Machine-learning-boosted ab-initio study of the thermal conductivity of Janus PtSTe van der Waals heterostructures

Lijun Pan,[1, 2] Jesús Carrete,[3, 2, *] Zhao Wang,[1] and Georg K. H. Madsen[2]

[1] *Department of Physics, Guangxi University, Nanning 530004, China*
[2] *Institute of Materials Chemistry, TU Wien, 1060 Vienna, Austria*
[3] *Instituto de Nanociencia y Materiales de Aragón (INMA),*
CSIC-Universidad de Zaragoza, E-50009 Zaragoza, Spain

Calculating the thermal conductivity of heterostructures with multiple layers presents a significant challenge for state-of-the-art ab-initio methods. In this study we introduce an efficient neural-network force field (NNFF) to explore the thermal transport characteristics of van der Waals heterostructures based on PtSTe, using both the phonon Boltzmann transport equation and molecular dynamics (MD) simulations. Besides demonstrating a remarkable level of agreement with both theoretical and experimental data, our predictions reveal that heterogeneous combinations like PtSTe-PtTe$_2$ display a notable reduction in thermal conductivity at room temperature, primarily due to broken out-of-plane symmetries and the presence of weak van der Waals interactions. Furthermore, our study highlights the superiority of MD simulations with NNFFs in capturing higher-order anharmonic phonon properties. This is demonstrated through the analysis of the temperature-dependent thermal conductivity curves of PtSTe-based van der Waals heterostructures, and advances our understanding of phonon transport in those materials.

I. INTRODUCTION

Finding materials with a low thermal conductivity (κ) is a necessary step to create thermoelectric (TE) devices able to efficiently scavenge electrical power from waste heat. In this context, monolayer transition metal dichalcogenides (TMDC) have emerged as promising candidates [1] since they also exhibit large charge-carrier effective masses. Both factors contribute to raising the dimensionless thermoelectric figure of merit (zT), the key descriptor of thermoelectric performance. A particularly interesting variation on those basic monolayer structures are Janus TMDCs with the formula MXY (where M is a transition metal and X and Y represent two different chalcogens) and a broken reflection symmetry in the out-of-plane direction. A number of experimental and theoretical studies have looked into the prospects of 2D Janus TMDC monolayers as TE materials and their connections with symmetry [2–9]. For instance, the Janus monolayer MoSSe has been experimentally synthesized [8] and found to have a thermal conductivity significantly lower than that of MoS$_2$ and only slightly higher than that of MoSe$_2$ despite the much heavier average atomic mass of the chalcogens in the latter [6]. Furthermore, monolayer PtSTe shows a lower thermal conductivity than both of its parent structures (PtTe$_2$ and PtS$_2$) [9].

A large set of additional degrees of freedom for tuning TE performance comes from the possibility of stacking those monolayers, as the transport properties of weakly bound layered structures are strongly dependent on their thickness and stacking sequence. For instance, the κ of bilayer graphene is much lower than that of monolayer graphene [10], a change driven by new phonon scattering channels caused by the weak interlayer coupling. More generally, van der Waals Heterostructures (vdWHSs) of 2D layered materials can display emerging functionalities distinct from those of the individual components. The zT of WS$_2$/WTe$_2$ bilayers has been theoretically predicted to reach the high value of 2.4 at room temperature due to its ultralow thermal conductivity [11]. This creates a strong motivation to design and test novel heterostructures, and 2D monolayer Janus TMDCs stand out as promising building blocks for those.

In this work, we focus on the thermal conductivity of different bi- and trilayer vdWHSs composed of PtS$_2$, PtTe$_2$, and Janus PtSTe monolayers at room temperature. For crystalline semiconducting systems, where lattice thermal transport is dominant, solving the Boltzmann transport equation (BTE) for phonons from first principles has emerged as one of the most popular and robust approaches for obtaining predictive estimates of the thermal conductivity [12, 13]. However, the complexity and diversity of the vdWHSs pose a computational challenge when attempting to calculate all the necessary components for solving the BTE. To overcome this obstacle while maintaining accuracy, we employ a neural-network force field (NNFF) [14, 15] trained on a smaller first-principles dataset composed primarily of less complex structures and strategically enriched to improve transferability. Even with the significant acceleration afforded by the machine-learning calculation of energies and forces, ordinary BTE-based methods have limited scalability with the number of atoms in the unit cell. This has been alleviated recently by several generalizations of the formalism [16–18] to encompass diffusive thermal transport mediated by the coupling of non-propagating vibrational modes. These generalizations have created quantitative frameworks and calculation workflows that include diffusons (introduced by Allen and Feldmann [19]) as heat carriers in addition to phonons, and are applicable to com-

---

* jcarrete@unizar.es

plex unit cells as well as disordered systems. Here, based on the recent work by Baroni and coworkers [20–22], we explore the alternative of using equilibrium molecular dynamics (MD) simulations with the same force field, which also lift all restrictions regarding the order and importance of anharmonicity. To validate our model, we compare the results with data obtained from first-principles calculations for the monolayer. To get further insight into the differences between the BTE and MD results, we also perform a temperature-dependent study of the thermal conductivity of monolayer and bilayer Janus PtSTe-based vdWHSs. We then extend our calculations to trilayer nanostructures.

II. METHODS

The NNFF model sits at the center of our methodology, powering both the BTE-based workflow and the MD simulations. Ab-initio calculations are used to train the NNFF (in a series of refinement steps) and to perform the BTE calculations employed as tests/calibration. The following subsections describe each step of the process in detail.

A. Force-field architecture and training

We parameterize the potential-energy hypersurface of our systems using a committee of 5 NeuralIL FFs [14, 15], which also afford an estimate of the uncertainty in that prediction. Each member of the committee is a neural network patterned after the Behler-Parinello architecture [23] consisting of a first stage where the Cartesian coordinates of each atom are translated into a set of symmetry-adapted descriptors, a second one where a per-atom contribution to the energy (ϕ_j) is computed from those descriptors, and a sum over atoms at the end of the pipeline. NeuralIL uses the spherical Bessel descriptors [24] augmented with an element-based embedding, the Swish-1 activation [25] and a deep residual network (ResNet) core to improve upon the original implementation of this architecture [15]. Moreover, it is end-to-end algorithmically differentiable thanks to being implemented on top of JAX [26], so forces are automatically obtained from the gradient of the energy. The predictions of the ensemble are obtained as the average of the predictions of the individual members, and the standard deviation of that sample mean is taken as an estimate of the uncertainty.

Every member of the ensemble shares the same characteristics: a cutoff of 4.5 Å and a maximum radial order of 6 for the spherical Bessel descriptors, and ResNet core layer widths of 64, 32, 16 and 16 (note that each ResNet layer actually involves several layers of neurons plus normalization stages; see Ref. 15 for details). The loss function is formulated as

$$
\begin{aligned}
\mathcal{L} = &\frac{1}{2} \left\langle \frac{0.2}{n_{\mathrm{atoms}}} \times \right. \\
&\left. \sum_{i=1}^{n_{\mathrm{atoms}}} \log \left[\cosh \left(\frac{\|f_{i,\mathrm{predicted}} - f_{i,\mathrm{reference}}\|_2}{0.2 \ \mathrm{eV\mathring{A}}^{-1}} \right) \right] \right\rangle \\
&+ \frac{1}{2} \left\langle 0.02 \log \left[\cosh \left(\frac{E_{\mathrm{pot}} - E_{\mathrm{pot,reference}}}{n_{\mathrm{atoms}} \times 0.02 \ \mathrm{eV\,atom}^{-1}} \right) \right] \right\rangle .
\end{aligned}
\tag{1}
$$

Although we have found forces to supply much more information in the training process, the second term takes care of aligning the origins of energies of the model and the ground-truth calculations. The log-cosh function performs a smooth transition between quadratic and linear behavior and prevents outliers from dominating the training. We train the ensemble for 600 epochs using the fully nonlinear VeLO optimizer [27].

B. Generation and structure of the ab-initio dataset

The workflow (Fig. 1 (a)) begins with the building of a dataset. That initial dataset, obtained from DFT calculations, is utilized for committee training to create the primary model, which is then employed in adversarial attacks to enrich the database. The initial dataset consisted of a set of monolayer configurations of Janus PtSTe and its parent structures, PtS_2 and $PtTe_2$, all of them in the 1T structure. Janus PtSTe, depicted in Fig. 1(b), belongs to the P3m1 space group, while the parent structures belong to P3m̄1 due to the presence of inversion symmetry. Starting with the equilibrium configurations of each structure, the dataset was populated by generating uniformly distributed random biaxial strains from -8% to $+8\%$ and applying normally distributed atomic displacements with a standard deviation of 0.05 Å. 700 configurations were created in this manner for the Janus monolayer, which were complemented with configurations for the parent monolayers until 1500 configurations were available in total. The interatomic forces and total energies of those configurations were then computed by DFT (see details below). This dataset was then randomly split into training (80% of the data) and validation (20% of the data) subsets. Following standard practice, the former was used to calculate the loss and its gradients, while the latter was reserved to evaluate the statistics describing the performance of the model at each step of the training process.

Building upon this foundation, we then augmented the dataset in three consecutive steps to extend the range of applicability of our model. In the first of those steps, we added data from 700 configurations of bilayer vdWHSs. Their generation follows the same template as in the case of the monolayers (random strain and random small displacements) but with an additional degree of freedom in
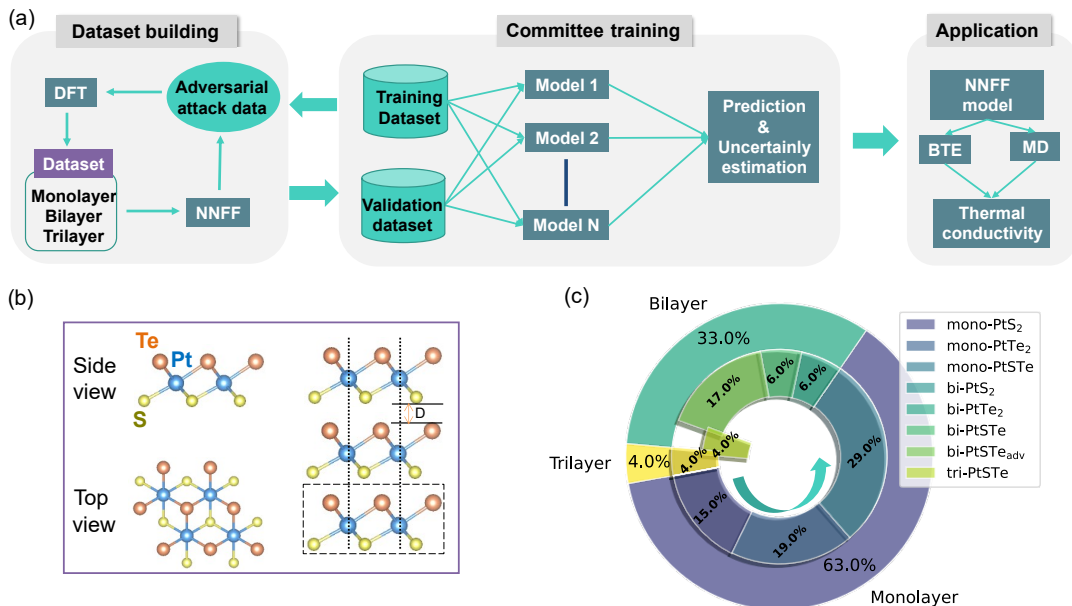
FIG. 1. (a) Schematic representation of the development of a committee of neural-network force fields for application in this study. (b) Monolayer and multilayer Janus PtSTe structures. (c) Distribution of structures in the training dataset.

the form of the interlayer distance, which was sampled at random within an interval with a half width of $0.5\,\text{Å}$ centered around the equilibrium values. The purpose of this procedure is to make the model aware of the characteristics of the weak interlayer interactions, which are naturally absent from the monolayers.

The second step was more targeted and aimed at finding and patching possible blind spots of the model with minimal additional data. The technique, detailed in Ref. 15, utilized adversarial attacks to locate configurations poorly characterized by the NN committee but statistically likely and involves optimizing the differentiable, likelihood-weighted uncertainty metric,

$$\mathcal{L}_{\text{adv}} = \sigma_{\boldsymbol{f}}^2 \exp\left(-\frac{E_{\text{pot}}}{k_{\text{B}}T}\right), \qquad (2)$$

through gradient-based optimization. T is set to $300\,\text{K}$ in this work. Thereby local optima maximizing NN committee uncertainty while considering thermodynamic likelihood are identified. Through this method, we added 100 new Janus bilayer configurations. Interatomic forces and total energies are calculated for these new configurations and integrated into the NN training, enhancing extrapolation abilities and enabling comprehensive exploration of the configuration space. The final step of our active-learning strategy consisted in adding 100 trilayer configurations with random strain and displacements.

Fig. 1(c) shows a breakdown of the composition of the final data set, including details about the origin of each subset of configurations. The successive improvements in the quality of the force field as data was added are clearly visible in Fig. 2, a parity plot of the forces over the

training and validation data sets as predicted at each of the stages of training: the prediction of the step-0 neural-network potential show a profusion of outliers pointing to poor transferability, while those of step 3 are tightly grouped around the perfect-prediction line.

On a quantitative level, the accuracy of the NNFF is reflected in root mean square errors (RMSEs) of $0.0194\,\text{eV atom}^{-1}$ and $0.0249\,\text{eV Å}^{-1}$ for the training energies and the training forces, respectively. For the validation set, the corresponding RMSEs for energies and atomic forces are $0.0196\,\text{eV atom}^{-1}$ and $0.0250\,\text{eV Å}^{-1}$, respectively (see Fig. S1 in the Supplemental Material). To further assess the model's predictive capabilities, we extended the comparison to include additional trilayer and tetralayer vdWHS systems. As shown in Fig. S2 in the Supplemental Material, the RMSEs values for atomic forces and energies are approximately $0.031\,\text{eV Å}^{-1}$ and $0.023\,\text{eV atom}^{-1}$, respectively.

C. First-principles calculations

Ab-initio computations of atomic forces and energies were performed in the framework of density functional theory (DFT) using the Vienna Ab-initio Simulation Package (VASP) [28], the Perdew-Burker-Ernzerhof generalized-gradient approximation [29] to the exchange and correlation components of the electronic Hamiltonian. In order to improve the description of the interlayer interactions, the DFT-D3 of Grimme's empirical dispersion interaction was included in the calculations
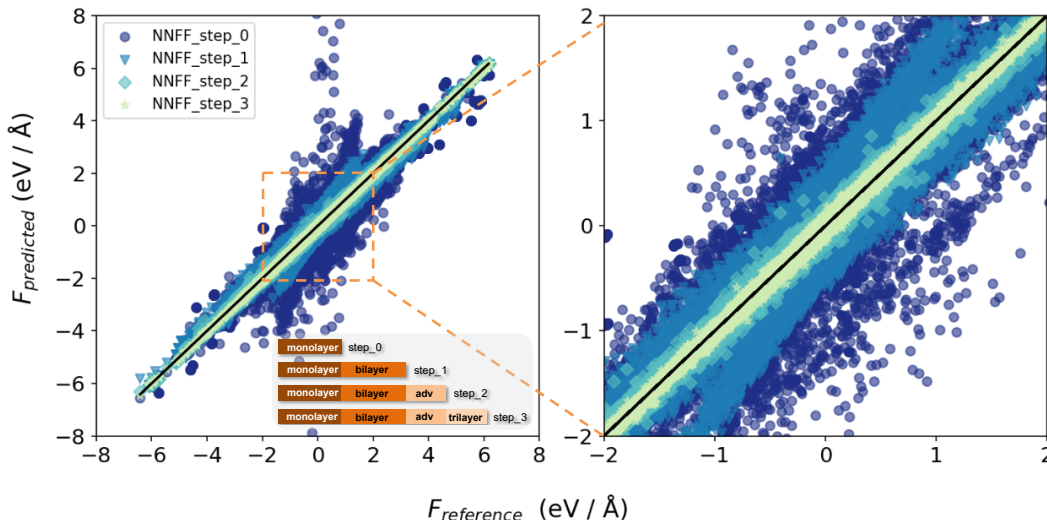
FIG. 2. Comparison between the forces on atoms predicted by the NNFF model and ground-truth values from DFT calculations. The right panel contains a zoomed-in view of the area enclosed by the orange dashed box in left panel. The inset summarizes the stepwise construction of the dataset.

[2, 30–32]. All electronic self-consistent-field calculations were run with a convergence criterion of 1×10^{-8} eV and a plane-wave energy cutoff of 550 eV. Our optimized structural parameters for the monolayer are listed in Tbl. I; we re-optimized them for the multilayer structures but found them to change very little. We also list the the equilibrium interlayer distances for homogeneous stacking sequences. For the heterogeneous stacks we obtained a minimum of 2.39 Å and a maximum of 2.42 Å.

TABLE I. Lattice parameter (a), Pt-S and Pt-Te bond lengths and interlayer distance (D) of monolayers and homogeneous multilayers, compared with literature values.

Materials	a (Å)	$d_{\mathrm{Pt-S}}$ (Å)	$d_{\mathrm{Pt-Te}}$ (Å)	D (Å)
PtS_2	3.58 (3.57^a)	2.40 (2.42^b)	–	2.25
PtSTe	3.80 (3.80^c)	2.50 (2.49^c)	2.63 (2.62^c)	2.30
$PtTe_2$	4.02 (4.01^d)	–	2.70	2.39

^{a} From Ref. [33], ^{b} From Ref. [34].
^{c} From Ref. [35].
^{d} From Ref. [36].

D. Calculation of the thermal conductivity using the BTE

We obtain $\boldsymbol{\kappa}$ in the framework of the linearized BTE as

$$\kappa^{\alpha\beta} = \frac{k_{\mathrm{B}}}{N_q \Omega_{\mathrm{uc}}} \sum_\lambda \left(\frac{\hbar \omega_\lambda}{k_{\mathrm{B}} T} \right)^2 n_\lambda^0 (n_\lambda^0 + 1) v_\lambda^\alpha F_\lambda^\beta, \quad (3)$$

where λ runs over all phonon branches and over all N_q points on a regular grid covering the Brillouin zone, each phonon mode is characterized by its angular frequency ω_λ, its group velocity $\boldsymbol{v}_\lambda$ and its equilibrium (Bose-Einstein) occupancy n_λ, and k_{B}, Ω_{uc} and T are the Boltzmann constant, the volume of unit cell and the temperature, respectively. α and β denote Cartesian axes, and $\boldsymbol{F}_\lambda$ is a the vector of coefficients that connects the deviation from equilibrium of each mode's occupancy to the temperature gradient in this linear regime.

The details of the workflow are documented in detail in Ref. 12. The essential ingredients are sets of second-and third-order derivatives of the potential energy evaluated at the equilibrium positions, the so-called interatomic force constants (IFCs). Second-order IFCs are required to calculate harmonic quantities like ω_λ and $\boldsymbol{v}_\lambda$, while the third-order IFCs characterize the strength of three-phonon scattering, the dominant influence determining $\boldsymbol{F}_\lambda$. We obtain the second-order IFCs using a finite-displacement method as implemented in Phonopy [37], with either VASP or our NNFF as the backend to calculate forces. We further use hiPhive [38] as a post-processing step to enforce the rotational symmetry of free space and obtain the proper quadratic behavior of the ZA branches in these quasi-2D systems [39]. To calculate the anharmonic third-order IFCs we employ Phono3py [40, 41] with $4 \times 4 \times 1$ supercells, a finite displacement of 0.03 Å and a cutoff distance of 6.0 Å (up to 6th nearest neighbors). We also use Phono3py to solve the BTE and calculate κ, evaluating Eq. (3) on a q-point mesh of $30 \times 30 \times 1$.

E. Calculation of the thermal conductivity through equilibrium MD

The thermal conductivity can be obtained using the Green-Kubo formula for the time autocorrelation function of the heat flux $\boldsymbol{J}$:

$$\kappa^{\alpha\beta} = \frac{\Omega_{\text{sim}}}{k_{\text{B}}T^2} \int_0^\infty \left\langle J^\alpha(t)J^\beta(0) \right\rangle dt, \quad (4)$$

where Ω_{sim} is the volume of the simulation box. This integral can also be extracted as the zero-frequency component of the power spectrum of $\boldsymbol{J}$. However, both the time-domain and frequency-domain formulations face convergence problems. Not only do those require costly long MD trajectories, but they can also introduce artifacts in the results even in that case. Moreover, the definition of $\boldsymbol{J}$ for a general form of the interaction potential is, a priori, ambiguous, and depends on how energy is apportioned among atoms. Work by Baroni and coworkers [20, 21] has helped overcome both classes of obstacles. Firstly, it shows the existence of a gauge principle guaranteeing that different partitioning strategies for the potential energy lead to the same estimate of the thermal conductivity. Secondly, it presents an efficient algorithm to obtain that estimate, along with a measure of its uncertainty, from short trajectories. Moreover, it puts the whole formalism on a firm footing for periodic multicomponent systems even in the presence of diffusion. The central quantity in this formulation is the cepstrum of $\boldsymbol{J}$, i.e., the inverse Fourier transform of the logarithm of its spectrum. Denoting the power spectrum of $\boldsymbol{J}$ at linear frequency f by $\boldsymbol{S}(f)$, the thermal conductivity is reconstructed from the cepstral components of S as

$$\kappa = \frac{\Omega_{\text{sim}}}{2k_{\text{B}}\text{T}^2} \exp[\log S(f=0)]$$
$$\simeq \frac{\Omega_{\text{sim}}}{2k_{\text{B}}\text{T}^2} \exp\left[\hat{C}_0 + 2\sum_{n=0}^{P^*-1} \hat{C}_n - \left\langle \log \hat{\xi}_0 \right\rangle\right], \quad (5)$$

where the Cartesian indices have been dropped to avoid cluttering the notation. Here, $\hat{C}_0 \dots \hat{C}_{P^*-1}$ are the aforementioned cepstral components, and the cutoff P^* is chosen objectively according to the Akaike information criterion. $\hat{\xi}_0$ is one of the independent and identically distributed random variables connecting the sample power spectrum with its expected value in this formulation, and the $\left\langle \log \hat{\xi}_0 \right\rangle$ term is calculated as $\psi(\ell - M + 1) - \log(\ell - M + 1)$, where ψ is the digamma function [42], N is the number of time steps, M is the number of conserved fluxes in the calculation, and ℓ is the number of Cartesian components of the flux (e.g. two in our subperiodic systems). The uncertainty in $\boldsymbol{\kappa}$ is estimated from the expression

$$\frac{\Delta\kappa}{\kappa} = \sigma_0\sqrt{\frac{4P^* - 2}{N}}, \quad (6)$$

where $\sigma_0^2 = \psi'(\ell - M + 1)$ and ψ' is the trigamma function.

We perform our MD simulations using a time step of $1\,\text{fs}$ and $5 \times 5 \times 1$ supercells. For each system, we start by running a canonical (NVT) trajectory for $200\,\text{ps}$ to equilibrate the system at room temperature ($300\,\text{K}$). We then run a $2500\,\text{ps}$-long microcanonical (NVE) trajectory during which we calculate the heat flux as

$$\boldsymbol{J}(t) = \boldsymbol{J}_{\text{kin}} + \boldsymbol{J}_{\text{pot}}$$
$$= \frac{1}{\Omega_{\text{sim}}} \left[\sum_i \left(\phi_i + \frac{1}{2}m_i\boldsymbol{v}_i^2\right)\boldsymbol{v}_i - \sum_{ij}(\boldsymbol{r}_i - \boldsymbol{r}_j)\frac{\partial\phi_j}{\partial\boldsymbol{r}_i} \cdot \boldsymbol{v}_i\right]. \quad (7)$$

In this expression, i and j run over all atoms in the system, whose masses, positions, velocities and contributions to the potential energy are denoted by m_i, $\boldsymbol{r}_i$, $\boldsymbol{v}_i$ and ϕ_i, respectively. We take ϕ_i directly from the output of the second stage of the NNFF, before the sum over atoms. Just like the total forces, the detailed Jacobian $\frac{\partial\phi_j}{\partial\boldsymbol{r}_i}$ is calculated using algorithmic differentiation. The velocity Verlet and thermostat routines from the atomic simulation environment (ASE [43]) take care of the integration steps.

Figure 3 shows, in detail, a representative example of our results, specifically for the PtSTe monolayer. Panel (a) contains a plot of the sample power spectrum as a function of the frequency. Following the prescription by Baroni and coworkers [20], we choose a cutoff frequency (f^*) for the cepstral analysis that leaves below it the main prominent feature of the smoothed power spectrum; that frequency is signaled by a vertical arrow. As expected from this method, the values of κ and its uncertainty converge quickly over the course of the simulation: this is illustrated by panel (b).

III. RESULTS AND DISCUSSION

As a first test of the applicability of our NNFF to the vibrational dynamics of vdWHSs, we show a comparison between the NN- and DFT-backed phonon spectra of PtSTe$_{\parallel}$ and the PtSTe monolayer in the top panels of Fig. 4. PtSTe$_{\parallel}$ denotes a PtSTe bilayer where the second layer is vertically flipped with respect to the first and its structure is depicted in the inset of Fig. 4(d). The comparison reveals a remarkable level of agreement between the two approaches, validating the accuracy and reliability of the calculations based on the NNFF potential. This agreement is observed both in the high-frequency range and in the low-frequency regions of the branches close to the Γ point. For the bilayer, both acoustic and low-lying optical modes can be observed close to the Γ point, reflecting that interlayer interactions are also captured. Despite the low strength of those interactions,
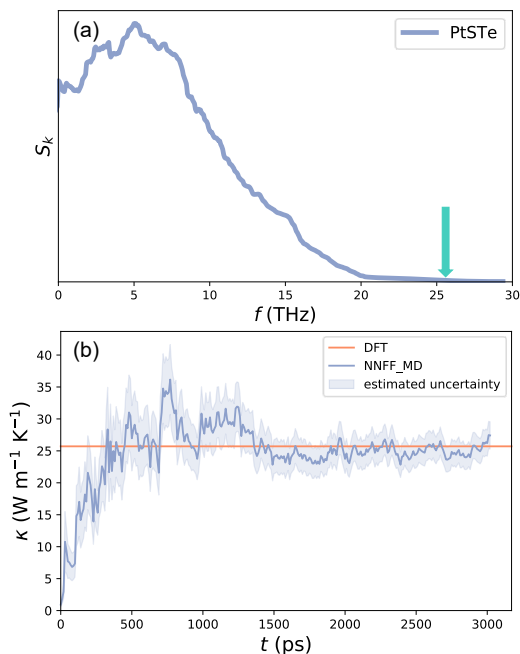
FIG. 3. Average sample power spectrum of the heat flux computed from MD trajectories for monolayer PtSTe (with a length of 100 ps) (a) and convergence of the thermal conductivity of PtSTe calculated from an NNFF-backed MD simulation (b). The orange line indicates the thermal conductivity obtained from the DFT-backed BTE calculation.

they are enough to make phonon modes more extended. The weaker average interactions depress the overall frequencies, and the high-energy optical bands become more dispersive as the more intense hybridation distributes those vibrations over larger sets of atoms. Accordingly, the gaps between bunches of phonon branches are reduced.

The bottom panels (c and d) of Fig. 4 show the other main ingredient for the study of thermal transport in the BTE framework outlined above, i.e. the three-phonon scattering rates, for the same systems. Although some deviations from DFT are observed in areas where the results depend on the subtler features of the branches, the NNFF also performs well in this test, especially taking into account the fact that these scattering rates span many order of magnitude. The aggregate effect of those deviations will be assessed when both sets of scattering rates are used for the thermal conductivity calculation below. While both limits are reproduced by the NNFF, it is noticeable that the monolayer and the bilayer stand in stark contrast in terms of their behavior in the $\omega \longrightarrow 0$ limit. Only the rates for the monolayer diverge towards infinity, which can be explained in terms of the out-of-plane symmetry properties of each system, just like in the case of monolayers [10, 44]. Specifically, the selection rules protecting the ZA branch from scattering in the PtS_2 and $PtTe_2$ monolayers are lifted in the asymmetric

Janus PtSTe monolayer [9]. These are restored due to the symmetric arrangement of the $PtSTe_{\uparrow\downarrow}$ bilayer, which shows the expected vanishing behavior in the $\omega \longrightarrow 0$ limit.

After this detailed look at the ability of the NNFF to reproduce the harmonic and anharmonic features of lattice vibrations in our systems, in Fig. 5 we provide a more global comparison of the three methods used here when it comes to predicting our main quantity of interest, the lattice thermal conductivity. Included in the figure are results at $T = 300$ K for all the homogeneous systems; the thermal conductivity of the monolayers and bilayers was calculated using the MD and BTE workflows, with the latter backed by both direct DFT calculations and by the ML potential, while for the trilayers we only ran MD simulations. Additionally, we include experimental results for PtS_2 from Ref. 45. The level of agreement between the NNFF and DFT results is very satisfactory, again lending support to the use of our NNFF as a surrogate model for lattice thermal conductivity calculations. Furthermore the agreement between MD and BTE based calculations suggests a limited effect of higher-order anharmonicity at 300 K despite the noncovalent bonding between layers. For the two cases where experimental data is available, its values are correctly predicted by our workflow, with a slight overestimation consistent, for instance, with some elastic phonon scattering coming from the residual defects present in the samples.

The thermal conductivity of the monolayers follows the sequence $PtS_2 > PtTe_2 > PtSTe$, as previously analyzed in Ref. 9. This order emerges from the combination of the higher mass of tellurium and the breakdown of the symmetry-induced selection rules in the Janus monolayer. The thermal conductivity of the multilayers is lower than that of the corresponding monolayer; this can be attributed to an increase in scattering channels caused by the breakdown of degeneracy of vibrational modes mediated by interlayer interactions. The decrease is more noticeable from the monolayer to the bilayer, with the trilayer already approaching the limit of this effect. Despite this general trend, there is a crucial dependence of the thermal conductivity on the stacking pattern even among homogenous multilayers, evident in the case of PtS_2 where we predict that an AA monolayer carries heat almost twice as efficiently as the experimentally characterized AB bilayer [details of AB stacking can be found in Fig. S3 of Supplemental Material]. Further stacking of layers is less effective at hindering phonons in Janus PtSTe than in the symmetric parent structures. For instance, in $PtTe_2$ we find that $\kappa_{\text{bilayer}} = 0.57\kappa_{\text{monolayer}}$ and $\kappa_{\text{trilayer}} = 0.90\kappa_{\text{bilayer}}$, but in PtSTe the approximate ratios are $\kappa_{\text{bilayer}} = 0.90\kappa_{\text{monolayer}}$ and $\kappa_{\text{trilayer}} = 0.93\kappa_{\text{bilayer}}$. This matches the general observation that it is much harder to depress the thermal conductivity in systems with already intense phonon scattering.

Analyzing the temperature dependence of the thermal conductivity (Fig. 6) sheds more light both on the performance of the different methods employed here and on
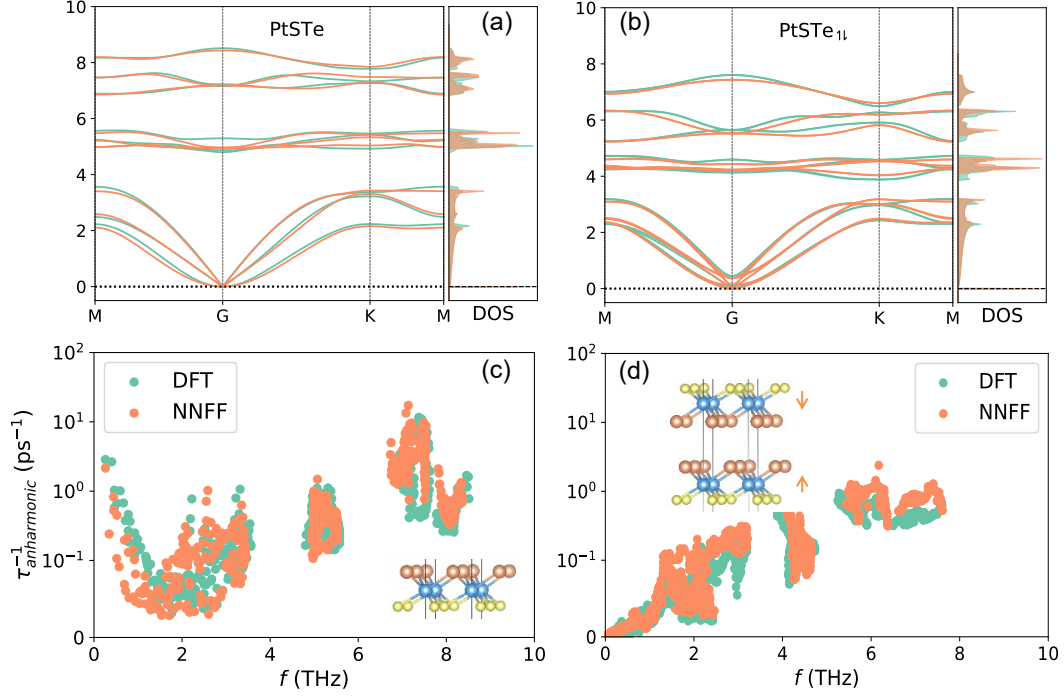
FIG. 4. Phonon dispersions and anharmonic scattering rates of monolayer PtSTe (a,c) and PtSTe$_{1\downarrow}$ (b,d). The insets in panels (c) and (d) contain side views of the corresponding structures.
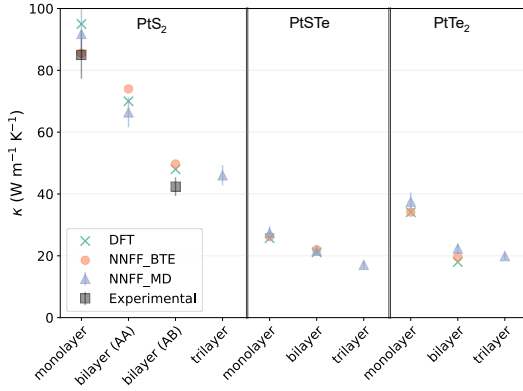

FIG. 5. Room-temperature κ for homogeneous systems obtained using three different methods: the BTE based on DFT- (crosses) or NNFF-calculated (dots) IFCs and equilibrium MD (triangles). The experimental points (squares) come from Ref. 45.

the physics of phonons in these materials. When anharmonicity is only included in the form of three-phonon processes, κ decreases with temperature according to the well known $\propto T^{-1}$ dependence. This is evident in both panels of Fig. 6, which present the data for PtSTe (a) and PtSTe$_{1\downarrow}$ (b). An almost perfect match between the DFT- and NNFF-backed BTE calculations is found for

the monolayer, while for the bilayer the curves are still parallel but offset with respect to each other. Although the agreement is still fair, this points to some degree of difficulty of the NN model when it comes to reproducing the finer details of the weak interlayer interactions. Perhaps more interestingly, the MD curves depart from the $\propto T^{-1}$ trend and show a more accentuated decay with temperature, with exponents around -1.15 and -1.2 for the monolayer and bilayer, respectively. It is reasonable to infer a nonnegligible importance of n-phonon processes with $n > 3$ for $T > 300$ K. The significance of higher-order IFCs, including four-phonon scattering, on the thermal conductivity at high temperatures has been noted in previous research, such as the work by Yang et al. [46]. Their findings suggest that four-phonon scattering contributes to a deviation towards a T^{-2} trend, resulting in a more substantial decrease in thermal conductivity as temperature increases. Here the importance of those processes may be enhanced by van-der-Waals interactions; however, the overall quantitative change in the predicted κ due to this departure is still less than 10% in the temperature range studied, and hence the BTE results remain predictive. A feature of this disagreement that may not be explained by resistive scattering alone is the higher value of the MD conductivity at lower temperatures in Fig. 6 (b) with respect to the BTE solution using the same force field. A first hypothesis about its possible origin is the use of Bose-Einstein occupations in the BTE treatment, which affect both the mode spe-
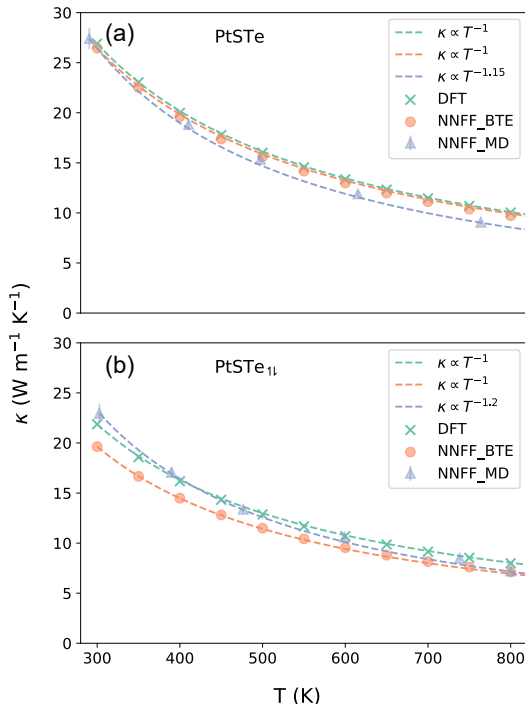
FIG. 6. Temperature-dependent thermal conductivity of monolayer PtSTe (a) and $PtSTe_{1\parallel}$ (b) obtained using the three different methods discussed in the main text, along with power-law fits to the temperature dependence (dashed lines).

cific heat and the scattering rates, versus the classical (equipartition) occupations that would emerge from MD. To test this possibility, we created a patch for Phono3py to use equipartition occupations [47]. A plot illustrating the change this introduces, as well as a modified version of Fig. 6 (b), are provided as part the Supplemental Materials; see our data availability statement for details on how to obtain the patch. Despite the increased mode specific heat, the net effect of equipartition occupations is to decrease the thermal conductivity, and therefore this effect cannot be responsible for the differences between the BTE and MD around room temperature in Fig. 6 (b). A more plausible alternative is therefore the contribution to thermal transport from diffusons [19–22], which has been shown to be non-negligible even in ordered systems .

At $300\,\mathrm{K}$, where the more prominent temperature dependence plays an even smaller role, MD results can be used interchangeably with BTE results as computational estimates of κ for this family of systems. We therefore use MD to complete our study of the family with more complex structures. The complete set of room-temperature results is displayed in Fig. 7. With respect to Fig. 5, this now includes heterogeneous multilayer (bilayer and trilayer) vdWHSs.
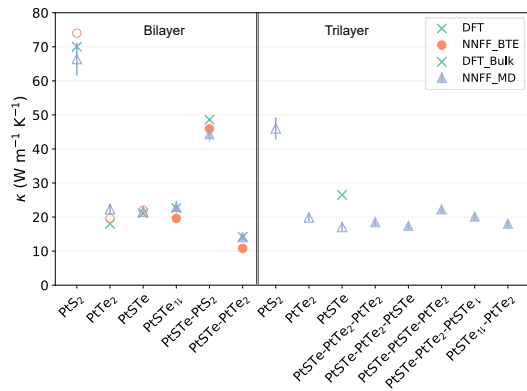

FIG. 7. The thermal conductivity of multilayer vdWHs calculated through a solution of the BTE based on DFT-(crosses) or NNFF-calculated (dots) IFCs, and through MD based on the NNFF (triangles). Empty symbols correspond to homogeneous systems; solid symbols denote heterogeneous structures. The bulk value for PtSTe was evaluated using the DFT+BTE workflow and added for reference.

Among the observations that can be drawn from Fig. 7 is the fact that the subseries containing PtS_2 tends to have a higher thermal conductivity, just like the PtS_2 monolayer itself. As can be seen in the details of the phonon dispersion included in the Supplemental Material, this can be traced directly to the phonon spectrum and ultimately to the lower atomic mass of sulphur. On the other hand, the values of κ for the $PtTe_2$ and Janus-PtSTe series of bilayers are mostly around $20 \pm 2\ \mathrm{W\,m^{-1}\,K^{-1}}$. The $PtSTe-PtTe_2$ combination exhibits the lowest thermal conductivity, with a value of $14 \pm 0.9\ \mathrm{W\,m^{-1}\,K^{-1}}$. However, adding more layers to the structure does not lead to an even lower thermal conductivity, which instead oscillates in a range of values close to the bulk structure. This observation can be rationalized based on the notion that, when interlayer interactions are weak, layers quickly reach a bulk-like environment, but with a residual confinement effect.

IV. CONCLUSIONS

In summary, we introduce a NNFF model, trained on a relatively small dataset derived from DFT calculations. This model accurately predicts the thermal-transport properties of layered van der Waals heterostructures (vd-WHSs) based on PtSTe, employing both the linearized phonon BTE method and equilibrium Green-Kubo MD simulations. Our results demonstrate that NNFF predictions effectively capture van der Waals interactions, showing excellent agreement with DFT calculations and experimental data at room temperature. Importantly, NNFF offers a substantial computational speed advantage over traditional ab initio calculations. For instance, while one time step on 20 cores for a system with 75

atoms takes approximately 3000 s for DFT to calculate atomic forces, NNFF predicts those in just 0.31 s.

Furthermore, our study emphasizes the advantage of MD simulations in capturing higher-order anharmonic properties, particularly at high temperatures. The temperature-dependent thermal conductivity curves of monolayer PtSTe and bilayer $PtSTe_{\uparrow\downarrow}$ both deviate from $\kappa \propto T^{-1}$, with decay ratios of -1.15 and -1.2, respectively. This departure suggests that the method employed can capture behavior beyond third-order scattering. Notably, among the PtSTe-based vdWHSs, PtSTe-$PtTe_2$ stands out with its remarkably low thermal conductivity. This behavior can be attributed to several factors, including broken out-of-plane symmetries and weak van der Waals interactions. However, it is important to note that trilayer vdWHSs do not exhibit further decreases in thermal conductivity. This is mainly due to weak interlayer interactions, leading to oscillations in thermal conductivity within the range of values observed in the corresponding bulk materials.

V. DATA AVAILABILITY STATEMENT

ACKNOWLEDGMENTS

[1] S. Bharadwaj, A. Ramasubramaniam, and L. R. Ram-Mohan, Nanoscale **14**, 11750 (2022).

[2] C. Wang, Y.-X. Chen, G. Gao, K. Xu, and H. Shao, Appl. Surf. Sci. **593**, 1 (2022).

[3] A. Patel, D. Singh, Y. Sonvane, P. B. Thakor, and R. Ahuja, ACS Appl. Mater. Interfaces **12**, 46212 (2020).

[4] H. T. T. Nguyen, V. T. T. Vi, T. V. Vu, N. V. Hieu, D. V. Lu, D. P. Rai, and N. T. T. Binh, RSC Adv. **10**, 44785 (2020).

[5] R. Gupta, B. Dongre, C. Bera, and J. Carrete, J. Phys. Chem. C Nanomater Interfaces **124**, 17476 (2020).

[6] S. D. Guo, Phys. Chem. Chem. Phys. **20**, 7236 (2018).

[7] R. Chaurasiya, S. Tyagi, N. Singh, S. Auluck, and A. Dixit, J. Alloys Compd. **855**, 157304 (2021).

[8] A. Y. Lu, H. Zhu, J. Xiao, C. P. Chuu, Y. Han, M. H. Chiu, C. C. Cheng, C. W. Yang, K. H. Wei, Y. Yang, Y. Wang, D. Sokaras, D. Nordlund, P. Yang, D. A. Muller, M. Y. Chou, X. Zhang, and L. J. Li, Nat. Nanotechnol. **12**, 744 (2017).

[9] L. Pan, Z. Wang, J. Carrete, and G. K. H. Madsen, Phys. Rev. Mater. **6**, 084005 (2022).

[10] L. Lindsay, D. A. Broido, and N. Mingo, Phys. Rev. B **83**, 235428 (2011).

[11] R. Hu, Z. Zhou, C. Sheng, S. Han, H. Yuan, and H. Liu, Phys. Rev. A **430**, 1 (2022).

[12] W. Li, J. Carrete, N. A. Katcho, and N. Mingo, Comput. Phys. Commun. **185**, 1747 (2014).

[13] J. Carrete, B. Vermeersch, A. Katre, A. van Roekeghem, T. Wang, G. K. H. Madsen, and N. Mingo, Comput. Phys. Commun. **220**, 351 (2017).

[14] H. Montes-Campos, J. Carrete, S. Bichelmaier, L. M. Varela, and G. K. H. Madsen, J. Chem. Inf. Model. , 88 (2021).

[15] J. Carrete, H. Montes-Campos, R. Wanzenbock, E. Heid, and G. K. H. Madsen, J. Chem. Phys. **158**, 204801 (2023).

[16] L. Isaeva, G. Barbalinardo, D. Donadio, and S. Baroni, Nat. Commun. **10**, 3853 (2019).

[17] M. Simoncelli, N. Marzari, and F. Mauri, Nat. Phys. **15**, 809 (2019).

[18] A. Fiorentino, P. Pegolo, and S. Baroni, Npj Comput. Mater. **9**, 157 (2023).

[19] P. B. Allen and J. L. Feldman, Phys. Rev. B **48**, 12581 (1993).

[20] L. Ercole, A. Marcolongo, and S. Baroni, Sci. Rep. **7**, 15835 (2017).

[21] S. Baroni, R. Bertossa, L. Ercole, F. Grasselli, and A. Marcolongo, Heat transport in insulators from ab initio green-kubo theory, in *Handbook of Materials Modeling* (2018) Book section Chapter 12-1, pp. 1–36.

[22] P. Pegolo, S. Baroni, and F. Grasselli, npj Computational Materials **8**, 24 (2022).

[23] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).

[24] E. Kocer, J. K. Mason, and H. Erturk, AIP Adv. **10**, 015021 (2020).

[25] P. Ramachandran, B. Zoph, and Q. V. Le, arXiv preprint arXiv:1710.05941 (2017).

[26] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, JAX: composable transformations of Python+NumPy programs (2018).

[27] L. Metz, J. Harrison, C. D. Freeman, A. Merchant, L. Beyer, J. Bradbury, N. Agrawal, B. Poole, I. Mordatch, A. Roberts, and J. Sohl-Dickstein, arXiv preprint arXiv:2211.09760 , 1 (2022).

[28] G. Kresse and J. Furthmuller, Phys. Rev. B Condens. Matter **54**, 11169 (1996).

[29] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[30] X. Zhao, M. Dang, W. Niu, Q. Zhao, X. Dai, and S. Wei, Phys. E: Low-Dimens. Syst. Nanostructures. **124**,

1 (2020).

[31] H. Zhao, F. Xie, Y. Liu, B. Bian, G. Yang, Y. Ding, Y. Gu, Y. Yu, X. Zhang, X. Huo, B. Hua, X. Ni, Q. Fan, and X. Gu, Mater. Sci. Semicond. **123**, 105588 (2021).

[32] T. Morawietz and J. Behler, J. Phys. Chem. A **117**, 7356 (2013).

[33] X. C. Liang, X. J. He, Y. C. Ding, Y. j. Hao, and J. Zhu, Phys. E: Low-Dimens. Syst. Nanostructures. **132**, 114744 (2021).

[34] G. Liu, Y. Gan, R. G. Quhe, and P. F. Lu, Chem. Phys. Lett. **709**, 65 (2018).

[35] P. A. L. Sino, L.-Y. Feng, R. A. B. Villaos, H. N. Cruzado, Z.-Q. Huang, C.-H. Hsu, and F.-C. Chuang, Nanoscale Adv. **3**, 6608 (2021).

[36] B. Lei, Y. Y. Zhang, and S. X. Du, Chin. Phys. B **29**, 058104 (2020).

[37] A. Togo and I. Tanaka, Scr. Mater. **108**, 1 (2015).

[38] F. Eriksson, E. Fransson, and P. Erhart, Adv. Theory Simul. **2**, 1800184 (2019).

[39] J. Carrete, W. Li, L. Lindsay, D. A. Broido, L. J. Gallego, and N. Mingo, Mater. Res. Lett. **4**, 204 (2016).

[40] A. Togo, L. Chaput, and I. Tanaka, Phys. Rev. B **91**, 094306 (2015).

[41] A. Togo, J. Phys. Soc. Jpn. **92**, 012001 (2023).

[42] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Dover, New York, 1964).

[43] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, J. Phys.: Condens. Matter **29**, 273002 (2017).

[44] R. Yang, S. Yue, Y. Quan, and B. Liao, Phys. Rev. B **103**, 184302 (2021).

[45] H. A. H. Mohammed, G. M. Dongho-Nguimdo, and D. P. Joubert, Mater. Today Commun. **21**, 1 (2019).

[46] X. Yang, J. Tiwari, and T. Feng, Mater. Today Phys. **24**, 1 (2022).

[47] Y. He, I. Savić, D. Donadio, and G. Galli, Phys. Chem. Chem. Phys. **14**, 16209 (2012).