

# Machine-Learning-Based Interatomic Potentials for Group IIB to VIA Semiconductors: Toward a Universal Model

Jianchuan Liu,<sup>△</sup> Xingchen Zhang,<sup>△</sup> Tao Chen, Yuzhi Zhang, Duo Zhang, Linfeng Zhang, and Mohan Chen\*



Cite This: *J. Chem. Theory Comput.* 2024, 20, 5717–5731



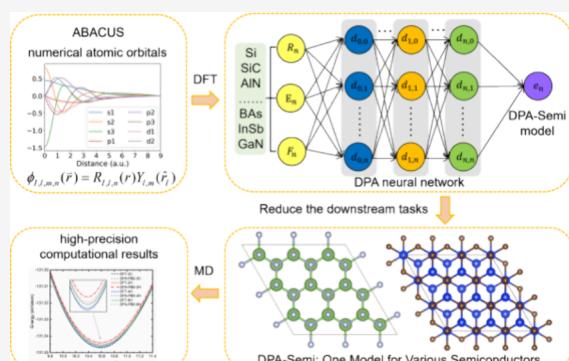
Read Online

ACCESS |

Metrics & More

Article Recommendations

**ABSTRACT:** Rapid advancements in machine-learning methods have led to the emergence of machine-learning-based interatomic potentials as a new cutting-edge tool for simulating large systems with ab initio accuracy. Still, the community awaits universal interatomic models that can be applied to a wide range of materials without tuning neural network parameters. We develop a unified deep-learning interatomic potential (the DPA-Semi model) for 19 semiconductors ranging from group IIB to VIA, including Si, Ge, SiC, BAs, BN, AlN, AlP, AlAs, InP, InAs, InSb, GaN, GaP, GaAs, CdTe, InTe, CdSe, ZnS, and CdS. In addition, independent deep potential models for each semiconductor are prepared for detailed comparison. The training data are obtained by performing density functional theory calculations with numerical atomic orbitals basis sets to reduce the computational costs. We systematically compare various properties of the solid and liquid phases of semiconductors between different machine-learning models. We conclude that the DPA-Semi model achieves GGA exchange-correlation functional quality accuracy and can be regarded as a pretrained model toward a universal model to study group IIB to VIA semiconductors.



## 1. INTRODUCTION

Semiconductor materials play a crucial role in the development of modern society. In particular, the IIB to VIA group semiconductors, which are compound semiconductors composed of elements from groups IIB to VIA of the periodic table, possess excellent optoelectronic properties and are widely used in photovoltaic, optoelectronics, thermoelectrics, and other energy conversion fields.<sup>1–4</sup> For example, silicon carbide (SiC) has found widespread industrial applications because of its excellent wear resistance, corrosion resistance, elevated temperature strength, as well as its high thermal conductivity and wide band gap.<sup>5–8</sup> Boron arsenide (BAs) was initially synthesized in 1958<sup>9</sup> but was recently confirmed to possess high charge carrier mobility and thermal conductivity.<sup>10–12</sup> Therefore, there is a promising prospect of utilizing this material to alleviate the current bottleneck issue in chip cooling. On the other hand, state-of-the-art simulation tools can complement experiments by elucidating experimental phenomena or predicting experimental outcomes, thereby providing invaluable information or better design principles.

Among the simulation tools, the atomistic-level simulation tools can describe interactions of semiconductors from a microscopic perspective, providing valuable insights into the fundamental processes governing the behavior of semiconductor materials. In particular, quantum-mechanics-based first-

principles methods are able to predict the properties of semiconductors without reliance on experimental data. Among them, density functional theory (DFT)<sup>13,14</sup> is one of the most widely used methods that can predict various properties of semiconductor materials.<sup>15–17</sup> Taking SiC as an example, it exists in various polytypes such as the 3C, 2H, 4H, and 6H structures, etc. Among them, the 3C structure is a cubic crystal lattice, while the 2H, 4H, and 6H polytypes refer to hexagonal crystal lattices with different stacking sequences of Si–C bilayers.<sup>18–26</sup> The relationship between atomic structures and mechanical properties in single crystal and polycrystalline solid phases,<sup>22,27,28</sup> as well as the stability and mobility of screw dislocations in 3C, 2H, and 4H-SiC can be calculated by DFT.<sup>8</sup> In addition, the compression mechanical properties<sup>25</sup> and the elastic properties<sup>26</sup> of 3C, 4H, and 6H-SiC at ambient and high pressure, and the electronic structures and optical properties of vacancy-doped 3C-SiC systems were also studied by DFT.<sup>29</sup>

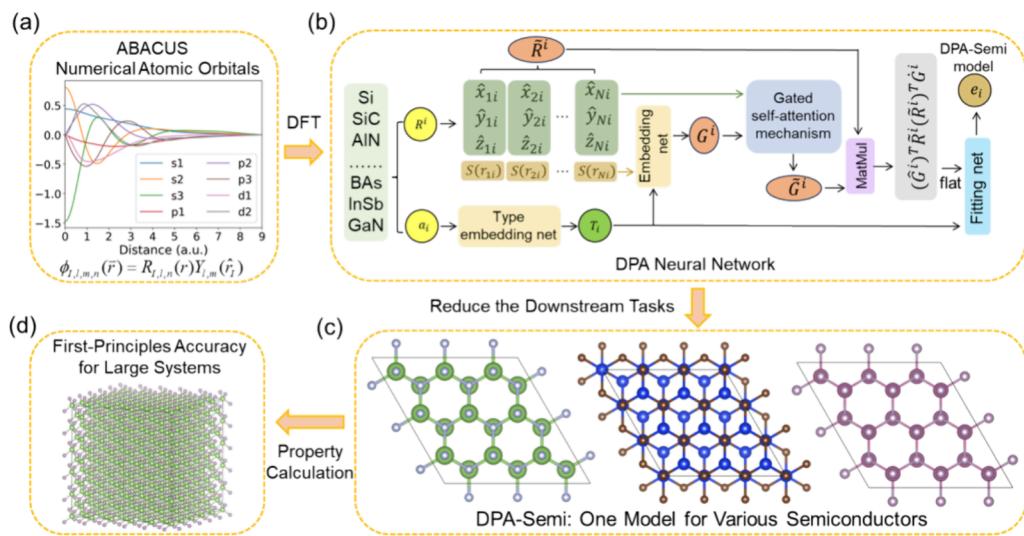
Received: December 1, 2023

Revised: June 2, 2024

Accepted: June 10, 2024

Published: June 20, 2024





**Figure 1.** Procedures for developing the DPA-Semi model. (a) Generate atomic data sets using the ABACUS package based on the numerical atomic orbitals as basis set; (b) Generate the DPA-Semi model via the Gated self-attention mechanism based on the DFT atomic data sets; (c) DPA-Semi model can be used for various kinds of semiconductors, and reduce the computational costs of downstream tasks; and (d) DPA-semi model is readily applied to calculate properties of large-systems with GGA quality accuracy.

However, simulations of semiconductor materials for industrial applications, which often demand a large number of atoms (typically millions or more), remain a formidable challenge when using the DFT method. Recently, rapid advancements in machine-learning methods have led to the emergence of machine-learning-based interatomic potentials. Among them, the deep potential (DP) model<sup>30</sup> with first-principles accuracy<sup>31–36</sup> has emerged as a cutting-edge tool for simulating large-size systems<sup>37</sup> and long-time-scale properties across a wide range of materials. Taking semiconductor materials as examples, all of the following works adopted the DP models. The thermal conductivity for the crystalline, liquid, and amorphous phases of Si,<sup>38</sup> as well as the thermal conductivity and phonon transport properties of  $\beta$ -Ga<sub>2</sub>O<sub>3</sub> were accurately predicted.<sup>39</sup> In addition, the temperature-dependent microwave dielectric permittivity of  $\beta$ -Ga<sub>2</sub>O<sub>3</sub> was calculated.<sup>40</sup> For the SiC materials, the thermal transport and mechanical properties were systematically investigated,<sup>41</sup> and the infrared resonance frequency and phonon line width were accurately predicted.<sup>20</sup> The DP models were also utilized in research related to heat transfer of semiconductor interfaces such as the Si/Ge interface.<sup>42</sup> These research outcomes demonstrate that the machine-learning-based DP method plays a key role in atomistic simulations of semiconductor materials.

However, when addressing similar properties of a wide range of materials, generating a DP model for each material is not only computationally demanding but also considerably time-consuming. In addition, some inconsistencies may exist among the models due to the different training samples used. In this regard, there is a strong demand in the community for generating a universal model that can simulate a wide range of materials. In 2022, a deep potential model with an attention mechanism was proposed by Zhang et al.,<sup>43</sup> i.e., the DPA-1 model. The model was trained on a large number of atomic data sets in terms of a variety of elements and showed satisfactory accuracy. Nowadays, the DPA-1 model can be readily applied to study real scientific issues with a small amount of additional effort.

Machine learning interatomic potential functions such as the DP and DPA-1 models are typically trained on a large number of DFT data sets. Typically, DFT calculations are performed with the Plane-Wave (PW) basis set.<sup>27</sup> The PW basis set has several advantages. For instance, the accuracy of DFT calculations can be controlled by a single value of energy cutoff. In addition, the PW basis sets are orthogonal, and no Pulay forces need to be evaluated. Notably, the PW basis set exhibits unfavorable scaling when the system size is large, typically around a few hundred atoms, resulting in large computational costs for generating training data. In this work, to reduce the computational costs of DFT, numerical atomic orbitals (NAO)<sup>44–48</sup> as a basis set are adopted to solve the Kohn–Sham equation. Specifically, we utilize the Atomic-orbital Based Ab-initio Computation at UStc (ABACUS) package,<sup>49</sup> which supports the NAO basis set.<sup>44,45</sup> The computational costs are relatively smaller than the PW basis set. The NAO basis set has been used in several applications and is suitable for studying large systems.<sup>50–53</sup>

In this work, we generated first-principles data for 19 bulk semiconductors ranging from group IIB to VIA, namely, Si, Ge, SiC, BAs, BN, AlN, AlP, AlAs, InP, InAs, InSb, GaN, GaP, GaAs, CdTe, InTe, CdSe, ZnS, and CdS. We used the ABACUS 3.2 package based on the numerical atomic orbitals basis set with the Perdew–Burke–Ernzerhof (PBE)<sup>54</sup> exchange-correlation functional generated atomic data sets to reduce the production cost of the data (Section 2.1). We adopted the atomic data sets as training data to generate an attention-based deep potential model using the DPA-1 method, which we named as the DPA-Semi model (Section 2.2). The procedures to generate the DPA-Semi model are shown in Figure 1. For comparison, we also generated DP models for each of the 19 semiconductors mentioned above (Section 2.2). The calculated lattice constants, bulk moduli, shear moduli, and Young's moduli of different semiconductors indicate that the results of the DPA-Semi model are consistent with the values of our DFT calculation. (Section 3.1). Furthermore, the DPA-Semi model was utilized to investigate the liquid and amorphous structures (Section 3.2) and melting temperatures (Section 3.3) of various semi-

conductor materials, and the results are consistent with experiments and other calculations.

## 2. METHODS

**2.1. Density Functional Theory.** All of the DFT calculations were performed with the ABACUS 3.2 package.<sup>49</sup> The generalized gradient approximation (GGA) in the form of the Perdew–Burke–Ernzerhof (PBE)<sup>54</sup> was used for the exchange-correlation functional. The norm-conserving pseudo-potentials<sup>55,56</sup> were employed, and the valence configurations for the elements were [B]2s<sup>2</sup>2p<sup>1</sup>, [C]2s<sup>2</sup>2p<sup>2</sup>, [N]2s<sup>2</sup>2p<sup>3</sup>, [Al]3s<sup>2</sup>3p<sup>1</sup>, [Si]3s<sup>2</sup>3p<sup>2</sup>, [P]3s<sup>2</sup>3p<sup>3</sup>, [S]3s<sup>2</sup>3p<sup>4</sup>, [Zn]-3s<sup>2</sup>3p<sup>6</sup>3d<sup>10</sup>4s<sup>2</sup>, [Ga]4s<sup>2</sup>4p<sup>1</sup>, [Ge]4s<sup>2</sup>4p<sup>2</sup>, [As]4s<sup>2</sup>4p<sup>3</sup>, [Cd]-4s<sup>2</sup>4p<sup>6</sup>4d<sup>10</sup>5s<sup>2</sup>, [In]4d<sup>10</sup>5s<sup>2</sup>5p<sup>1</sup>, [Sb]5s<sup>2</sup>5p<sup>3</sup>, and [Te]4d<sup>10</sup>5s<sup>2</sup>5p<sup>4</sup>, respectively. The energy cutoff was set to 100 Ry. The Monkhorst–Pack<sup>57</sup> *k*-points mesh was set with the spacing being 0.08 Bohr<sup>-1</sup>. We employed the Gaussian smearing method with a smearing width of 0.002 Ry. The electronic iteration convergence threshold was set to 10<sup>-6</sup>.

The triple- $\zeta$  plus double polarization (TZDP) numerical atomic orbitals basis sets were used for all of the DFT calculations. The numerical atomic orbitals were chosen to be 3s3p2d for B, 3s3p2d for C, 3s3p2d for N, 3s3p2d for Al, 3s3p2d for Si, 3s3p2d for P, 3s3p2d for S, 6s3p3d2f for Zn, s3p2d for Ga, 3s3p2d for Ge, 3s3p2d for As, 6s3p3d2f for Cd, 3s3p3d2f for In, 3s3p2d for Sb, and 3s3p3d2f for Te, respectively. The cutoffs of numerical atomic orbitals were all set to 9 a.u.

**2.2. Deep Potential Generation.** In this work, we constructed machine-learning-based models for 19 kinds of semiconductors ranging from group IIB to VIA, namely, Si, Ge, SiC, BAs, BN, AlN, AlP, AlAs, InP, InAs, InSb, GaN, GaP, GaAs, CdTe, InTe, CdSe, ZnS, and CdS. For each semiconductor, we first selected a variety of crystal structures, as detailed in Table 1. Next, random perturbations were performed on the atomic coordinates by adding values drawn from a uniform distribution in the range of [-0.01, 0.01]. We also changed the cell vectors by a symmetric deformation matrix constructed by adding random noise drawn from a uniform distribution in the range of [-0.03, 0.03]. Five steps of AIMD simulations were performed for all the perturbed structures to produce labeled data with energies, forces, and virial tensors calculated from DFT. These labeled data were used to form the initial data sets.

Next, we utilized the Deep Potential Generator (DP-GEN 0.11.1 package)<sup>58</sup> to generate Deep Potential (DP) models<sup>30</sup> for each semiconductor. The initial data were trained by the DeePMD-kit 2.2.8 package.<sup>59</sup> We adopted three hidden layers for the embedding network for the DP models, with sizes of 25, 50, and 100. In addition, three hidden layers with sizes of 240, 240, and 240 were selected for the fitting network. An exponentially decaying learning rate was chosen to change from  $1.0 \times 10^{-3}$  to  $3.5 \times 10^{-8}$ . During the optimization process, the prefactor of the energy (force) term in the loss function changes from 0.02 to 1 (1000 to 1). The DP model was trained for  $4.0 \times 10^5$  steps with the cutoff radius being 8 Å. Four DP models were generated for each training process, where the same reference data set was used, but the initial parameters for the deep neural network were different.

Finally, we performed MD simulations with the DP model with temperatures ranging from 0 K to twice the melting temperatures of each semiconductor (detailed temperature settings are shown in Table 1) and pressures ranging from 0 to 5 GPa to explore new configurations using the LAMMPS (23 Jun 2022) package.<sup>60</sup> We did not include perturbations in the initial

**Table 1. Crystal Structures Adopted for Generating First-principles Data for 19 Semiconductors, As Well As the Ranges of Temperatures Used for Generating the Corresponding Machine-learning-Based Models<sup>a</sup>**

system	temperature (K)	crystal structures
Si	0–3200	cubic ( $F\bar{d}3m$ ), hexagonal ( $P6_3/mmc$ ), hexagonal ( $P6/mmm$ ), tetragonal ( $I4_1/amd$ )
Ge	0–2600	cubic ( $F\bar{d}3m$ ), hexagonal ( $P6_3/mmc$ ), tetragonal ( $I4_1/amd$ )
SiC	0–5950	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mc$ )
BAs	0–4600	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mc$ )
BN	0–6600	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mmc$ )
AlN	0–7200	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mc$ ), hexagonal ( $P6_3/mmc$ )
AlP	0–4600	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mc$ )
AlAs	0–4200	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3/mmc$ )
InP	0–3200	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mc$ )
InAs	0–2600	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mc$ )
InSb	0–2200	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mc$ )
GaN	0–4000	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mc$ )
GaP	0–3500	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mc$ )
GaAs	0–3000	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mc$ )
CdTe	0–2650	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mc$ )
InTe	0–1950	cubic ( $Fm\bar{3}m$ ), cubic ( $Pm\bar{3}m$ )
CdSe	0–3250	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mc$ )
ZnS	0–3950	cubic ( $F\bar{4}3m$ )
CdS	0–4050	cubic ( $F\bar{4}3m$ ), hexagonal ( $P6_3mc$ )

<sup>a</sup>The pressure range is set to 0–5 GPa.

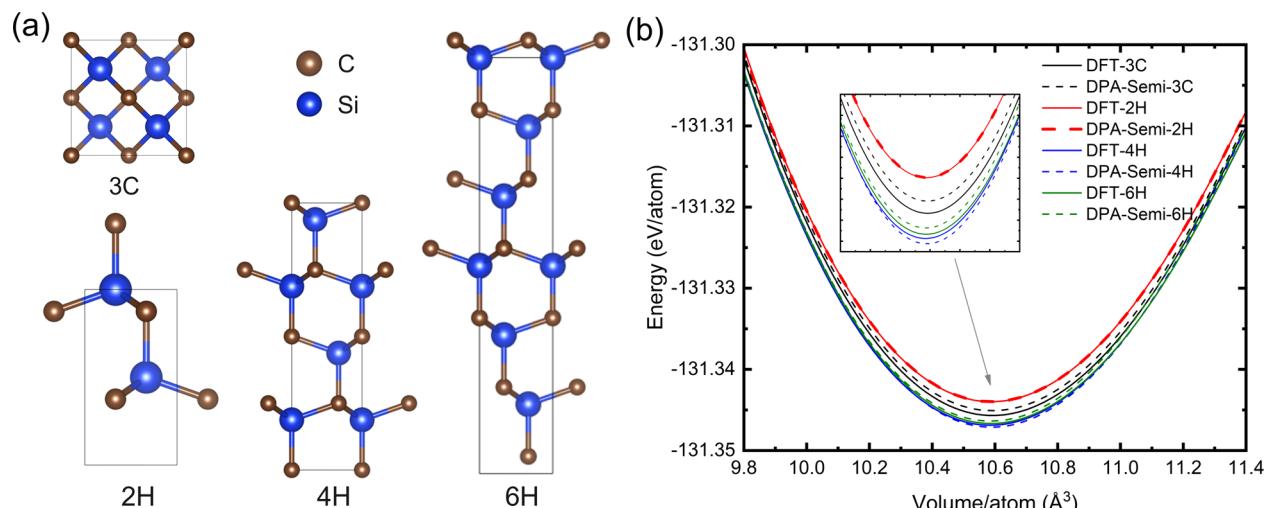
configurations. More details regarding the adopted crystal structures, as well as the selected ranges of temperatures and pressures for exploring the configuration space in each iteration, are shown in Table 1. During each iteration, a maximum of 60 candidate configurations were selected for each semiconductor crystal structure. These configurations were added to the training set for the next iteration after calculating their energies, atomic forces, and virial tensor using the DFT method. All iterations were done automatically with the DP-GEN 0.11.1 package. The iterations continued until the proportion of candidate configurations was less than 5% and remained almost unaltered for another few iterations. After the DP-GEN iterations were converged, we trained the collected data for  $1.2 \times 10^7$  steps with “se\_e2\_a” descriptors (DP model)<sup>59,61</sup> and “se\_atten” descriptors (DPA-Semi model)<sup>43</sup> using DeePMD-kit 2.2.8 package, respectively. The number of data sets for 19 semiconductors is shown in Table 2. In addition, the AlN system has the largest amount of data sets with 29,715 frames. Because its melting temperatures are high, the iterative temperature reaches twice its melting temperatures (Table 1), so the amount of data sets sampled is relatively large. The InSb system has the least amount of data sets, which is 2415 frames. This is caused by the fact that the melting temperatures of InSb are relatively small, indicating that the amount of data sampled is relatively small.

Furthermore, Table 2 shows the root-mean-square errors (RMSEs) of the energies and forces predicted by the DP and DPA-Semi models. Generally, smaller RMSEs of energy and forces imply a relatively higher accuracy of the machine learning models. We find the smallest RMSEs of energy (0.005 eV/atom) and forces (0.11 eV/Å) are found for the Ge system by using the DP model. For the DPA-Semi model, the energy RMSEs of the BAs system are the smallest, which is 0.0032 eV/atom, and the

**Table 2.** Root-Mean-square Errors (RMSEs) of the Total Energy (meV/atom) and Forces (meV/Å) Predicted By Each DP Model and the DPA-Semi Model, and the Number of Data Sets Generated for 19 Semiconductors after the DP-GEN Iterations<sup>a</sup>

	Si	Ge	SiC	BA	BN	AlN	AlP	AlAs	InP	InAs
energy (DP)	5.53	3.90	11.61	4.53	10.49	11.48	9.11	7.92	7.95	6.42
energy (DPA)	6.63	5.42	9.04	3.24	6.11	10.18	8.74	7.81	8.97	7.37
force (DP)	120.27	114.61	370.00	151.34	354.74	324.40	216.74	187.28	191.00	152.87
force (DPA)	181.22	120.06	303.34	142.37	367.91	283.25	176.01	163.56	182.47	139.87
frame	16,923	18,966	12,749	4236	18,045	29,715	6988	5043	8601	7477
	InSb	GaN	GaP	GaAs	CdTe	InTe	CdSe	ZnS	CdS	
energy (DP)	4.29	9.71	8.46	6.70	7.33	6.20	5.88	8.61	6.77	
energy (DPA)	6.12	4.98	6.45	6.53	7.94	7.75	6.94	10.36	7.24	
force (DP)	112.88	263.22	211.88	166.11	145.47	133.63	127.43	182.99	151.63	
force (DPA)	105.53	311.29	192.74	152.66	138.55	132.70	127.12	192.79	145.64	
frame	2415	9005	11,950	7045	7753	7375	11,164	8763	12,589	

<sup>a</sup>Note that the data to calculate RMSEs are from both solid and liquid phases.



**Figure 2.** (a) Crystal structures of the 3C-, 2H-, 4H-, and 6H-SiC polytypes. The blue and gray spheres represent the Si and C atoms, respectively. (b) Curves of energy vs average atomic volume for 3C-, 2H-, 4H-, and 6H-SiC polytypes calculated by the DFT (solid lines) and DPA-Semi (dashed lines) methods.

force RMSEs of the InSb system are the smallest, which is 0.11 eV/Å. On the contrary, for the DP model, the SiC system has the largest RMSEs of energy (0.012 eV/atom) and forces (0.37 eV/Å). For the DPA-Semi model, the largest RMSEs of energy (0.010 eV/atom) and force (0.37 eV/Å) come from the ZnS system and the BN system, respectively. The lattice constant and elastic constants computed by the DP and DPA models agree well with the DFT results. The data sets and the deep-learning potential models (DP and DPA-Semi) are available for download.<sup>62,63</sup> In addition, examples of the DP and DPA-Semi models can be run by Bohrium Notebook.<sup>64</sup>

**2.3. Elastic Constants.** Elastic constants ( $C_{ij}$ ) can be determined by performing a linear least-squares fit between stress and strain for a series of small deformations in the crystal lattice,<sup>65</sup> following Hooke's law:

$$\sigma_{ij} = C_{ijkl} \cdot \epsilon_{kl} \quad (1)$$

where  $\sigma_{ij}$  is the stress tensor,  $\epsilon_{kl}$  is the strain tensor, and  $C_{ijkl}$  is the fourth-rank elastic stiffness tensor, which can be further simplified using the two-indices Voigt notation as a  $6 \times 6$   $C_{ij}$  matrix.

In this study, the zero-temperature elastic constants ( $C_{ij}$ ) are evaluated by the DP-GEN software package,<sup>58</sup> which is based on the Python Materials Genomics (pymatgen) library.<sup>65,66</sup> The

bulk modulus ( $B$ ), shear modulus ( $G$ ), and Young's modulus ( $E$ ) are estimated using the corresponding equations based on the elastic constants.<sup>67,68</sup> For a cubic crystal structure, these properties are computed via the formulas of

$$\begin{aligned} B &= \frac{1}{3}(C_{11} + 2C_{12}), \mathbf{G} \\ &= \frac{1}{5}(C_{11} - C_{12} + 3C_{44}), \mathbf{E} \\ &= \frac{9B}{1 + \frac{3B}{G}} \end{aligned} \quad (2)$$

For a hexagonal crystal structure, the formulas are

$$\begin{aligned} \mathbf{B} &= \frac{1}{9}(2(C_{11} + C_{12}) + 4C_{13} + C_{33}), \\ \mathbf{G} &= \frac{1}{30}(C_{11} + C_{12} + 2C_{33} - 4C_{13} + 12C_{44} + 12C_{66}), \mathbf{E} \\ &= \frac{9B}{1 + \frac{3B}{G}} \end{aligned} \quad (3)$$

**2.4. Mean Square Displacements.** The mean square displacement (MSD) is defined as  $u^2(t)$  according to

**Table 3.** Lattice Constant ( $a$  and  $c$ , in Å), Elastic Constants  $C_{ij}$  (in GPa), Bulk Modulus ( $B$ , in GPa), Shear Modulus ( $G$ , in GPa), and Young's Modulus ( $E$ , in GPa) for the 3C-, 2H-, 4H-, and 6H-SiC Polytypes Structures Calculated by the ABACUS (Using NAO Basis Set with the PBE Functional, Labeled as NAO-PBE) Package, the DP Model, the DPA-Semi Model, the PW-PBE Method (Using PW Basis Set With the PBE Functional), and the PW-LDA Method<sup>a</sup>

methods	$a$	$c$	$C_{11}$	$C_{12}$	$C_{13}$	$C_{33}$	$C_{44}$	$C_{66}$	$B$	$G$	$E$
<b>3C</b>											
NAO-PBE	4.392		379	125			239		210	194	445
DP	4.392		385	121			242		212	195	448
DPA-semi	4.391		373	129			238		210	192	445
PW-PBE <sup>b</sup>	4.366		376	129			246		211	176	310
PW-PBE <sup>c</sup>	4.380		382	128			239				
PW-LDA <sup>d</sup>	4.328		369	138			226		215	181	425
exp.	4.360 <sup>f</sup>		395 <sup>n</sup>	123 <sup>n</sup>			236 <sup>n</sup>		225 <sup>h</sup>	192 <sup>i</sup>	469 <sup>j</sup>
<b>2H</b>											
NAO-PBE	3.100	5.087	490	96	47	529	150	196	210	187	433
DP	3.100	5.086	484	100	50	529	151	191	211	185	430
DPA-semi	3.100	5.082	497	101	55	526	158	197	214	190	440
PW-PBE <sup>e</sup>	3.079	5.053	506	92	46	542	154		213	191	441
PW-PBE <sup>c</sup>	3.088	5.083	490	93	52	533	153				
exp.	3.079 <sup>g</sup>	5.053 <sup>g</sup>									
<b>4H</b>											
NAO-PBE	3.102	10.156	490	93	49	528	156	194	210	189	437
DP	3.102	10.162	490	102	57	551	162	193	218	191	444
DPA-semi	3.103	10.158	483	105	49	529	159	189	211	189	437
PW-PBE <sup>b</sup>	3.084	10.096	497	97	49	529	154	199	213	184	475
PW-PBE <sup>c</sup>	3.090	10.178	498	91	52	535	153				
PW-LDA <sup>d</sup>	3.067	10.068	379	116			242		204	197	448
exp.	3.073 <sup>f</sup>	10.052 <sup>f</sup>	501 <sup>k</sup>	111 <sup>k</sup>	52 <sup>k</sup>	553 <sup>k</sup>	163 <sup>k</sup>		215 <sup>l</sup>	131 <sup>l</sup>	450 <sup>m</sup>
<b>6H</b>											
NAO-PBE	3.103	15.226	491	91	49	528	159	194	210	190	439
DP	3.103	15.227	486	102	54	532	164	191	214	190	440
DPA-semi	3.103	15.227	479	106	49	527	162	190	211	190	436
PW-PBE <sup>b</sup>	3.085	15.138	493	100	53	532	156	196	214	184	469
PW-LDA <sup>d</sup>	3.074	15.100	376	118			238		204	194	442
exp.	3.081 <sup>f</sup>	15.120 <sup>f</sup>	501 <sup>k</sup>	111 <sup>k</sup>	52 <sup>k</sup>	553 <sup>k</sup>	163 <sup>k</sup>		215 <sup>l</sup>	131 <sup>l</sup>	450 <sup>m</sup>

<sup>a</sup>Experimental results are also presented as comparisons. <sup>b</sup>ref 25. <sup>c</sup>ref 8. <sup>d</sup>ref 26. <sup>e</sup>ref 73. <sup>f</sup>ref 74. <sup>g</sup>ref 76. <sup>h</sup>ref 80. <sup>i</sup>ref 81. <sup>j</sup>ref 82. <sup>k</sup>ref 77. <sup>l</sup>ref 83. <sup>m</sup>ref 84. <sup>n</sup>ref 75.

$$u^2(t) = \frac{1}{N} \sum_{i=1}^N [R_i(t) - R_i(0)]^2 \quad (4)$$

where  $R_i(t)$  is the atom position of atom  $i$  after  $t$  time of the simulation.  $N$  is the total number of atoms. MSD provides a method to observe whether the system is in a solid or liquid state. If the system is in a solid state, then MSD oscillates around a constant value. This indicates that all atoms are confined to specific positions. However, if some atoms melt, the MSD linearly increases. Utilizing the DP model and DPA-Semi model, we calculated the melting temperature data of semiconductor materials based on the observation of variations in MSD.

**2.5. Radial Distribution Functions.** The radial distribution function  $g(r)$  describes the local structure of atoms and is defined as

$$g(r) = \frac{V}{4\pi r^2 N^2} \left\langle \sum_{i=1}^N \sum_{j=1, j \neq i}^N \delta(r - |\mathbf{r}_i - \mathbf{r}_j|) \right\rangle \quad (5)$$

where  $V$  is the cell volume,  $N$  is the number of atoms,  $\mathbf{r}_i$  and  $\mathbf{r}_j$  are atomic coordinates of atoms  $i$  and  $j$ , and  $\langle \dots \rangle$  means the time or ensemble average. In this manuscript, we computed the radial distribution function (RDF) for amorphous silicon, liquid silicon, and amorphous InSb systems utilizing the DPA-Semi

model. A comparative analysis with experimental data was undertaken to substantiate the accuracy of the DPA-Semi model.

### 3. RESULTS AND DISCUSSION

#### 3.1. Lattice Constants and Mechanical Properties.

Figure 2a illustrates the stable structures of silicon carbide (SiC), which include the 3C, 2H, 4H, and 6H structures with close energies.<sup>18–26</sup> Therefore, it is an ideal example for testing the accuracy of the DPA-Semi model. In detail, the SiC structures are based on different stacking patterns in the cubic and hexagonal diamond structures. As shown in Figure 2a, the 3C structure is a zinc blende structure with the ABC stacking, while the 2H, 4H, and 6H structures take the wurtzite structure and have the AB, ABCB, and ABCACB stackings.<sup>18,19</sup> In this regard, the 3C-, 2H-, 4H-, and 6H-SiC structures have 8, 4, 8, and 12 atoms per primitive unit cell, respectively. We find the cells shown in Figure 2a become more elongated along the cross-plane direction of  $n$ H-SiC with increasing  $n$ .

Figure 2b shows the computed total energy with respect to the volume per atom for the 3C-, 2H-, 4H-, and 6H-SiC polytypes calculated by the DFT and DPA-Semi methods. We find the calculated total energies satisfy the inequality relation:  $E_{2H} > E_{3C} > E_{6H} > E_{4H}$ . This data indicates that the 4H-SiC structure is

**Table 4.** Lattice Constants ( $a$ ), Bulk Moduli ( $B$ ), Shear Moduli ( $G$ ), and Young's Moduli ( $E$ ) for Semiconductor Structures<sup>a</sup>

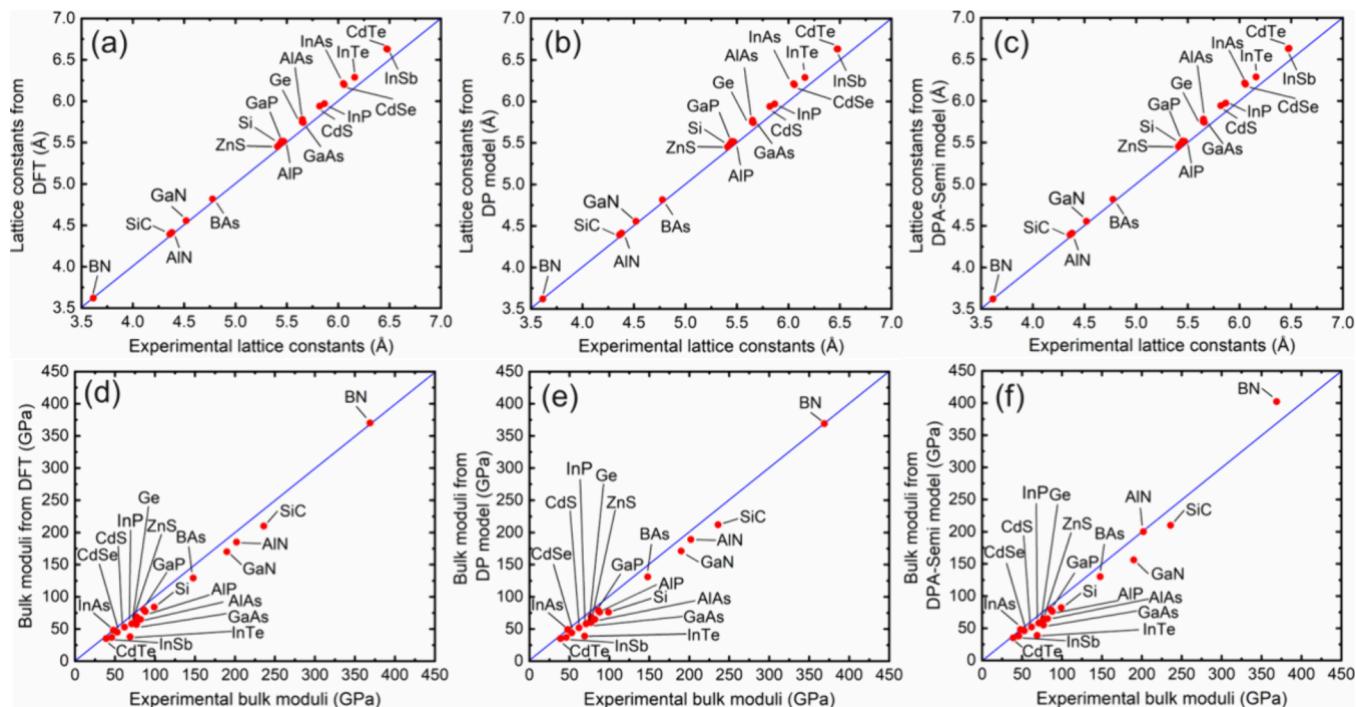
systems	methods	$a$ (Å)	$B$ (GPa)	$G$ (GPa)	$E$ (GPa)	systems	methods	$a$ (Å)	$B$ (GPa)	$G$ (GPa)	$E$ (GPa)						
Si	DFT	5.480	84	63	152		GaN	4.555	170	112	276						
	DP	5.482	76 (-8)	64 (1)	150 (-2)		DP	4.555	171 (1)	113 (1)	278 (2)						
	DPA-semi	5.484	82 (-2)	61 (-2)	145 (-7)		DPA-semi	4.554	156 (-14)	111 (-1)	267 (-9)						
	exp.	5.430 <sup>b</sup>	99 <sup>b</sup>				exp.	4.52 <sup>i</sup>	190 <sup>j</sup>								
Ge	DFT	5.779	57	42	103		GaP	5.517	77	54	132						
	DP	5.774	60 (3)	43 (1)	104 (1)		DP	5.518	76 (-1)	53 (-1)	130 (-2)						
	DPA-semi	5.781	55 (-2)	43 (1)	103 (0)		DPA-semi	5.517	76 (-1)	53 (-1)	128 (-4)						
	exp.	5.652 <sup>c</sup>	77 <sup>d</sup>				exp.	5.450 <sup>l</sup>	88 <sup>m</sup>								
SiC	DFT	4.392	210	194	445		GaAs	5.750	61	43	106						
	DP	4.392	212 (2)	195 (1)	448 (3)		DP	5.752	61 (0)	43 (0)	105 (-1)						
	DPA-semi	4.391	210 (0)	194 (0)	445 (0)		DPA-semi	5.751	61 (0)	39 (-4)	97 (-9)						
	exp.	4.360 <sup>e</sup>	225 <sup>f</sup>				exp.	5.650 <sup>n</sup>	76 <sup>n</sup>								
BAs	DFT	4.819	129	127	288		CdTe	6.628	35	15	40						
	DP	4.816	131 (2)	128 (1)	291 (3)		DP	6.630	35 (0)	15 (0)	39 (-1)						
	DPA-semi	4.819	130 (1)	134 (7)	300 (12)		DPA-semi	6.630	35 (0)	14 (-1)	38 (-2)						
	exp.	4.777 <sup>g</sup>	148 <sup>g</sup>				exp.	6.480 <sup>o</sup>	39 <sup>o</sup>								
BN	DFT	3.621	370	392	870		InTe	6.288	38	19	50						
	DP	3.621	369 (-1)	401 (9)	883 (13)		DP	6.289	39 (1)	19 (0)	49 (-1)						
	DPA-semi	3.621	402 (32)	389 (-3)	882 (12)		DPA-semi	6.288	39 (1)	19 (0)	50 (0)						
	exp.	3.615 <sup>h</sup>	369 <sup>h</sup>				exp.	6.160 <sup>p</sup>	69 <sup>q</sup>								
AlN	DFT	4.413	185	131	318		CdSe	6.213	45	18	47						
	DP	4.413	189 (4)	129 (-2)	317 (-1)		DP	6.213	44 (-1)	16 (-2)	42 (-5)						
	DPA-semi	4.410	200 (15)	129 (-2)	319 (1)		DPA-semi	6.215	46 (1)	17 (-1)	46 (-1)						
	exp.	4.380 <sup>i</sup>	202 <sup>j</sup>				exp.	6.050 <sup>o</sup>	53 <sup>o</sup>								
AlP	DFT	5.514	80	50	124		ZnS	5.453	69	37	95						
	DP	5.514	79 (-1)	50 (0)	124 (0)		DP	5.453	69 (0)	35 (-2)	89 (-6)						
	DPA-semi	5.513	79 (-1)	50 (0)	124 (0)		DPA-semi	5.453	65 (-4)	37 (0)	94 (-1)						
	exp.	5.470 <sup>k</sup>	86 <sup>k</sup>				exp.	5.410 <sup>r</sup>	77 <sup>r</sup>								
AlAs	DFT	5.742	65	42	105		CdS	5.939	53	20	55						
	DP	5.743	65 (0)	41 (-1)	101 (-4)		DP	5.939	52 (-1)	18 (-2)	49 (-6)						
	DPA-semi	5.744	65 (0)	43 (1)	106 (1)		DPA-semi	5.943	52 (-1)	20 (0)	56 (1)						
	exp.	5.660 <sup>k</sup>	82 <sup>k</sup>				exp.	5.820 <sup>o</sup>	62 <sup>o</sup>								
InP	DFT	5.970	58	33	84												
	DP	5.969	58 (0)	31 (-2)	79 (-5)												
	DPA-semi	5.973	58 (0)	33 (0)	85 (1)												
	exp.	5.868 <sup>s</sup>	71 <sup>t</sup>														
InAs	DFT	6.200	48	26	67												
	DP	6.201	49 (1)	27 (1)	68 (1)												
	DPA-semi	6.202	48 (0)	26 (0)	67 (0)												
	exp.	6.058 <sup>u</sup>	48 <sup>v</sup>														
InSb	DFT	6.631	37	20	52												
	DP	6.631	37 (0)	22 (2)	56 (4)												
	DPA-semi	6.628	38 (1)	18 (-2)	47 (-5)												
	exp.	6.473 <sup>w</sup>	46 <sup>x</sup>														

<sup>a</sup>The Si and Ge systems are calculated using the diamond structure. The InTe system adopts the B1 structure. Other semiconductors are calculated using the zinc blende structure. The ABACUS package with numerical atomic orbitals (NAO) is used for DFT calculations with the PBE functional. The values in parentheses are the deviations of the DP and DPA-Semi model's predictions from the reference DFT computed values. We do not include  $G$  and  $E$  values here because they are often not discussed in most of the semiconductor experimental literature. <sup>b</sup>ref 85. <sup>c</sup>ref 86. <sup>d</sup>ref 87. <sup>e</sup>ref 74. <sup>f</sup>ref 80. <sup>g</sup>ref 88. <sup>h</sup>ref 89. <sup>i</sup>ref 90. <sup>j</sup>ref 91. <sup>k</sup>ref 92. <sup>l</sup>ref 93. <sup>m</sup>ref 94. <sup>j</sup>ref 95. <sup>o</sup>ref 96. <sup>p</sup>ref 97. <sup>q</sup>ref 98. <sup>r</sup>ref 99. <sup>s</sup>ref 100. <sup>t</sup>ref 101. <sup>u</sup>ref 102. <sup>v</sup>ref 103. <sup>w</sup>ref 104. <sup>x</sup>ref 105.

more stable than the other three structures, which agrees with previous DFT results.<sup>25,26,69</sup> Although the energy differences between the polytypes are small, on the order of meV/atom, both DFT and DPA-Semi models correctly distinguish these structural energy differences, demonstrating that the DPA-Semi model owns a satisfactory GGA quality accuracy.

Table 3 lists the in- and cross-plane lattice constants ( $a$  and  $c$ ) and elastic constants ( $C_{11}$ ,  $C_{12}$ ,  $C_{13}$ ,  $C_{33}$ ,  $C_{44}$ , and  $C_{66}$ ) for the 3C-, 2H-, 4H- and 6H-SiC structures. The data are obtained by ABACUS<sup>49</sup> with the NAO basis and the PBE functional, the DP models,<sup>30</sup> the DPA-Semi mode,<sup>43</sup> DFT data from refs 70–72, and available experimental data. Our results for the lattice constants and elastic constants of the 3C-, 2H-, 4H-, and 6H-SiC structures agree well with other calculated results<sup>8,25,26,73</sup> and experimental data from refs 74–77.

Taking the 3C-SiC structure as an example. Table 3 lists the elastic constants ( $C_{11}$ ,  $C_{12}$ , and  $C_{44}$ ) of 3C-SiC. We find the  $C_{11}$  values obtained from the DP models, the DPA-Semi model, the ABACUS package, and the experiments are 385, 379, 379, and 395 GPa, respectively. In this regard, the maximum deviation between our calculations and the experimental results is 4.1%, which suggests that the DP models and the DPA-Semi model are sufficiently accurate. On the other hand, we compare the deviation of elastic constants ( $C_{11}$ ,  $C_{12}$ , and  $C_{44}$ ) between DP and DFT calculations. In particular, we observe that the discrepancies for the  $C_{12}$  value of the 3C, 2H, 4H, and 6H-SiC structures between DP and DFT are 3.3, 4.4, 9.4, and 10.8%, respectively. On the other hand, we also compare the discrepancies for the  $C_{12}$  value of the 3C, 2H, 4H, and 6H-SiC structures between DPA-Semi and DFT are 0.8, 3.1, 5.4, and



**Figure 3.** Predicted lattice constants of various semiconductors by (a) DFT calculations with NAO basis set, (b) the DP models, (c) the DPA-Semi model, and available experimental data. Predicted bulk moduli of various semiconductors by (d) DFT calculations with the NAO basis set, (e) the DP models, (f) the DPA-Semi model, and available experimental data. The DFT results are calculated by the ABACUS package with numerical atomic orbitals.

8.8%, respectively. It indicates that the DPA-Semi model predicts slightly closer values of the  $C_{12}$  elastic constants than the DP model when compared to the DFT results.

Table 3 also shows the bulk modulus ( $B$ ), shear modulus ( $G$ ), and Young's modulus ( $E$ ) of the 3C-, 2H-, 4H-, and 6H-SiC structures. As listed in Table 3, the DP and DPA-Semi models reproduce the elastic moduli of the SiC structures with DFT accuracy. For example, the maximum deviation of bulk modulus for 3C-, 2H-, 4H-, and 6H-SiC obtained from the DPA-Semi model and the DFT calculations is 1.9% (2H-SiC). These results demonstrate that the DP and DPA-Semi models are suitable for studying the mechanical properties of SiC polytypes. Besides the results for SiC, Table 4 shows the lattice constants and elastic moduli of 19 semiconductors from group IIB to VIA. We find that the results obtained from the DP and DPA-Semi models are in excellent agreement with the DFT results calculated by ABACUS. These results offer reliable evidence that the DP and DPA-Semi models can be employed to study the physical mechanisms of group IIB to VIA semiconductor systems with GGA quality accuracy. The DP/DPA-Semi results of bulk modulus are in excellent agreement with other DFT results.<sup>78,27,28</sup> Taking GaN, GaP, and GaAs as examples, the maximum deviation of the bulk modulus between DP/DPA-Semi and DFT is 1.1% for GaP. From Table 4, the  $B$  of GaN, GaP, and GaAs is 210, 89, and 76 GPa calculated by the DPA-Semi model, respectively. If compared to the experimental bulk modulus from ref 79., i.e., the  $B$  of GaN, GaP, and GaAs is 210, 89, and 76 GPa, respectively, the calculated values for the bulk modulus which were obtained in our calculations or other DFT results<sup>78</sup> are underestimated due to the PBE functional failure to account for nonlocal correlation effects among electrons in a material.<sup>78</sup> This effect is particularly substantial for materials

with metallic or covalent properties where electron–electron interactions are strong.

To facilitate comparison, Figure 3 shows the comparison between ABACUS, DP, DPA-Semi, and experimental data in terms of lattice constants and bulk modulus. We find that the maximum deviation of lattice constants between the computational methods and experimental data is the CdSe system (26.9%). The maximum deviation of bulk modulus between ABACUS/DP/DPA-Semi and experimental data is the InTe system (44.9%).

**3.2. Melting Temperatures.** Estimating the melting temperatures of semiconductors can help one to study the solid–liquid phase transition and related thermodynamic properties. Each semiconductor has a unique melting temperature, and in order to predict it, we used the ‘heat-until-melts’ method for various semiconductors and the two-phase coexistence method for Si and SiC. The “heat-until-melts” method was known to overestimate the melting temperatures due to the superheating effects.<sup>106,107</sup> However, it would be useful here to yield upper boundaries of the melting temperatures for the semiconductors by using machine-learning-based models. We then performed MD simulations using the DP and DPA-Semi models to predict the melting temperatures and compared them with the experimental data. On the other hand, the two-phase coexistence method yields a more accurate melting temperature but the method is more expensive, and we adopted the DPA-Semi model for these simulations.

The initial structure setups for using the “heat-until-melts” method are shown in Table 5. The number of atoms in these structures is selected to be at least 1000 atoms in order to eliminate the size effects. All of the crystal structures were simulated at an external pressure of 1 bar for 50 ps using the isothermal–isobaric (NPT) ensemble with a time step of 2 fs.

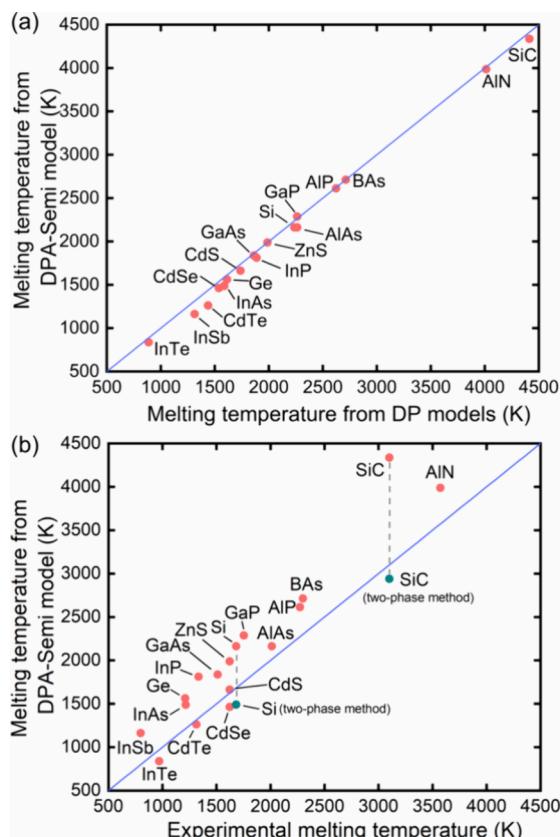
**Table 5. Initial Structures Used for Melting Temperatures Calculations of Semiconductors<sup>a</sup>**

systems	crystal structures	$\alpha$	$\beta$	$\gamma$	number of atoms
Si	cubic ( $F\bar{d}3m$ )	90	90	90	1000
Ge	cubic ( $F\bar{d}3m$ )	90	90	90	1000
SiC	cubic ( $F\bar{4}3m$ )	90	90	90	1000
BA <sub>3</sub>	cubic ( $F\bar{4}3m$ )	90	90	90	1000
AlN	hexagonal ( $P6_3/mmc$ )	90	90	120	1536
AlP	cubic ( $F\bar{4}3m$ )	90	90	90	1000
AlAs	cubic ( $F\bar{4}3m$ )	90	90	90	1000
InP	cubic ( $F\bar{4}3m$ )	90	90	90	1000
InAs	cubic ( $F\bar{4}3m$ )	90	90	90	1000
InSb	cubic ( $F\bar{4}3m$ )	90	90	90	1000
GaP	cubic ( $F\bar{4}3m$ )	90	90	90	1000
GaAs	cubic ( $F\bar{4}3m$ )	90	90	90	1000
CdTe	cubic ( $F\bar{4}3m$ )	90	90	90	1000
InTe	cubic ( $Fm\bar{3}m$ )	90	90	90	1000
CdSe	cubic ( $F\bar{4}3m$ )	90	90	90	1000
ZnS	cubic ( $F\bar{4}3m$ )	90	90	90	1000
CdS	cubic ( $F\bar{4}3m$ )	90	90	90	1000

<sup>a</sup>Labels  $\alpha$ ,  $\beta$ , and  $\gamma$  are the angles ( $^{\circ}$ ) between the crystallographic axes of a crystal, respectively. The number of atoms in the cell is also listed.

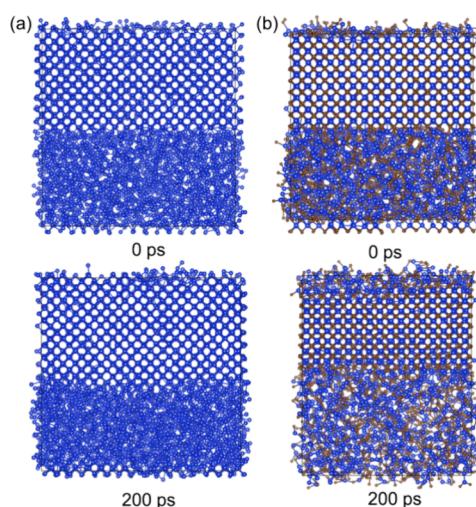
The melting temperature is obtained by changing the simulation temperatures until the liquid phase of the semiconductor is observed by the mean square displacement (MSD). Figure 4a compares the melting temperatures of semiconductors as obtained from the DP and DPA-Semi models. The results indicated that the DPA-Semi model well reproduces the results of DP potentials. Figure 4b further compares the melting temperatures predicted by the DPA-Semi model and those from experimental results.<sup>108</sup> In general, the melting temperatures predicted by the DPA-Semi model are in reasonable agreement with the experimental values, while most predicted melting points are higher than the experimental values. For example, the predicted melting points of Si and SiC (3C) are about 2162 and 4337 K, higher than the experimental melting points of 1687 and 3100 K,<sup>108</sup> respectively.

Next, we utilized the two-phase coexistence method<sup>109</sup> to predict a more accurate melting temperature of Si and SiC (3C) by using the DPA-Semi model. The method predicts the coexistence temperature of solid and liquid phases and can be used to obtain a more accurate melting temperature.<sup>107</sup> We initially set up a 2048-atom solid phase structure and a 2048-atom liquid phase for both Si and 3C-SiC structures. The solid phase was heated to 1000 K, while the liquid phase was heated slightly above the melting temperatures predicted through the above-mentioned “heat-until-melts” method. The simulations were run for 30 ps with the time step being 1 fs to reach equilibrium structures for both phases. We then combined the two phases to construct a 4096-atom cell and fixed the solid-phase atoms. Subsequently, a 30 ps simulation was run using the NVT ensemble at the experimentally determined melting temperature to reduce the interfacial energy between the solid and liquid phases. Finally, we conducted an NPT simulation of the entire system at a pressure of 1 bar. We incrementally increased the temperature with an interval of 5 K and performed simulations of 200 ps to determine the coexistence temperature of the solid and liquid phases. This coexistence temperature served as the melting temperatures obtained through the two-phase method.

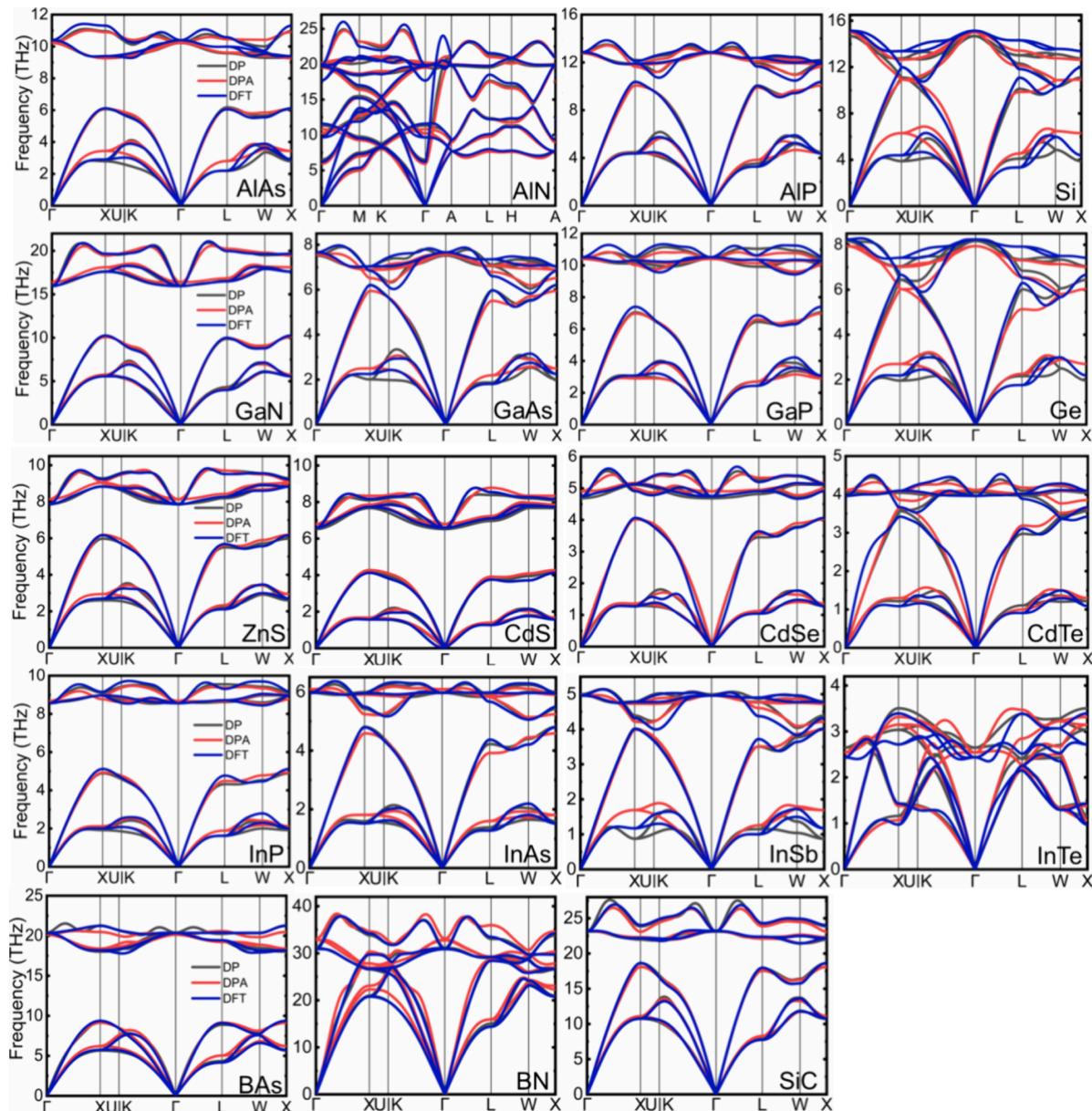


**Figure 4.** (a) Predicted melting temperatures of various semiconductors by the DPA-Semi and DP models. (b) Predicted melting temperatures of various semiconductors by the DPA-Semi and available experimental data.<sup>108</sup> In addition, the melting points of Si (diamond structure) and SiC (3C structure) obtained from the two-phase coexistence method are shown for comparison (green dots).

Figure 5a,b shows the initial structure and a structure of 200 ps for the solid–liquid coexistence DPMD simulations of Si and



**Figure 5.** Two-phase coexistence method with the DPA-Semi model to predict the melting points of Si and SiC (3C) structures. (a) At 1485 K, the Si system with 4096 atoms (blue) exhibit coexistence of solid and liquid phases during a 200 ps DPMD simulations. (b) At 2925 K, the 3C-SiC system with 4096 atoms (Si: blue, C: brown) show solid–liquid coexistence during a 200 ps DPMD.



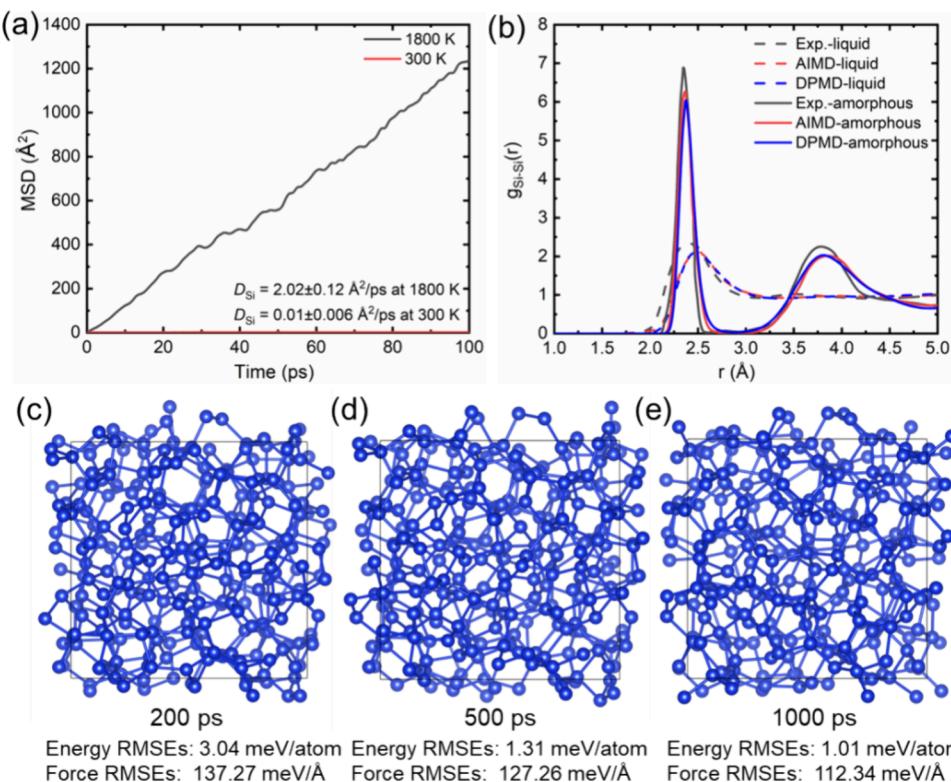
**Figure 6.** Phonon spectra of 19 semiconductors calculated by the DFT, DPA-Semi, and DP methods. The Si and Ge systems adopt the diamond structure while the AlN system adopts the B1 structure; other semiconductors are calculated using the zinc blende structure. The special  $k$ -points are  $\Gamma(0,0,0)$ ,  $M(0.5,0,0)$ ,  $K(1/3,1/3,0)$ ,  $A(0,0,0.5)$ ,  $L(0.5,0,0.5)$ , and  $H(1/3,1/3,0.5)$  for the B1 structure, and  $\Gamma(0,0,0)$ ,  $X(0.5,0,0.5)$ ,  $U(0.625,0.25,0.625)$ ,  $K(0.375,0.375,0.75)$ ,  $L(0.5,0.5,0.5)$ , and  $W(0.5,0.25,0.75)$  for the diamond and zinc blende structures.

SiC (3C), respectively. In the case of Si, the DPA-Semi model yields a coexistence temperature close to 1485 K, which is also shown in Figure 4b. When compared to the experimental data of 1687 K,<sup>108</sup> the coexistence method yields a better melting temperature than the “heat-until-melts” method. This result is also consistent with 1485 K from a recent DPMD simulation using the two-phase coexistence method.<sup>110</sup> On the other hand, the DPA-Semi model for 3C-SiC yields a solid–liquid coexistence temperature of ~2925 K, as shown in Figure 4b. Compared with the overheating effect of the “heat-until-melts” method, the data is closer to the experimental melting temperature of 3100 K.<sup>108</sup>

Note that we do not include the melting temperatures of GaN and BN in the melting temperature simulations because the experimental results for the two structures are still inconclusive. On the one hand, the phase diagram of GaN has not yet been

well understood through research because significant phase transformations occur.<sup>111</sup> For instance, the liquid GaN may coexist with N<sub>2</sub> molecules, which poses challenges to the model. Currently, both DP and DP-Semi models do not include training sets of N<sub>2</sub> molecules. However, it is possible to fine-tune the DP-Semi model with additional data of N<sub>2</sub> molecules for downstream tasks but beyond the scope of current work.

On the other hand, measuring the melting point of BN remains a significant challenge due to the difficulties associated with heating optically transparent diamonds through the absorption of intense laser radiation in a diamond anvil cell, as well as the practical obstacles in attaining high temperatures within a large-volume press using conventional resistance heating techniques.<sup>112</sup> Besides, BN easily sublimes at its melting temperature.<sup>113</sup> Due to the challenges in simulating the liquid



**Figure 7.** (a) Mean square displacements (MSD) and diffusion coefficients ( $D$ ) of crystalline Si at 300 K (red line) and liquid Si at 1800 K (black line) from the DPMD trajectories with the DPA-Semi model. (b) Radial distribution functions  $g(r)$  of Si–Si for liquid Si at 1800 K and amorphous Si at 300 K, where the black dashed line indicates the experiment results,<sup>116,117</sup> the red dashed line depicts AIMD (PBE) results,<sup>38,118</sup> and the blue dashed line represents the DPMD results with the DPA-Semi model. (c), (d), and (e) show the snapshots of amorphous Si at 300 K when the simulation time is 200, 500, and 1000 ps, respectively. RMSEs of the total energy (meV/atom) and forces (meV/\text{\AA}) for the snapshots are also shown.

phases of GaN and BN, we do not include the predicted melting temperatures of GaN and BN in this work.

**3.3. Phonon Spectra.** The phonon spectrum provides crucial insights into the vibrational properties of a material, shedding light on its thermal, electrical, and mechanical behavior. Furthermore, the phonon band structure offers valuable information on the dispersion relation of phonon modes. In this regard, Figure 6 shows the computed phonon spectra of 19 semiconductors utilizing the DP, DPA-Semi, and DFT calculations based on the Phonopy 2.20.0 package.<sup>114,115</sup> For both DP and DPA-Semi calculations, a  $5 \times 5 \times 5$  supercell was employed to ensure convergence, while for DFT-based computations, we utilized a  $4 \times 4 \times 4$  supercell structure that is large enough to converge the phonon spectra. We find the phonon spectra obtained from the three methods for AlAs, AlN, AlP, GaN, GaAs, GaP, ZnS, CdS, CdSe, CdTe, InP, InAs, Bas, BN, and SiC agree well. However, some deviations still exist for the Si, Ge, InSb, and InTe structures. In general, we consider that the DPA-Semi model is able to yield satisfactory results for the phonon spectra of semiconductors.

**3.4. Liquid and Amorphous Structures.** The DPA-Semi model was adopted to predict properties of various phases of Si.<sup>38,119–122</sup>

We first prepared a crystalline Si with a  $3 \times 3 \times 3$  supercell and 216 Si atoms arranged in a diamond lattice configuration. The system was first heated from 300 to 2500 K for 1000 ps to yield a liquid structure, then simulated at 2500 K for another 1000 ps to ensure an equilibration state within the isothermal–isobaric (NPT) ensemble and a time step of 1.0 fs. A controlled cooling period was subsequently initiated, gradually lowering the

temperature to 1800 K throughout at least 500 ps to achieve the liquid phase and to 300 K over a minimum of 1500 ps to obtain the amorphous phase. Equilibration runs were performed in the NPT ensemble at 0 bar and 1800 K/300 K for at least 1000 ps to obtain properties of the liquid/amorphous phases of Si.

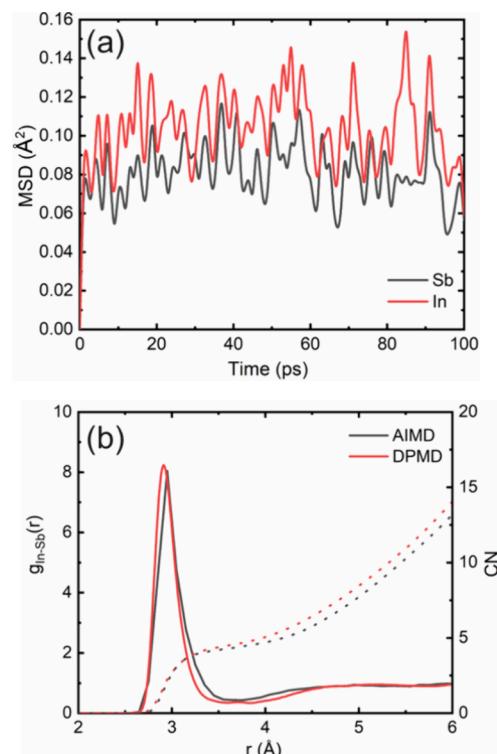
The mean square displacements (MSD) of liquid and crystalline phases of Si are shown in Figure 6a. Here, we conducted a 1000 ps NPT simulation on crystalline Si at 300 K. By using the formula  $\frac{1}{6} \lim_{\Delta t \rightarrow \infty} \frac{\text{MSD}(\Delta t)}{\Delta t}$ , the diffusion coefficient ( $D$ ) of crystalline Si at 300 K and liquid Si at 1800 K is computed to be  $0.01 \pm 0.006$  and  $2.02 \pm 0.12 \text{ \AA}^2/\text{ps}$ .

Figure 7b compares the radial distribution functions  $g(r)$  of Si–Si for liquid Si at 1800 K and amorphous Si at 300 K. We find that the first peak position of the  $g_{\text{Si-Si}}(r)$  of DPMD with the DPA-Semi model for liquid phase or amorphous phase agrees well with previous AIMD results with the PBE functional.<sup>38,118</sup> To further evaluate the accuracy of the DPA-Semi model for amorphous Si, Figure 7c–e shows a few snapshots from the amorphous Si trajectory, and we compared the energy and force obtained from the DPA-Semi model to those obtained from the DFT method. The deviations in energy and force are relatively small, demonstrating the DPA-Semi owns an accuracy similar to the DFT method in describing the amorphous structures of Si. However, it is worth noting that simulation results exhibit some deviations when compared to the experimental data,<sup>116,117</sup> as the peak positions from simulations are slightly larger for the liquid phase of Si,<sup>116</sup> while the peak heights from simulations are slightly lower than the experimental value for the amorphous phase of Si.<sup>117</sup> The above deviations have been attributed to the

inaccuracy of PBE in simultaneously capturing covalent and metallic bonding in Si.<sup>118,78</sup>

We also prepared an amorphous phase of InSb at a density of 6.09 g/cm<sup>3</sup> containing 64/216 atoms, which used the experimental density (6.1 g/cm<sup>3</sup>) of amorphous InSb measured by X-ray reflectivity.<sup>123</sup> The method to generate the amorphous structure of InSb is similar to the above-mentioned amorphous Si. Concurrently, a 10 ps (64 atoms) and 1000 ps (216 atoms) DPMD trajectory were performed with the DFT method and the DPA-Semi model. The canonical ensemble (*NVT*) was selected with a time step of 1.0 fs at a temperature of 300 K.

Similarly, we also calculated the diffusion coefficients (*D*) of In and Sb atoms. The MSD is shown in Figure 8a, and the *D* is



**Figure 8.** (a) Mean square displacements (MSD) for In and Sb atom at 300 K based on the DPMD simulations with the DPA-Semi model. (b) Radial distribution functions  $g_{\text{In-Sb}}(r)$  (solid lines) and coordination numbers (CN, dashed lines) for amorphous InSb at 300 K. The black line indicates our AIMD results (64 atoms) using the PBE functional, and the red line denotes the DPMD with the DPA-Semi model results (216 atoms).

close to zero for In and Sb atoms, indicating that the In and Sb atoms are in an amorphous phase structure. Radial distribution functions and the coordination number (CN) for the two methods are compared in Figure 8b. We observed that the CN of In–Sb at 300 K is about 4.9. Although the  $g_{\text{In-Sb}}(r)$  of DPMD with the DPA-Semi model agrees well with that of AIMD, those are also shifted to slightly smaller distances due to DPA-Semi model bias. The positions of the first peak of  $g_{\text{In-Sb}}(r)$  are at 2.96 Å from AIMD and 2.91 Å from DPMD with the DPA-Semi model, respectively. Those results are slightly larger than the experimental value (2.82 Å) obtained from X-ray measurement data.<sup>123</sup> We utilized the DPA-Semi model to assess the RMSEs of energy and force for various configurations in AIMD, revealing a non-negligible discrepancy (the energy and force RMSEs are 58.3 meV/atom and 201.8 meV/Å, respectively).

This may explain the deviation between the RDF peak values obtained from DPA-Semi model calculations and those from DFT calculations.

## 4. CONCLUSIONS

In this work, we generated an attention-based deep potential model (i.e., DPA-Semi model) and a series of DP models that allow for large-scale pretraining on atomistic data sets to study the various properties of 19 semiconductors ranging from group IIB to VIA, namely, Si, Ge, SiC, BAs, BN, AlN, AlP, AlAs, InP, InAs, InSb, GaN, GaP, GaAs, CdTe, InTe, CdSe, ZnS, CdS. Importantly, the DPA-Semi model, along with data sets, exhibits a high degree of generalization capability for downstream tasks. Researchers can adopt our semiconductor data sets by augmenting them with a small amount of new data to effectively simulate defect and dopant-incorporated semiconductors using the DPA method, thereby reducing the amount of data required in training the traditional DP model.

We focused on comparing the DP and DPA-Semi model results. For example, we calculated the lattice constants and elastic moduli for 19 semiconductors ranging from group IIB to VIA. Our results indicated that the DPA-Semi model can well reproduce the results of each DP model and are in excellent agreement with the DFT results calculated by the ABACUS package. Besides, we also adopted the “heat-until-melts” method and the “two-phase coexistence” method to predict the melting temperatures of semiconductors using the DP or DPA-Semi models. The results indicated that the DPA-Semi model can accurately reproduce the results of each DP model. The melting temperatures predicted by the DPA-Semi model are close to the experimental values.

We took SiC as an example and found the energy differences of the 3C-, 2H-, 4H-, and 6H-SiC polytypes calculated by DFT and DPA-Semi model were small, on the order of meV/atom. Besides, the total energies per atom satisfied the inequality relation:  $E_{2\text{H}} > E_{3\text{C}} > E_{6\text{H}} > E_{4\text{H}}$ , which agreed with the previous theoretical results. It indicated that our DFT calculations based on ABACUS with the LCAO method and DPA-Semi calculation results could correctly distinguish these structural energy differences. Besides, we also compared the lattice constants and elastic constants of the 3C-, 2H-, 4H-, and 6H-SiC structures calculated by ABACUS, DP model, and DPA-Semi model. The results also agreed well with the results from other packages (CASTEP, Quantum Espresso, ABINIT<sup>72</sup>) and experimental data, demonstrating that the DP and DPA-Semi models were adequate for studying the mechanical properties of SiC polytypes. In addition, we found that the radial distribution functions of Si–Si for liquid Si and radial distribution functions of In–Sb for the amorphous phase of InSb calculated by DPMD with the DPA-Semi model were in reasonable agreement with those from AIMD.

In conclusion, our work provided reliable evidence that the DPA-Semi model can be readily employed to study the scientific issues of group IIB to VIA semiconductor systems with GGA quality accuracy. In the future, the DPA-Semi model can be potentially useful for various downstream tasks and substantially reduce computational costs.

## ■ AUTHOR INFORMATION

### Corresponding Author

Mohan Chen – College of Engineering, Peking University,  
Beijing 100871, P. R. China; AI for Science Institute, Beijing

100080, P. R. China; [orcid.org/0000-0002-8071-5633](https://orcid.org/0000-0002-8071-5633); Email: [mohanchen@pku.edu.cn](mailto:mohanchen@pku.edu.cn)

## Authors

Jianchuan Liu – School of Electrical Engineering and Electronic Information, Xihua University, Chengdu 610039, P. R. China; [orcid.org/0000-0001-9507-9732](https://orcid.org/0000-0001-9507-9732)

Xingchen Zhang – College of Engineering, Peking University, Beijing 100871, P. R. China

Tao Chen – HEDPS, CAPT, School of Physics, Peking University, Beijing 100871, P. R. China

Yuzhi Zhang – DP Technology, Beijing 100080, P. R. China

Duo Zhang – AI for Science Institute, Beijing 100080, P. R. China; DP Technology, Beijing 100080, P. R. China; Center for Machine Learning Research, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, P. R. China

Linfeng Zhang – DP Technology, Beijing 100080, P. R. China; AI for Science Institute, Beijing 100080, P. R. China; [orcid.org/0000-0002-8470-5846](https://orcid.org/0000-0002-8470-5846)

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.jctc.3c01320>

## Author Contributions

▀ J.L. and X.Z. contributed equally to this work.

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

This work is supported by the National Science Foundation of China under Grant No. 12122401 and No.12074007. The numerical simulations were performed on the high-performance computing platform of CAPT and the “Bohrium” cloud computing platform of DP Technology Co., LTD.

## REFERENCES

- (1) Sze, S. M. *Semiconductor Devices: Physics and Technology*; John Wiley & Sons, 2008.
- (2) Green, M. A.; Bremner, S. P. J. N. M. Energy Conversion Approaches and Materials for High-Efficiency Photovoltaics. *Nat. Mater.* **2017**, *16*, 23–34.
- (3) Polman, A.; Knight, M.; Garnett, E.; Ehrler, B.; Sinke, W. Photovoltaic Materials: Present Efficiencies and Future Challenges. *Science* **2016**, *352*, No. aad4424.
- (4) Green, M.; Dunlop, E.; Hohl-Ebinger, J.; Yoshita, M.; Kopidakis, N.; Hao, X. J. P. I. P. R. Solar Cell Efficiency Tables (Version 57). *Progress in photovoltaics: research and applications* **2021**, *29*, 3–15.
- (5) Wu, R.; Zhou, K.; Yue, C. Y.; Wei, J.; Pan, Y. J. P. I. M. S. Recent Progress in Synthesis, Properties and Potential Applications of SiC Nanomaterials. *Prog. Mater. Sci.* **2015**, *72*, 1–60.
- (6) Wang, J.; Zhou, Y.; Wang, Z.; Rasmita, A.; Yang, J.; Li, X.; von Bardeleben, H. J.; Gao, W. J. N. C. Bright Room Temperature Single Photon Source at Telecom Range in Cubic Silicon Carbide. *Nat. Commun.* **2018**, *9*, 4106.
- (7) Lukin, D. M.; Dory, C.; Guidry, M. A.; Yang, K. Y.; Mishra, S. D.; Trivedi, R.; Radulaski, M.; Sun, S.; Vercruyse, D.; Ahn, G. H.; Vučković, J. 4h-Silicon-Carbide-on-Insulator for Integrated Quantum and Nonlinear Photonics. *Nat. Photonics* **2020**, *14*, 330–334.
- (8) Pizzagalli, L. Stability and Mobility of Screw Dislocations in 4h, 2h and 3c Silicon Carbide. *Acta Mater.* **2014**, *78*, 236–244.
- (9) Stukel, D. Electronic Structure and Optical Spectrum of Boron Arsenide. *Phys. Rev. B* **1970**, *1*, 3458.
- (10) Kang, J. S.; Li, M.; Wu, H.; Nguyen, H.; Hu, Y. J. S. Experimental Observation of High Thermal Conductivity in Boron Arsenide. *Science* **2018**, *361*, 575–578.
- (11) Tian, F.; Song, B.; Chen, X.; Ravichandran, N. K.; Lv, Y.; Chen, K.; Sullivan, S.; Kim, J.; Zhou, Y.; Liu, T.-H. Unusual High Thermal Conductivity in Boron Arsenide Bulk Crystals. *Science* **2018**, *361*, 582–585.
- (12) Li, S.; Zheng, Q.; Lv, Y.; Liu, X.; Wang, X.; Huang, P. Y.; Cahill, D. G.; Lv, B. High Thermal Conductivity in Cubic Boron Arsenide Crystals. *Science* **2018**, *361*, 579–581.
- (13) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864.
- (14) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133.
- (15) Ōba, F.; Kumagai, Y. Design and Exploration of Semiconductors from First Principles: A Review of Recent Advances. *Appl. Phys. Express* **2018**, *11*, No. 060101.
- (16) Sato, K.; et al. First-Principles Theory of Dilute Magnetic Semiconductors. *Rev. Mod. Phys.* **2010**, *82*, 1633–1690.
- (17) Van de Walle, C. G.; Neugebauer, J. First-Principles Calculations for Defects and Impurities: Applications to III-Nitrides. *J. Appl. Phys.* **2004**, *95*, 3851–3879.
- (18) Kackell, P.; Wenzien, B.; Bechstedt, F. Influence of Atomic Relaxations on the Structural Properties of SiC Polytypes from Ab Initio Calculations. *Phys. Rev. B Condens Matter.* **1994**, *50*, 17037–17046.
- (19) Mujica, A.; Rubio, A.; Munoz, A.; Needs, R. High-Pressure Phases of Group-IV, III–V, and II–VI Compounds. *Rev. Mod. Phys.* **2003**, *75*, 863.
- (20) Chen, W.; Li, L.-S. The Study of the Optical Phonon Frequency of 3c-SiC by Molecular Dynamics Simulations with Deep Neural Network Potential. *J. Appl. Phys.* **2021**, *129*, 244104 DOI: [10.1063/5.0049464](https://doi.org/10.1063/5.0049464).
- (21) Choyke, W. J.; Pensl, G. Physical Properties of SiC. *MRS Bull.* **1997**, *22*, 25–29.
- (22) Yang, B.; Deng, Q.; Su, Y.; Peng, X.; Huang, C.; Lee, A.; Hu, N. The Effects of Atomic Arrangements on Mechanical Properties of 2h, 3c, 4h and 6h-SiC. *Comput. Mater. Sci.* **2022**, *203*, No. 111114.
- (23) Stockmeier, M.; Müller, R.; Sakwe, S. A.; Wellmann, P. J.; Magerl, A. On the Lattice Parameters of Silicon Carbide. *J. Appl. Phys.* **2009**, *105*, No. 033511.
- (24) Sinelnik, A. V.; Semenov, A. V. Theoretical Study of the Band Structure of 2h-SiC and 4h-SiC of Silicon Carbide Polytypes. *Condens. Matt. Phys.* **2021**, *24*, 23706.
- (25) Peivaste, I.; Alahyarizadeh, G.; Minuchehr, A.; Aghaie, M. Comparative Study on Mechanical Properties of Three Different SiC Polytypes (3c, 4h and 6h) under High Pressure: First-Principle Calculations. *Vacuum* **2018**, *154*, 37–43.
- (26) Konstantinova, E.; Bell, M. J. V.; Anjos, V. Ab Initio Calculations of Some Electronic and Elastic Properties for SiC Polytypes. *Intermetallics* **2008**, *16*, 1040–1042.
- (27) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169.
- (28) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59*, 1758.
- (29) Lu, X.; Yang, P.; Luo, J.; Guo, X.; Ren, J.; La, P. Density Functional Theory Capture of Electronic Structures and Optical Properties of Vacancy Doped 3c-SiC Systems. *Mater. Res. Express* **2019**, *6*, No. 115905.
- (30) Zhang, L.; Han, J.; Wang, H.; Car, R.; E, W. Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics. *Phys. Rev. Lett.* **2018**, *120*, No. 143001.
- (31) Wen, T.; Zhang, L.; Wang, H.; E, W.; Srolovitz, D. J. Deep Potentials for Materials Science. *Mater. Futures* **2022**, *1*, No. 022601.
- (32) Dai, F.-Z.; Wen, B.; Sun, Y.; Xiang, H.; Zhou, Y. Theoretical Prediction on Thermal and Mechanical Properties of High Entropy ( $Zr_{0.2}Hf_{0.2}Ti_{0.2}Nb_{0.2}Ta_{0.2}$ ) C by Deep Learning Potential. *J. Mater. Sci. Technol.* **2020**, *43*, 168–174.
- (33) Liang, W.; Lu, G.; Yu, J. Molecular Dynamics Simulations of Molten Magnesium Chloride Using Machine-Learning-Based Deep Potential. *Adv. Theor. Simul.* **2020**, *3*, No. 2000180.

- (34) Huang, J.; Zhang, L.; Wang, H.; Zhao, J.; Cheng, J.; E, W. Deep Potential Generation Scheme and Simulation Protocol for the  $\text{Li}_{10}\text{GeP}_2\text{S}_{12}$ -Type Superionic Conductors. *J. Chem. Phys.* **2021**, *154*, No. 094703.
- (35) Wen, T.; Wang, R.; Zhu, L.; Zhang, L.; Wang, H.; Srolovitz, D. J.; Wu, Z. Specialising Neural Network Potentials for Accurate Properties and Application to the Mechanical Response of Titanium. *npj Comput. Mater.* **2021**, *7*, 206.
- (36) Wang, X.; Wang, Y.; Zhang, L.; Dai, F.; Wang, H. A Tungsten Deep Neural-Network Potential for Simulating Mechanical Property Degradation under Fusion Service Environment. *Nucl. Fusion* **2022**, *62*, No. 126013.
- (37) Lu, D.; Jiang, W.; Chen, Y.; Zhang, L.; Jia, W.; Wang, H.; Chen, M. Dp Compress: A Model Compression Scheme for Generating Efficient Deep Potential Models. *J. Chem. Theory Comput.* **2022**, *18*, 5559–5567.
- (38) Li, R.; Lee, E.; Luo, T. A Unified Deep Neural Network Potential Capable of Predicting Thermal Conductivity of Silicon in Different Phases. *Mater. Today Phys.* **2020**, *12*, No. 100181.
- (39) Li, R.; Liu, Z.; Rohskopf, A.; Gordiz, K.; Henry, A.; Lee, E.; Luo, T. A Deep Neural Network Interatomic Potential for Studying Thermal Conductivity of  $\text{B}-\text{Ga}_2\text{O}_3$ . *Appl. Phys. Lett.* **2020**, *117*, 152102.
- (40) Li, Z.; Duan, X.; Liu, L.; Yang, J. Y. Temperature-Dependent Microwave Dielectric Permittivity of Gallium Oxide: A Deep Potential Molecular Dynamics Study. *J. Appl. Phys.* **2023**, *133*, 224103.
- (41) Fu, B.; Sun, Y.; Zhang, L.; Wang, H.; Xu, B. Deep Learning Inter-Atomic Potential for Thermal and Phonon Behaviour of Silicon Carbide with Quantum Accuracy. *arXiv preprint arXiv:10843* **2021**, DOI: [10.48550/arXiv.2110.10843](https://doi.org/10.48550/arXiv.2110.10843).
- (42) Jin, S.; Zhang, Z.; Guo, Y.; Chen, J.; Nomura, M.; Volz, S. Optimization of Interfacial Thermal Transport in Si/Ge Heterostructure Driven by Machine Learning. *Int. J. Heat Mass Transfer* **2022**, *182*, No. 122014.
- (43) Zhang, D.; Bi, H.; Dai, F.-Z.; Jiang, W.; Zhang, L.; Wang, H. Dpa-1: Pretraining of Attention-Based Deep Potential Model for Molecular Simulation. *arXiv preprint arXiv:2208.08236* **2022**, DOI: [10.48550/arXiv.2208.08236](https://doi.org/10.48550/arXiv.2208.08236).
- (44) Chen, M.; Guo, G.; He, L. Systematically Improvable Optimized Atomic Basis Sets for Ab Initio Calculations. *J. Phys.: Condens. Matter* **2010**, *22*, No. 445501.
- (45) Chen, M.; Guo, G.; He, L. Electronic Structure Interpolation Via Atomic Orbitals. *J. Phys.: Condens. Matter* **2011**, *23*, No. 325501.
- (46) Junquera, J.; Paz, Ó.; Sánchez-Portal, D.; Artacho, E. Numerical Atomic Orbitals for Linear-Scaling Calculations. *Phys. Rev. B* **2001**, *64*, No. 235111.
- (47) Ozaki, T. Variationally Optimized Atomic Orbitals for Large-Scale Electronic Structures. *Phys. Rev. B* **2003**, *67*, No. 155108.
- (48) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. Ab Initio Molecular Simulations with Numeric Atom-Centered Orbitals. *Comput. Phys. Commun.* **2009**, *180*, 2175–2196.
- (49) Li, P.; Liu, X.; Chen, M.; Lin, P.; Ren, X.; Lin, L.; Yang, C.; He, L. Large-Scale Ab Initio Simulations Based on Systematically Improvable Atomic Basis. *Comput. Mater. Sci.* **2016**, *112*, 503–517.
- (50) Liu, Y.; Ding, X.; Chen, M.; Xu, S. A Caveat of the Charge-Extrapolation Scheme for Modeling Electrochemical Reactions on Semiconductor Surfaces: An Issue Induced by a Discontinuous Fermi Level Change. *Phys. Chem. Chem. Phys.* **2022**, *24*, 15511–15521.
- (51) Chen, D.; Liu, Y.; Zheng, Y.; Zhuang, H.; Chen, M.; Jiao, Y. Disordered Hyperuniform Quasi-One-Dimensional Materials. *Phys. Rev. B* **2022**, *106*, No. 235427.
- (52) Zheng, D.; Shen, Z.-X.; Chen, M.; Ren, X.; He, L. Retention and Recycling of Deuterium in Liquid Lithium-Tin Slab Studied by First-Principles Molecular Dynamics. *J. Nucl. Mater.* **2021**, *543*, No. 152542.
- (53) Zheng, Y.; Liu, L.; Nan, H.; Shen, Z. X.; Zhang, G.; Chen, D.; He, L.; Xu, W.; Chen, M.; Jiao, Y.; Zhuang, H. Disordered Hyperuniformity in Two-Dimensional Amorphous Silica. *Sci. Adv.* **2020**, *6*, No. eaba0826.
- (54) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- (55) Hamann, D. Optimized Norm-Conserving Vanderbilt Pseudopotentials. *Phys. Rev. B* **2013**, *88*, No. 085117.
- (56) Schlipf, M.; Gygi, F. Optimization Algorithm for the Generation of Oncv Pseudopotentials. *Comput. Phys. Commun.* **2015**, *196*, 36–44.
- (57) Monkhorst, H. J.; Pack, J. D. Special Points for Brillouin-Zone Integrations. *Phys. Rev. B* **1976**, *13*, 5188.
- (58) Zhang, Y.; Wang, H.; Chen, W.; Zeng, J.; Zhang, L.; Wang, H.; E, W. Dp-Gen: A Concurrent Learning Platform for the Generation of Reliable Deep Learning Based Potential Energy Models. *Comput. Phys. Commun.* **2020**, *253*, No. 107206.
- (59) Wang, H.; Zhang, L.; Han, J.; E, W. Deepmd-Kit: A Deep Learning Package for Many-Body Potential Energy Representation and Molecular Dynamics. *Comput. Phys. Commun.* **2018**, *228*, 178–184.
- (60) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19.
- (61) Zhang, L.; Han, J.; Wang, H.; Saidi, W.; Car, R. End-to-End Symmetry Preserving Inter-Atomic Potential Energy Model for Finite and Extended Systems. *Adv. Neural Inf. Process. Syst.* **2018**, *31*.
- (62) DPA-Semi Model Web Site: <https://www.aissquare.com/models/detail?pageType=models&name=DPA-Semi&id=153> (accessed April 27, 2024).
- (63) DPA-Semi Data Set Web Site: <https://www.aissquare.com/datasets/detail?pageType=datasets&name=Semiconductor&id=151>. (Accessed on April 27, 2024).
- (64) Examples of the DPA-Semi Models Web Site: <https://nb.bohrium.dp.tech/detail/2412744832>. (Accessed on April 27, 2024).
- (65) Le Page, Y.; Saxe, P. Symmetry-General Least-Squares Extraction of Elastic Data for Strained Materials from Ab Initio Calculations of Stress. *J. Phys. Rev. B* **2002**, *65*, No. 104104.
- (66) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (Pymatgen): A Robust, Open-Source Python Library for Materials Analysis. *J. Comp. Mater. Sci.* **2013**, *68*, 314–319.
- (67) Voigt, W., Lehrbuch Der Kristallphysik (Teubner, Leipzig, 1928); Macmillan New York, 1928.
- (68) Mayer, B.; Anton, H.; Bott, E.; Methfessel, M.; Sticht, J.; Harris, J.; Schmidt, P. Ab-Initio Calculation of the Elastic Constants and Thermal Expansion Coefficients of Laves Phases. *Intermetallics* **2003**, *11*, 23–32.
- (69) Park, C. H.; Cheong, B. H.; Lee, K. H.; Chang, K. J. Structural and Electronic Properties of Cubic, 2h, 4h, and 6h SiC. *Phys. Rev. B Condens Matter* **1994**, *49*, 4485–4493.
- (70) Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. I.; Refson, K.; Payne, M. C. First Principles Methods Using Castep. *Zeitschrift für kristallographie-crystalline materials* **2005**, *220*, 567–570.
- (71) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Dal Corso, A.; de Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. Quantum Espresso: A Modular and Open-Source Software Project for Quantum Simulations of Materials. *J. Phys.: Condens. Matter* **2009**, *21*, No. 395502.
- (72) Gonze, X.; Beuken, J.-M.; Caracas, R.; Detraux, F.; Fuchs, M.; Rignanese, G.-M.; Sindic, L.; Verstraete, M.; Zerah, G.; Jollet, F. First-Principles Computation of Material Properties: The Abinit Software Project. *Comput. Mater. Sci.* **2002**, *25*, 478–492.
- (73) Nuruzzaman, M.; Islam, M. A.; Alam, M. A.; Shah, M.; Karim, A. Structural, Elastic and Electronic Properties of 2H- and 4H-SiC. *Int. J. Eng. Res. Appl.* **2015**, *5*, 48–52.
- (74) Madelung, O.; Schulz, M.; Weiss Landoldt-Bornstein III; Springer: Berlin, 1982.
- (75) Djemla, P.; Roussigné, Y.; Dirras, G. F.; Jackson, K. M. Elastic Properties of B-SiC Films by Brillouin Light Scattering. *J. Appl. Phys.* **2004**, *95*, 2324–2330.

- (76) Schulz, H.; Thiemann, K. Structure Parameters and Polarity of the Wurtzite Type Compounds SiC-2h and ZnO. *Solid State Commun.* **1979**, *32*, 783–785.
- (77) Kamitani, K.; Grimsditch, M.; Nipko, J. C.; Loong, C. K.; Okada, M.; Kimura, I. The Elastic Constants of Silicon Carbide: A Brillouin-Scattering Study of 4h and 6h SiC Single Crystals. *J. Appl. Phys.* **1997**, *82*, 3152–3154.
- (78) Harl, J.; Schimka, L.; Kresse, G. Assessing the Quality of the Random Phase Approximation for Lattice Constants and Atomization Energies of Solids. *Phys. Rev. B* **2010**, *81*, No. 115126.
- (79) Paier, J.; Marsman, M.; Kresse, G. Why Does the B3lyp Hybrid Functional Fail for Metals? *J. Chem. Phys.* **2007**, *127*, No. 024103.
- (80) Harrison, W. A. *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond*; Courier Corporation, 2012.
- (81) Carnahan, R. Elastic Properties of Silicon Carbide. *J. Am. Ceram. Soc.* **1968**, *51*, 223–224.
- (82) Liu, X.; Li, L.; Li, Q.; Li, Y.; Lu, F. Optical and Mechanical Properties of C, Si, Ge, and 3c-SiC Determined by First-Principles Theory Using Heyd–Scuseria–Ernzerhof Functional. *Mater. Sci. Semicond. Process.* **2013**, *16*, 1369–1376.
- (83) Luo, X.; Goel, S.; Reuben, R. L. A Quantitative Assessment of Nanometric Machinability of Major Polytypes of Single Crystal Silicon Carbide. *J. Eur. Ceram. Soc.* **2012**, *32*, 3423–3434.
- (84) Wolfenden, A.; Oliver, K.; Singh, M. Measurements of Elastic and Anelastic Properties of Reaction-Formed Silicon Carbide Ceramics. *J. Mater. Sci.* **1996**, *31*, 6073–6076.
- (85) Kittel, C. *Introduction to Solid State Physics*; John Wiley & Sons, Inc: New York, 1996; 402.
- (86) Donohue, J. *Structures of the Elements*; osti, 1974.
- (87) McSkimin, H. Measurement of Elastic Constants at Low Temperatures by Means of Ultrasonic Waves—Data for Silicon and Germanium Single Crystals, and for Fused Silica. *J. Appl. Phys.* **1953**, *24*, 988–997.
- (88) Greene, R. G.; Luo, H.; Ruoff, A. L.; Trail, S. S.; DiSalvo, F. J., Jr Pressure Induced Metastable Amorphization of Bas: Evidence for a Kinetically Frustrated Phase Transformation. *Phys. Rev. Lett.* **1994**, *73*, 2476.
- (89) Knittle, E.; Wentzcovitch, R. M.; Jeanloz, R.; Cohen, M. L. Experimental and Theoretical Equation of State of Cubic Boron Nitride. *Nature* **1989**, *337*, 349–352.
- (90) Trampert, A.; Brandt, O.; Ploog, K. Crystal Structure of Group III Nitrides. In *Semiconductors and Semimetals*; Elsevier: 1997; *50*, 167–192.
- (91) Sherwin, M. E.; Drummond, T. J. Predicted Elastic Constants and Critical Layer Thicknesses for Cubic Phase Aln, Gan, and Inn on B-SiC. *J. Appl. Phys.* **1991**, *69*, 8423–8425.
- (92) Hellwege, K.; Madelung, O. Semi-Conductor, Intrinsic Properties of Group IV Elements and III–V, II–VI and I–VII Compounds. *Landolt-Bornstein New Series, Group III*, 1987, *22*, 9.
- (93) Weil, R.; Groves, W. O. The Elastic Constants of Gallium Phosphide. *J. Appl. Phys.* **1968**, *39*, 4049–4051.
- (94) Boyle, W.; Sladek, R. Elastic Constants and Lattice Anharmonicity of GaSb and GaP from Ultrasonic-Velocity Measurements between 4.2 and 300 K. *Phys. Rev. B* **1975**, *11*, 2933.
- (95) Ahmed, R.; Hashemifar, S. J.; Akbarzadeh, H.; Ahmed, M. Ab Initio Study of Structural and Electronic Properties of III-Arsenide Binary Compounds. *Comput. Mater. Sci.* **2007**, *39*, 580–586.
- (96) Madelung, O. Numerical Data and Functional Relationships in Science and Technology. *Landolt Bornstein, New Series, Group III*, 1982, *22*, 117.
- (97) Sclar, C.; Garrison, L.; Schwartz, C. Indium Telluride (Ii'): Transitory Intermediate Phase in the Transformation Inte (Ii) to Inte (I). *Science* **1965**, *147*, 1569–1571.
- (98) Chattopadhyay, T.; Santandrea, R.; Von Schnerring, H. A New High Pressure Phase of InTe. *Phys. B+C* **1986**, *139*, 353–355.
- (99) Chen, X.-R.; Li, X.-F.; Cai, L.-C.; Zhu, J. Pressure Induced Phase Transition in ZnS. *Solid State Commun.* **2006**, *139*, 246–249.
- (100) Giesecke, G.; Pfister, H. Präzisionsbestimmung Der Gitterkonstanten Von Aiiibv-Verbindungen. *Acta Crystallogr.* **1958**, *11*, 369–371.
- (101) Nichols, D.; Rimai, D.; Sladek, R. Elastic Anharmonicity of InP: Its Relationship to the High Pressure Transition. *Solid State Commun.* **1980**, *36*, 667–669.
- (102) Ozolin'sh, J.; Averkieva, G.; Ievin'sh, A.; Gozyunova, N. An X-Ray Study of Some A3b5 Compounds Which Display Deviation from Stoichiometry. *Sov. Phys. Cryst.* **1963**, *7*, 691.
- (103) Gerlich, D. Elastic Constants of Single-Crystal Indium Arsenide. *J. Appl. Phys.* **1963**, *34*, 2915–2915.
- (104) Sirota, N. Heats of Formation and Temperatures and Heats of Fusion of Compounds Aiii Bv. In *Semiconductors and Semimetals*; Elsevier, 1968; *4*, 35–162.
- (105) Van Camp, P.; Van Doren, V.; Devreese, J. Pressure Dependence of the Electronic Properties of Cubic Iii-V in Compounds. *Phys. Rev. B* **1990**, *41*, 1598.
- (106) Chokappa, D. K.; Clancy, P. A Computer Simulation Study of the Melting and Freezing Properties of a System of Lennard-Jones Particles: I. Melting the Solid. *Mol. Phys.* **1987**, *61*, 597–615.
- (107) Chen, M.; Hung, L.; Huang, C.; Xia, J.; Carter, E. A. The Melting Point of Lithium: An Orbital-Free First-Principles Molecular Dynamics Study. *Mol. Phys.* **2013**, *111*, 3448–3456.
- (108) Lide, D. R. *Crc Handbook of Chemistry and Physics*; CRC press, 2004; 85.
- (109) Morris, J. R.; Wang, C.; Ho, K.; Chan, C. T. Melting Line of Aluminum from Simulations of Coexisting Phases. *Phys. Rev. B* **1994**, *49*, 3109.
- (110) Li, R.; Lee, E.; Luo, T. A Unified Deep Neural Network Potential Capable of Predicting Thermal Conductivity of Silicon in Different Phases. *J. Mater. Today Phys.* **2020**, *12*, No. 100181.
- (111) Piechota, J.; Kruckowski, S.; Sadovyi, B.; Sadovyi, P.; Porowski, S.; Grzegory, I. Melting Versus Decomposition of Gan: Ab Initio Molecular Dynamics Study and Comparison to Experimental Data. *J. Chem. Mater.* **2023**, *35*, 7694–7707.
- (112) Liang, A.; Liu, Y.; Shi, L.; Lei, L.; Zhang, F.; Hu, Q.; He, D. Melting Temperature of Diamond and Cubic Boron Nitride at 15 Gigapascals. *Phys. Rev. Res.* **2019**, *1*, No. 033090.
- (113) Gleiman, S.; Chen, C.-K.; Datye, A.; Phillips, J. Melting and Spheroidization of Hexagonal Boron Nitride in a Microwave-Powered, Atmospheric Pressure Nitrogen Plasma. *J. Mater. Sci.* **2002**, *37*, 3429–3440.
- (114) Togo, A.; Chaput, L.; Tadano, T.; Tanaka, I. Implementation Strategies in Phonopy and Phono3py. *J. Phys.: Condens. Matter* **2023**, *39*, 353001 DOI: [10.1088/1361-648X/acd831](https://doi.org/10.1088/1361-648X/acd831).
- (115) Togo, A. First-Principles Phonon Calculations with Phonopy and Phono3py. *J. Phys. Soc. Jpn.* **2023**, *92*, No. 012001.
- (116) Waseda, Y.; Shinoda, K.; Sugiyama, K.; Takeda, S.; Terashima, K.; Toguri, J. M. High Temperature X-Ray Diffraction Study of Melt Structure of Silicon. *Jpn. J. Appl. Phys.* **1995**, *34*, 4124.
- (117) Laaziri, K.; Kycia, S.; Roorda, S.; Chicoin, M.; Robertson, J.; Wang, J.; Moss, S. High Resolution Radial Distribution Function of Pure Amorphous Silicon. *Phys. Rev. Lett.* **1999**, *82*, 3460.
- (118) Remsing, R. C.; Klein, M. L.; Sun, J. Dependence of the Structure and Dynamics of Liquid Silicon on the Choice of Density Functional Approximation. *Phys. Rev. B* **2017**, *96*, No. 024203.
- (119) Qian, X.; Peng, S.; Li, X.; Wei, Y.; Yang, R. Thermal Conductivity Modeling Using Machine Learning Potentials: Application to Crystalline and Amorphous Silicon. *Mater. Today Phys.* **2019**, *10*, No. 100140.
- (120) Babaei, H.; Guo, R.; Hashemi, A.; Lee, S. Machine-Learning-Based Interatomic Potential for Phonon Transport in Perfect Crystalline Si and Crystalline Si with Vacancies. *Phys. Rev. Mater.* **2019**, *3*, No. 074603.
- (121) Oyake, T.; Feng, L.; Shiga, T.; Isogawa, M.; Nakamura, Y.; Shiomi, J. Ultimate Confinement of Phonon Propagation in Silicon Nanocrystalline Structure. *Phys. Rev. Lett.* **2018**, *120*, No. 045901.
- (122) McGaughey, A. J. H.; Jain, A.; Kim, H. Y.; Fu, B. Phonon Properties and Thermal Conductivity from First Principles, Lattice Dynamics, and the Boltzmann Transport Equation. *J. Appl. Phys.* **2019**, *125*, No. 011101.

(123) Krbal, M.; Kolobov, A. V.; Hyot, B.; André, B.; Fons, P.; Simpson, R. E.; Uruga, T.; Tanida, H.; Tominaga, J. Amorphous Insb: Longer Bonds yet Higher Density. *J. Appl. Phys.* **2010**, *108*, No. 023506.