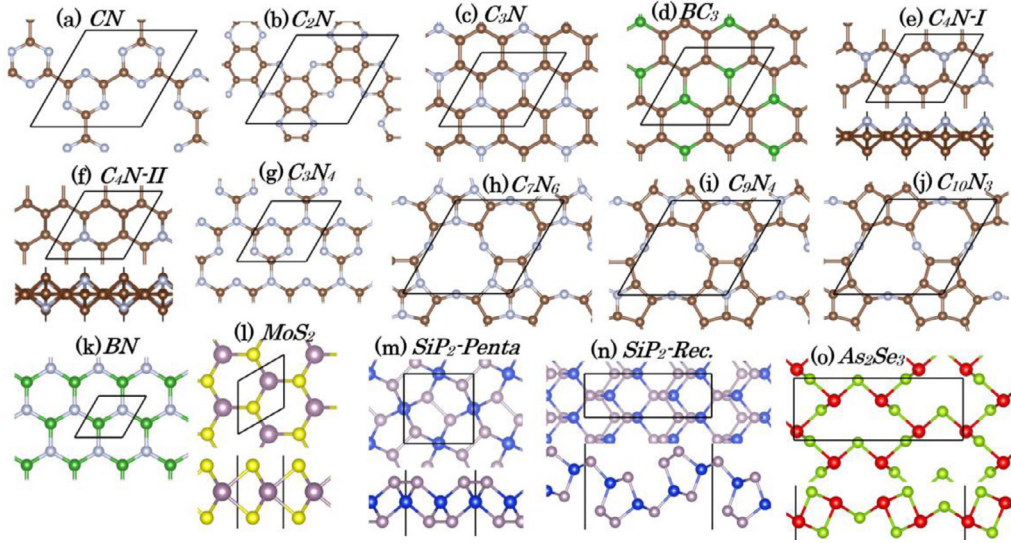
tion, when comparing with prior studies [44,45], our method could more accurately reproduce the acoustic modes in graphene's PDR, which is an important finding because these modes are the dominant heat carriers in graphene.

We next consider the binary 2D systems shown in Fig. 4. We note that various carbon-nitride 2D systems, like CN[41],  $C_2N$ [46],  $C_3N_4$ [47] and  $C_3N$ [48] are among the most attractive 2D semiconductors that have been experimentally fabricated and exhibit promising performances for a wide range of applications.  $BC_3$ [49] has been also experimentally fabricated,  $C_4N$  and  $C_7N_6$ ,  $C_9N_4$  and  $C_{10}N_3$  have been theoretically predicted by Li et al. [50] and Mortazavi et al. [51], respectively. We also consider carbon-free binary structures of hexagonal boron-nitride (BN), 2H transition metal dichalcogenides ( $MX_2$ ,  $M = Mo, W$  and  $X = S, Se, Te$ ),  $SiP_2$  [52] and

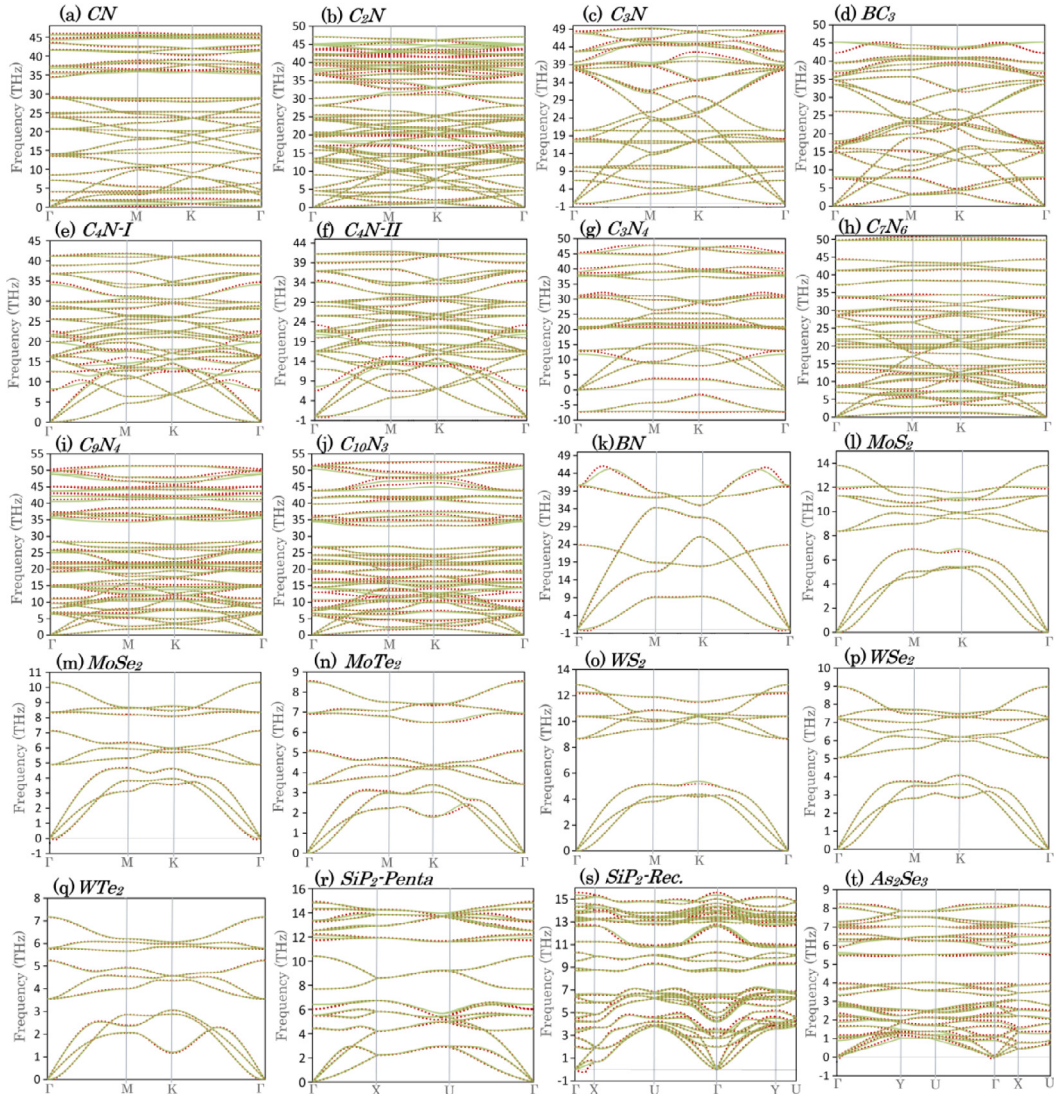
$As_2Se_3$ [53] monolayers. Similar to the case of monoelemental 2D lattices, we took MTPs with 1009 parameters and trained them on 1ps-long AIMD trajectories at the temperature of 50 K.

The predicted PDRs by MTPs are compared with those by the DFPT method for the binary 2D lattices in Fig. 5. As it is clear, passively-trained MTPs can very accurately reproduce the PDRs for the considered binary lattices. It is noticeable that PDRs on the basis of MTPs are free of the slightly imaginary frequencies that have occurred around the  $\Gamma$  points in the DFPT results for some samples. The  $C_3N_4$  monolayer with a full flat structure shows conspicuous imaginary frequencies, accurately reproduced also by the MTP-based method. These results are highly promising taking into account the complexity of the considered structures, especially in view of rectangular  $SiP_2$  and  $As_2Se_3$  exhibiting intricate buckled

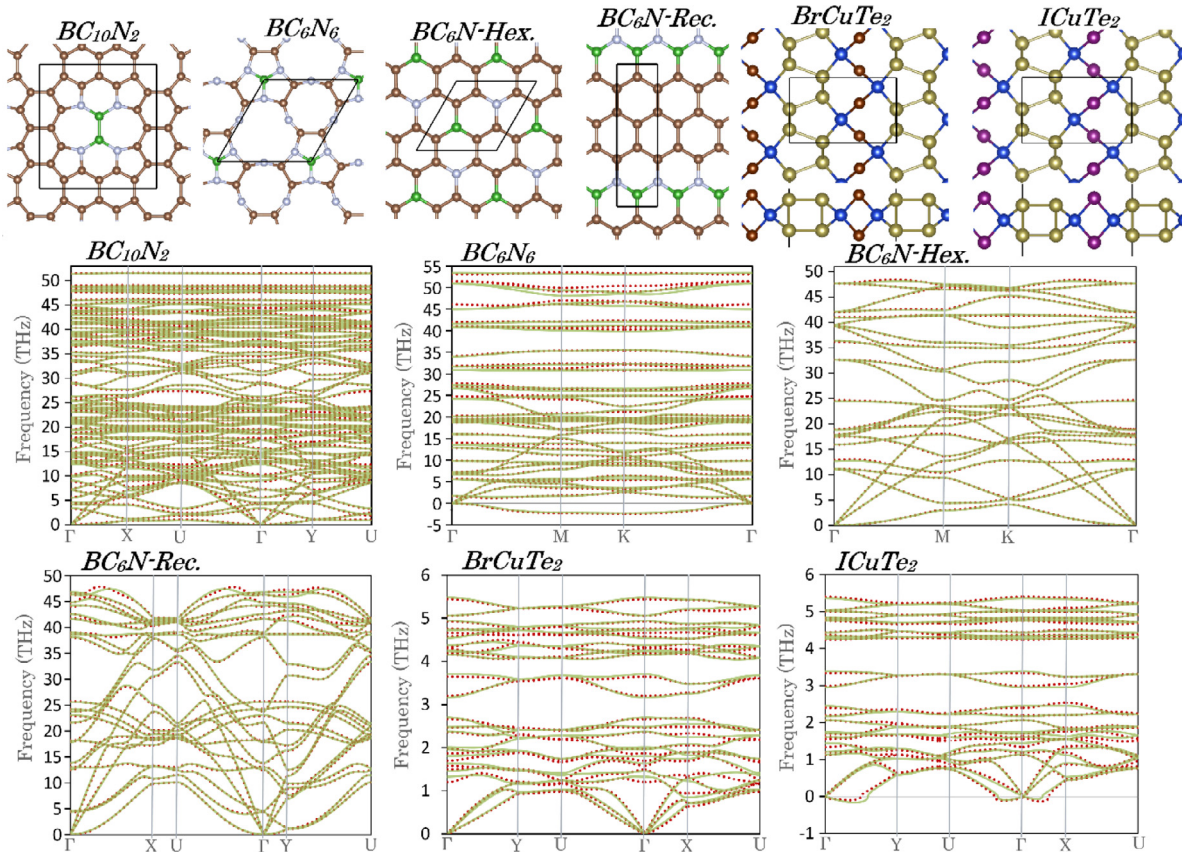




**Fig. 4.** Top views of atomic structures of binary 2D systems. Side views are shown only for non-planar lattices. Black lines illustrate the primitive unit cell.



**Fig. 5.** Phonon dispersion relations of binary 2D lattices acquired by the DFPT (red-dotted lines) and first-attempt MTPs (continuous green lines). The corresponding atomic lattices are shown in Fig. 4.



**Fig. 6.** Phonon dispersion relations for six different ternary 2D lattices acquired by the DFPT (red-dotted lines) and first-attempt MTPs (continuous green lines). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

lattices. It should also be noted that while the acoustic modes are very closely reproduced for all considered samples, slight deviations are observed for some of the optical modes in a few samples, particularly for the  $C_2N$  and  $BC_3$  monolayers. As discussed for the case of graphene, to better reproduce the optical modes, the training set should be improved by including the AIMD trajectories at higher temperatures.

We next explore the accuracy of the PDRs predictions by the MTP method for ternary 2D lattices. In Fig. 6 we compare the PDRs for six different ternary lattices by the DFPT and MTP methods. We note that  $BC_{10}N_2$  is recently predicted by Tromer et al.[54],  $BC_6N$  shows hexagonal and rectangular atomic lattices,  $BrCuTe_2$  and  $ICuTe_2$  are also predicted according to their bulk counterparts, and  $BC_6N_6$  is a novel ternary lattice whose stability we examine. It is noticeable that imaginary frequencies in the DFPT results are also well reproduced by the passively trained MTPs. These results, consistent with our previous observations, highlight that the accuracy of MTPs are insensitive to the number of elements' types and/or complexity of the structure. Although we have limited our AIMD simulations to less than 1 ps at 50 K, results presented up to this stage reveal that for the most of cases no additional AIMD trajectories are required to assess the phonon dispersions. Nonetheless, if required, the incorporation of additional high temperature AIMD trajectories in the training set is expected to improve the accuracy and stability as well.

Besides the phonon dispersions, the phonons group velocity is another important phononic property, which can provide useful information concerning the lattice thermal conductivity. In Fig. 7 we compare the phonon group velocities predicted by the MTP method with those obtained within the DFPT using the PHONOPY code[34]. Remarkably, the passively trained MTPs can very closely

reproduce the phonon group velocities for the studied samples. For some of the samples like graphene, penta-graphene and  $-SiP_2$ ,  $CN$ ,  $BC_6N_6$  and transition metal dichalcogenides ( $MX_2$ ,  $M = Mo, W$  and  $X = S, Se, Te$ ) excellent agreement between the MTP method and DFPT-based estimations are observed. It is clear that in the worst cases, like  $BrCuTe_2$ ,  $BC_3$ , or  $C_2N$  the group velocities are slightly shifted up or down and the results are free of substantial inaccuracies.

Last but not least, we examine the agreement between the MTP- and DFPT-based approaches to calculate the free energy, heat capacity, and entropy from their statistical thermodynamic expressions using the PHONOPY code[34]. We compared the MTP- and DFPT-based estimations for the aforementioned thermal properties in Fig. 8. As it is clear, for these results the agreement between DFPT and MTP results is excellent.

Presented results highlight the accuracy and stability of the MTP-based method to explore the phononic properties of complex 2D lattices. The required AIMD simulations to assess the phononic properties can be conducted relatively fast and are computationally less expensive than those commonly performed to examine the thermal stability (which usually requires more than 10 ps-long AIMD simulations). In comparison with DFPT method, AIMD simulations show lower sensitivity to the K-point grid and plane-wave cutoff energy. Moreover, the computational cost for the examination of supercell size effect on the PDRs is marginal for MTP-based approach, which is in contrast the bottleneck of DFPT method. One of the most salient advantages of the MTP-based method is that it typically yields smooth dispersions for the acoustic modes near the  $\Gamma$ -point. Differently when using the DFPT method for 2D materials, the convergence of flexural acoustic mode near the  $\Gamma$ -point may turn into a difficult task [55]. Nonphysical imaginary



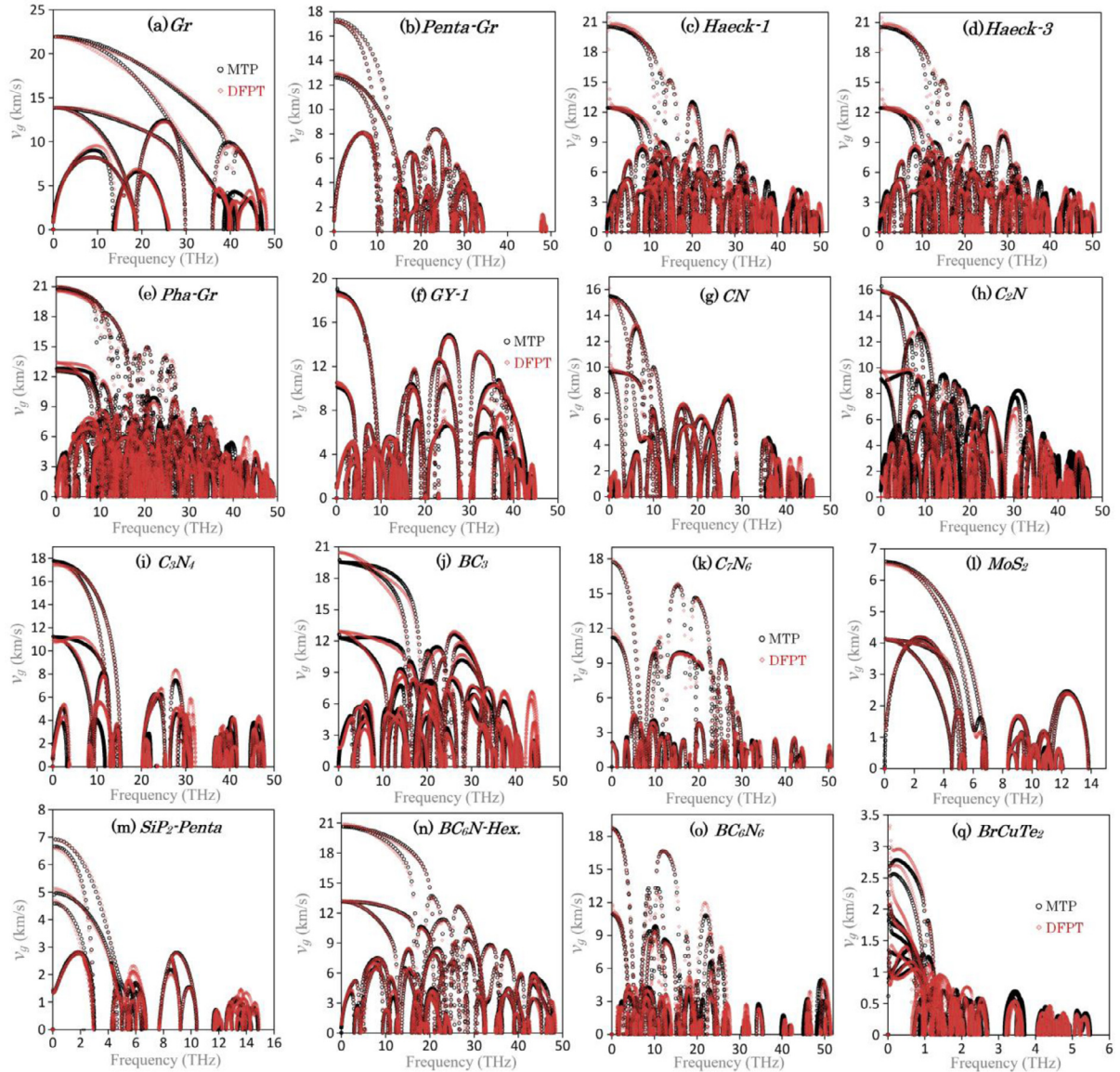
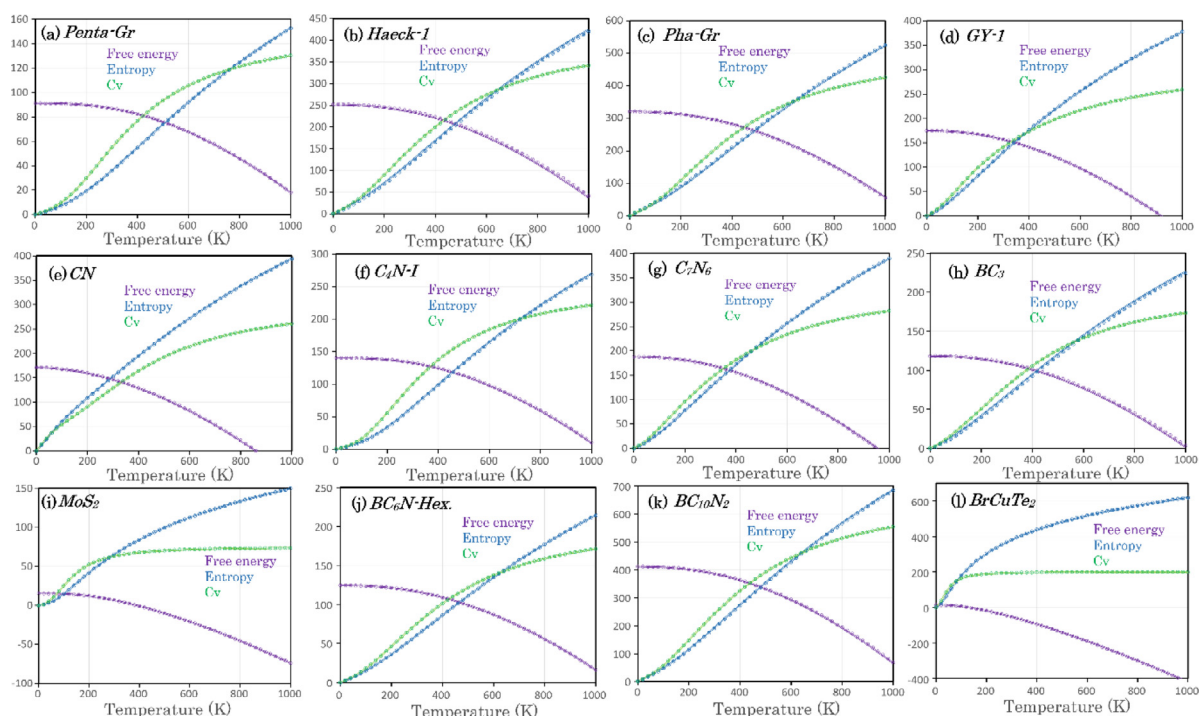


Fig. 7. Phonon group velocity ( $v_g$ ) predicted by MTP and DFPT methods.

frequencies in the acquired PDRs may result in the instabilities for evaluating other thermal properties, particularly when calculating the lattice thermal conductivity. Worthy to point out that one of the reasons for appearance of imaginary frequencies in the PDRs might stem from slight inaccuracies in the atomic lattice. When using the DFPT method, small structural changes in the primitive unit cell, either in the atomic position or strain in the lattice size, demand entirely new calculations, whereas these effects can be checked with negligible computation costs using the MTP-based method. We should also observe that for highly symmetrical structures like graphene, transition metal dichalcogenides and penta-graphene, the standard DFPT method is by a large extent more computationally efficient. Actually the computational efficiency of MTP significantly enhances as the symmetry decreases, like that for haeckelites and  $C_7N_6$  structures or for materials with large primitive unit cells, such as: metal- or conductive-organic frameworks and graphyne/graphdiyne lattices. In the previous works [44,45] it was demonstrated that MLIPs can accurately reproduce the phononic properties of graphene as compared with the standard DFPT method, but with substantially higher computational costs, questioning the practical application of MLIPs for

this purpose. However, our results for extensive 2D systems clearly confirm that for low-symmetry and highly porous atomic lattices, the MTP-based approach offers an accurate, simple and convenient alternative to commonly employed DFPT approach to access the phononic properties. This is therefore a novel application area of MLIPs in the materials science as they can accurately and efficiently substitute the standard DFPT method for the valuation of phononic properties of complex structures. Although, the majority of our results on the basis of short AIMD trajectories at the low temperature of 50 K reveal outstanding accuracy, it is nonetheless highly recommended to include AIMD trajectories at higher temperatures in the training set to ensure improved stability and accuracy as well. Moreover, to more accurately reproduce the high frequency optical modes, employment of larger supercells in the AIMD calculations is highly recommended. In this case, the increase in the computational cost can be partially compensated by shortening the AIMD simulation. In this work PBE method was employed in the AIMD calculations for the evaluation of phononic properties. However we note that the proposed MTP-based approach can be also effectively used for other exchange-correlation functionals. Worthy to note that since AIMD simulations are avail-



**Fig. 8.** Free energy, heat capacity, and entropy with the units of kJ/mol, J/K/mol, and J/K/mol, respectively, on the basis of MTP (continuous lines) and DFPT (circles) methods acquired using the PHONOPY code [34].

able by a vast number of free and open-source first-principles packages, the methodology proposed in this study can be efficiently employed to conveniently examine the phononic properties of complex structures.

#### 4. Conclusion

Our extensive results for a wide variety of 2D materials highlight that machine-learning interatomic potentials trained over short ab-initio molecular dynamics trajectories are able to reproduce the phononic properties in a strikingly close agreement with those by DFPT-based results. The proposed methodology could therefore play a pivotal role to conveniently explore the phononic properties of a large variety of low-symmetry and porous nanomembranes, with a high level of accuracy and reproducibility. In recent years astonishing advances have been achieved on the synthesis of low-symmetry and highly porous atomic lattices, like metal- or conductive-organic frameworks and graphdiyne nanomembranes. It is clear that phononic properties of these novel systems can now be effectively explored using the approach proposed in this work, which would otherwise require excessive computational resources with DFT-based methods. Our study therefore proposes that MTP-based approach can accurately and efficiently substitute the standard DFPT method for the valuation of phononic properties of complex structures. To facilitate the practical application, the full computational details of the proposed approach are included in the data availability section.

#### Data availability

The following data are available to download: (1) a brief guide for the installation of MLIP package, (2) all energy minimized lattices, (3) passive training approach and related commands, (4) VASP input script for the AIMD simulations, (5) VASP output file (vasprun.xml) for the DFT calculations, (6) created training sets and trained MTPs (p.mtp), (7) the C++ code to calculate

the force constants using the PHONOPY and MTP for the input geometries and force calculator, respectively and (8) PHONOPY related scripts to adjust the outputs, from: GITLAB: [https://gitlab.com/ivannovikov/mlip\\_phonopy](https://gitlab.com/ivannovikov/mlip_phonopy) and Mendeley: <http://dx.doi.org/10.17632/7ppcf7cs27.1>.

#### Funding

All of the sources of funding for the work described in this publication are acknowledged below: B.M. and X.Z. appreciate the funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy within the Cluster of Excellence PhoenixD (EXC 2122, Project ID 390833453). E.V.P. I.S.N., and A.V.S. were supported by the Russian Science Foundation (Grant No 18-13-00479). ICN2 is supported by the Severo Ochoa program from Spanish MINECO (Grant No. SEV-2017-0706) and funded by the CERCA Programme/Generalitat de Catalunya

#### Intellectual property

We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property

#### Research ethics

We further confirm that any aspect of the work covered in this manuscript that has involved human patients has been conducted with the ethical approval of all relevant bodies and that such approvals are acknowledged within the manuscript



## Authorship

The International Committee of Medical Journal Editors (ICMJE) recommends that authorship be based on the following four criteria:

- 1 Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND
- 2 Drafting the work or revising it critically for important intellectual content; AND
- 3 Final approval of the version to be published; AND
- 4 Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

All those designated as authors should meet all four criteria for authorship, and all who meet the four criteria should be identified as authors. For more information on authorship, please see <http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html#two>

## Declaration of Competing Interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome

## CRediT authorship contribution statement

**Bohayra Mortazavi:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Writing - review & editing. **Ivan S. Novikov:** Methodology, Software, Formal analysis, Investigation, Writing - review & editing. **Evgeny V. Podryabinkin:** Methodology, Software, Writing - review & editing. **Stephan Roche:** Supervision, Formal analysis, Writing - original draft. **Timon Rabczuk:** Supervision, Investigation, Methodology. **Alexander V. Shapeev:** Methodology, Software, Supervision, Funding acquisition, Writing - review & editing. **Xiaoying Zhuang:** Resources, Investigation, Methodology, Supervision, Funding acquisition.

## Acknowledgment

B.M. and X.Z. appreciate the funding by the [Deutsche Forschungsgemeinschaft](#) (DFG, German Research Foundation) under Germany's Excellence Strategy within the Cluster of Excellence PhoenixD (EXC 2122, Project ID 390833453). E.V.P., I.S.N., and A.V.S. were supported by the [Russian Science Foundation](#) (Grant No 18-13-00479). ICN2 is supported by the Severo Ochoa program from Spanish MINECO (Grant No. [SEV-2017-0706](#)) and funded by the CERCA Programme/Generalitat de Catalunya.

## References

- [1] M.T. Dove, Introduction to the theory of lattice dynamics, École thématique la Société Française la Neutron (2011), doi:10.1051/sfn/201112007.
- [2] M.T. Dove, Introduction to lattice dynamics, *Introd. Lattice Dyn.* (1993), doi:10.1119/1.17708.
- [3] K.S. Novoselov, et al., Electric field effect in atomically thin carbon films, *Science* 306 (2004) 666–669.
- [4] A.K. Geim, K.S. Novoselov, The rise of graphene, *Nat. Mater.* 6 (2007) 183–191.
- [5] T.T. Tran, K. Bray, M.J. Ford, M. Toth, I. Aharonovich, Quantum emission from hexagonal boron nitride monolayers, *Nat. Nanotechnol.* 11 (2016) 37–41.
- [6] M. Raeisi, S. Ahmadi, A. Rajabpour, Modulated thermal conductivity of 2D hexagonal boron arsenide: a strain engineering study, *Nanoscale* 11 (2019) 21799–21810.
- [7] Y. Wang, Y. Ding, Strain-induced self-doping in silicene and germanene from first-principles, *Solid State Commun.* 155 (2013) 6–11.
- [8] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, A. Kis, Single-layer MoS<sub>2</sub> transistors, *Nat. Nanotechnol.* 6 (2011) 147–150.
- [9] R.H. Baughman, H. Eckhardt, M. Kertesz, Structure-property predictions for new planar forms of carbon: layered phases containing sp<sup>2</sup> and sp atoms, *J. Chem. Phys.* 87 (1987) 6687.
- [10] D. Jariwala, V.K. Sangwan, L.J. Lauhon, T.J. Marks, M.C. Hersam, Emerging device applications for semiconducting two-dimensional transition metal dichalcogenides, *ACS Nano* 8 (2014) 1102–1120.
- [11] T. Wang, C. Zhang, H. Snoussi, G. Zhang, Machine learning approaches for thermoelectric materials research, *Adv. Funct. Mater.* 30 (2020) 1906041.
- [12] B. Sun, A.S. Barnard, Visualising multi-dimensional structure/property relationships with machine learning, *J. Phys. Mater.* 2 (2019) 34003.
- [13] H. Oda, S. Kiyohara, T. Mizoguchi, Machine learning for structure determination and investigating the structure-property relationships of interfaces, *J. Phys. Mater.* 2 (2019) 34005.
- [14] G.R. Schleder, A.C.M. Padilha, C.M. Acosta, M. Costa, A. Fazzio, From DFT to machine learning: recent approaches to materials science—a review, *J. Phys. Mater.* 2 (2019) 32001.
- [15] J. Schmidt, M.R.G. Marques, S. Botti, M.A.L. Marques, Recent advances and applications of machine learning in solid-state materials science, *npj Comput. Mater.* 5 (2019) 83.
- [16] G. Carleo, et al., Machine learning and the physical sciences, *Rev. Mod. Phys.* 91 (2019) 45002.
- [17] Han, B. et al. Deep-learning-enabled fast optical identification and characterization of two-dimensional materials. (2019).
- [18] E.V. Podryabinkin, A.V. Shapeev, Active learning of linearly parametrized interatomic potentials, *Comput. Mater. Sci.* 140 (2017) 171–180.
- [19] K. Gubaev, E.V. Podryabinkin, G.L.W. Hart, A.V. Shapeev, Accelerating high-throughput searches for new alloys with active learning of interatomic potentials, *Comput. Mater. Sci.* 156 (2019) 148–156.
- [20] E.V. Podryabinkin, E.V. Tikhonov, A.V. Shapeev, A.R. Oganov, Accelerating crystal structure prediction by machine-learning interatomic potentials with active learning, *Phys. Rev. B* 99 (2019) 064114.
- [21] V.V. Ladygin, P.Y. Korotaev, A.V. Yanilkin, A.V. Shapeev, Lattice dynamics simulation using machine learning interatomic potentials, *Comput. Mater. Sci.* 172 (2020) 109333.
- [22] B. Mortazavi, et al., Efficient machine-learning based interatomic potentials for exploring thermal conductivity in two-dimensional materials, *J. Phys. Mater.* 3 (2020) 02LT02, doi:10.1088/2515-7639/ab7cbb.
- [23] P. Korotaev, I. Novoselov, A. Yanilkin, A. Shapeev, Accessing thermal conductivity of complex compounds by machine learning interatomic potentials, *Phys. Rev. B* 100 (2019) 144308.
- [24] A.V. Shapeev, Moment tensor potentials: a class of systematically improvable interatomic potentials, *Multiscale Model. Simul.* 14 (2016) 1153–1173.
- [25] K. Schutt, et al., Schnet a continuous-filter convolutional neural network for modeling quantum interactions, *Advances in neural information processing systems* (2017) 991–1001.
- [26] V. Botu, R. Batra, J. Chapman, R. Ramprasad, Machine learning force fields: construction, validation, and outlook, *J. Phys. Chem. C* 121 (2017) 511–522.
- [27] Y. Zuo, et al., Performance and cost assessment of machine learning interatomic potentials, *J. Phys. Chem. A* 124 (2020) 731–745.
- [28] J. Behler, Perspective: Machine learning potentials for atomistic simulations, *J. Chem. Phys.* 145 (2016) 170901.
- [29] G. Kresse, J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, *Comput. Mater. Sci.* 6 (1996) 15–50.
- [30] G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B* 54 (1996) 11169–11186.
- [31] G. Kresse, From ultrasoft pseudopotentials to the projector augmented-wave method, *Phys. Rev. B* 59 (1999) 1758–1775.
- [32] J. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.* 77 (1996) 3865–3868.
- [33] H. Monkhorst, J. Pack, Special points for Brillouin zone integrations, *Phys. Rev. B* 13 (1976) 5188–5192.
- [34] A. Togo, I. Tanaka, First principles phonon calculations in materials science, *Scr. Mater.* 108 (2015) 1–5.
- [35] K. Gubaev, E.V. Podryabinkin, A.V. Shapeev, Machine learning of molecular properties: locality and active learning, *J. Chem. Phys.* 148 (2018) 241727.
- [36] I.S. Novikov, A.V. Shapeev, Improving accuracy of interatomic potentials: more physics or more data? A case study of silica, *Mater. Today Commun.* 18 (2019) 74–80.
- [37] S. Zhang, et al., Penta-graphene: a new carbon allotrope, in: *Proceedings of the National Academy of Sciences of the United States of America*, 2015, doi:10.1073/pnas.1416591112.
- [38] H. Terrones, et al., New metallic allotropes of planar and tubular carbon, *Phys. Rev. Lett.* 84 (2000) 1716–1719.
- [39] Z. Wang, et al., Phagraphene: a low-energy graphene allotrope composed of 5–6–7 carbon rings with distorted dirac cones, *Nano Lett.* 15 (2015) 6182–6186.
- [40] R.H. Baughman, H. Eckhardt, M. Kertesz, R.H. Baughman, H. Eckhardt, Structure-property predictions for new planar forms of carbon: layered phases containing sp<sup>2</sup> and sp atoms Structure-property predictions for new planar forms of carbon : layered phases containing sp<sup>2</sup> and sp atoms, *J. Chem. Phys.* 11 (1987) 6687–6699.
- [41] V.N. Popov, P. Lambin, Theoretical Raman fingerprints of  $\sqrt{3}\times\sqrt{3}$ -graphyne, *Phys. Rev. B* 88 (2013) 75427.

- [42] H. Zhan, Y. Nie, Y. Chen, J.M. Bell, Y. Gu, Thermal transport in 3D nanostructures, *Adv. Funct. Mater.* 30 (2020) 1903841.
- [43] H. Zhan, et al., Thermal conductivity of a new carbon nanotube analog: the diamond nanothread, *Carbon N. Y.* 98 (2016) 232–237.
- [44] P. Rowe, G. Csányi, D. Alfè, A. Michaelides, Development of a machine learning potential for graphene, *Phys. Rev. B* 97 (2018) 54303.
- [45] M. Wen, E.B. Tadmor, Hybrid neural network potential for multilayer graphene, *Phys. Rev. B* 100 (2019) 195419.
- [46] J. Mahmood, et al., Nitrogenated holey two-dimensional structures, *Nat. Commun.* 6 (2015) 1–7.
- [47] G. Algara-Siller, et al., Triazine-based graphitic carbon nitride: a two-dimensional semiconductor, *Angew. Chemie - Int. Ed.* 53 (2014) 7450–7455.
- [48] J. Mahmood, et al., Two-dimensional polyaniline ( $C_3N$ ) from carbonized organic single crystals in solid state, *Proc. Natl. Acad. Sci.* 113 (2016) 7414–7419.
- [49] H. Tanaka, et al., Novel macroscopic BC<sub>3</sub>honeycomb sheet, *Solid State Commun* (2005), doi:10.1016/j.ssc.2005.06.025.
- [50] L. Li, et al., Carbon-rich carbon nitride monolayers with dirac cones: dumbbell  $C_4N$ , *Carbon N. Y.* 118 (2017) 285–290.
- [51] B. Mortazavi, M. Shahrokhi, A.V. Shapeev, T. Rabczuk, X. Zhuang, Prediction of  $C_7N_6$  and  $C_9N_4$ : stable and strong porous carbon-nitride nanosheets with attractive electronic and optical properties, *J. Mater. Chem. C* 7 (2019) 10908–10917.
- [52] F. Shojaei, B. Mortazavi, X. Zhuang, M. Azizi, Silicon diphosphide ( $SiP_2$ ) and silicon diarsenide ( $SiAs_2$ ): novel stable 2D semiconductors with high carrier mobilities, promising for water splitting photocatalysts, *Mater. Today Energy* 16 (2020) 100377.
- [53] B. Mortazavi, F. Shojaei, M. Azizi, T. Rabczuk, X. Zhuang,  $As_2S_3$ ,  $As_2Se_3$  and  $As_2Te_3$  nanosheets: superstretchable semiconductors with anisotropic carrier mobilities and optical properties, *J. Mater. Chem. C* 8 (2020) 2400–2410.
- [54] R.M. Tromer, I.M. Felix, A. Freitas, S. Azevedo, L.F.C. Pereira, Diboron-porphyrin monolayer: a new 2D semiconductor, *Comput. Mater. Sci.* 172 (2019) 109338.
- [55] Z. Zhu, J. Guan, D. Liu, D. Tománek, Designing isoelectronic counterparts to layered group V semiconductors, *ACS Nano* (2015), doi:10.1021/acs.nano.5b02742.