

Machine learning-assisted optimization of MTO basis sets

Zhiqiang Li(李志强), and Lei Wang(王蕾)

Citation: Chin. Phys. B, 2026, 35 (**1**): 016301. DOI: 10.1088/1674-1056/ae0a39
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Unveiling the role of high-order anharmonicity in thermal expansion: A first-principles perspective

Tianxu Zhang(张天旭), Kun Zhou(周琨), Yingjian Li(李英健), Chenhao Yi(易晨浩), Muhammad Faizan, Yuhao Fu(付钰豪), Xinjiang Wang(王新江), and Lijun Zhang(张立军)
Chin. Phys. B, 2025, 34 (**4**): 046301. DOI: 10.1088/1674-1056/adb94c

First-principles study of electronic and magnetic properties of self-intercalated van der Waals magnet $Cr_3Ge_2Te_6$

Jia-wan Li(李家万), Shi-Bo Zhao(赵世博), Lin Zhuang(庄琳), and Yusheng Hou(侯玉升)
Chin. Phys. B, 2025, 34 (**3**): 036301. DOI: 10.1088/1674-1056/adab67

Strain-tuned electronic and valley-related properties in Janus monolayers of SWSiX_2 (X = N, P, As)

Yunxi Qi(戚云西), Jun Zhao(赵俊), and Hui Zeng(曾晖)
Chin. Phys. B, 2024, 33 (**9**): 096302. DOI: 10.1088/1674-1056/ad6077

Two-dimensional $Cr_2Cl_3S_3$ Janus magnetic semiconductor with large magnetic exchange interaction and high-T_c

Lei Fu(伏磊), Shasha Li(李沙沙), Xiangyan Bo(薄祥奘), Sai Ma(马赛), Feng Li(李峰), and Yong Pu(普勇)
Chin. Phys. B, 2024, 33 (**9**): 096301. DOI: 10.1088/1674-1056/ad5538

Optical spectrum of ferrovalley materials: A case study of Janus H-VSSe

Chao-Bo Luo(罗朝波), Wen-Chao Liu(刘文超), and Xiang-Yang Peng(彭向阳)
Chin. Phys. B, 2024, 33 (**1**): 016303. DOI: 10.1088/1674-1056/acf660

Machine learning-assisted optimization of MTO basis sets

Zhiqiang Li(李志强) and Lei Wang(王蕾)[†]

Key Laboratory of Quantum Materials and Devices of Ministry of Education, School of Physics, Southeast University, Nanjing 211189, *China*

(Received 7 May 2025; revised manuscript received 19 August 2025; accepted manuscript online 23 September 2025)

First-principles calculations based on density functional theory (DFT) have had a significant impact on chemistry, physics, and materials science, enabling in-depth exploration of the structural and electronic properties of a wide variety of materials. Among different implementations of DFT, the plane-wave method is widely used for periodic systems because of its high accuracy. However, this method typically requires a large number of basis functions for large systems, leading to high computational costs. Localized basis sets, such as the muffin-tin orbital (MTO) method, have been introduced to provide a more efficient description of electronic structure with a reduced basis set, albeit at the cost of reduced computational accuracy. In this work, we propose an optimization strategy using machine-learning techniques to automate MTO basis-set parameters, thereby improving the accuracy and efficiency of MTO-based calculations. Default MTO parameter settings primarily focus on lattice structure and give less consideration to element-specific differences. In contrast, our optimized parameters incorporate both structural and elemental information. Based on these converged parameters, we successfully recovered missing bands for $CrTe_2$. For the other three materials — Si, GaAs, and CrI_3 — we achieved band improvements of up to 2 eV. Furthermore, the generalization of the machine-learned method is validated by perturbation, strain, and elemental substitution, resulting in improved band structures. Additionally, lattice-constant optimization for GaAs using the converged parameters yields closer agreement with experiment.

Keywords: first-principles calculations, muffin-tin orbital, machine learning

PACS: 63.20.dk, 89.20.Ff

DOI: 10.1088/1674-1056/ae0a39 **CSTR:** 32038.14.CPB.ae0a39

1. Introduction

First-principles calculations based on density functional theory (DFT) are formulated through the Kohn–Sham (KS) equations,[1] incorporating exchange–correlation functionals[2–8] and norm-conserving pseudopotentials.[9] With the continuous advancement of computational resources, DFT has emerged as a robust and widely adopted approach for determining the ground-state properties of materials. It enables the investigation of various structural and electronic properties and has had a profound impact on condensed matter physics, chemistry, and materials science.

In first-principles calculations, the choice of basis sets plays a critical role in determining both accuracy and computational efficiency. Two widely used types of basis sets are non-localized plane-wave basis sets and localized basis sets. Plane-wave methods, such as those implemented in the Vienna *ab initio* simulation package (VASP)[10,11] and WIEN2K,[12] offer systematically convergent and highly accurate results for periodic systems. However, the required large number of basis functions often limits the applicability of the plane-wave method to relatively small systems. Alternatively, various types of localized basis sets, which enable simulations of much larger systems — ranging up to tens of thousands of atoms — albeit at the cost of reduced systematic accuracy, are utilized in several electronic-structure codes, including those discussed below. SIESTA,[13–15] OPENMX,[16–18] and ABACUS[19,20] use numerical atomic orbitals, CP2K[21] uses Gaussian-type orbitals, and NMTO[22–24] uses muffin-tin orbitals (MTO). Specifically, a 500 eV plane-wave cutoff energy in VASP corresponds to a basis size of 48100, approximately 33 times larger than the 1440 basis functions used in the MTO approach with spd orbitals when taking a primitive cell of yttrium iron garnet (YIG) containing 80 atoms as an example.

Generally, the MTO method performs particularly well for close-packed systems such as Fe, Co, and Ni. However, for open or low-symmetry structures, the method encounters symmetry issues due to an inefficient choice of basis sets. To address this issue, additional "empty spheres" — non-physical atoms with atomic number 0 — are introduced to improve the basis-set representation. The many parameters depend sensitively on the positions and radii of all empty spheres and atomic spheres, making the optimization of basis sets a high-dimensional and nontrivial task, with manual tuning becoming inefficient and unreliable, especially for complex structures. To overcome this challenge, we incorporate machine-learning techniques to automate the optimization of basis-set parameters. This approach enables systematic exploration of the parameter space and leads to improved accuracy in MTO calculations, making them more competitive with plane-wave methods.

The organization of this paper is as follows. In Section 2,

[†]Corresponding author. E-mail: wanglei.icer@seu.edu.cn

© 2026 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

we introduce the choice of basis sets, atomic parameters, and the machine-learning strategy. Meanwhile, the relevant loss function and computational parameters are also provided to ensure reproducibility. In Section 3, we study four materials, including Si, GaAs, $CrTe_2$, and CrI_3. Improved electronic band structures are obtained across all systems, with notable enhancement in the unoccupied states. Furthermore, the generalization of the machine-learned parameters is demonstrated through studies of perturbation, strain, elemental substitution, and lattice-constant optimization. Finally, a conclusion is given in Section 4.

2. Method

2.1. MTO method

The MTO method has undergone continuous development and refinement over the years, resulting in several well-established variants, including the linear muffin-tin orbital (LMTO) method,[24] the exact muffin-tin orbital (EMTO) method,[25] and the N-th order muffin-tin orbital (NMTO) method.[22,23] The LMTO approach employs a basis set comprising two orbitals, whereas both EMTO and NMTO utilize three orbitals,[25] intersecting at points a_r and s_r in Fig. 1(a). Here, a_r denotes the matching point between ψ and φ, and s_r denotes the matching point between φ and φ. The explicit forms of the three orbitals are defined as follows:

$$\Phi_{RL}(\varepsilon, r_R) = \phi_{RL} + \psi_{RL} - \varphi_{RL}, \tag{1}$$

$$\phi_{RL}(\varepsilon, r_R) = \chi_{RL}(\varepsilon, r_R) Y_L(r_R), \tag{2}$$

$$\psi_{RL}(\kappa^2, r_R) = n_{RL}(\kappa^2, r_R)\,\delta_{RR'} + \sum_{R'L'} j_{R'L'}(\kappa^2, r_{R'})\, S_{R'L',RL}(\kappa^2), \tag{3}$$

$$\varphi_{RL}(\varepsilon, r_R) = n_{RL}(\kappa^2, r_R) + j_{RL}(\kappa^2, r_R)\, D_{RL}(\varepsilon). \tag{4}$$

Here, ϕ_{RL} denotes the spherical harmonic inside the potential sphere of radius s_r. ψ_{RL} represents the screened spherical wave (SSW) outside the hard sphere of radius a_r. $\kappa^2(=\varepsilon - V_{mtz})$ defines the kinetic energy, where V_{mtz} is the muffin-tin zero. The functions n_{RL} and j_{RL} are composed of spherical Bessel and Neumann functions, respectively. S is the corresponding screened slope matrix. φ is the free-electron solution, designed to handle the connectivity at the boundary. D_{RL} denotes the logarithmic derivative of the free-electron solution at a_r. The MTO approach provides a balance of computational efficiency and accuracy, particularly in the study of close-packed systems, where it performs comparably to plane-wave methods and can be applied in transport calculations.[26–30]

2.2. Machine learning

In this work, we choose the covariance matrix adaptation evolution strategy (CMA-ES), a type of machine-learning algorithm. This gradient-free, black-box optimization method explores the parameter space via adaptive step-size control and covariance matrix updates, improving search efficiency in high-dimensional and non-convex spaces. The algorithm is particularly well-suited for black-box optimization problems, as is the case here, where the mapping between atomic sphere parameters (e.g., positions and radii of atomic spheres) and band structure accuracy is non-analytic and not directly differentiable. The optimization is performed using the Optuna framework,[31,32] a widely used and flexible platform for automated hyperparameter optimization. The fundamental expression of the optimization scheme is given by

$$x_k^{(g+1)} \sim m^{(g)} + \sigma^{(g)} N(0, C^g) \quad \text{for} \quad k = 1, 2, \ldots, \lambda. \tag{5}$$

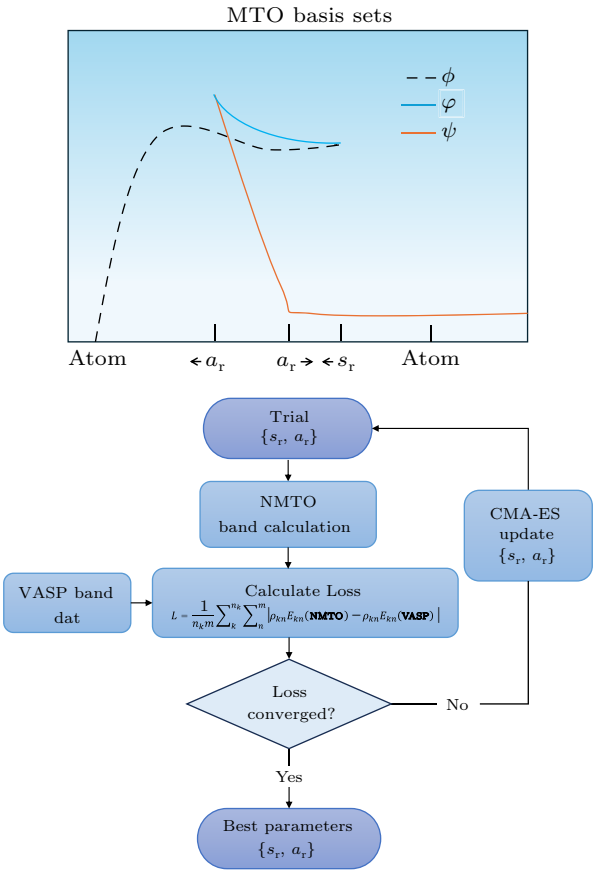


Fig. 1. (a) Three MTO basis sets in NMTO. (b) The flowchart of the algorithm, including the update of parameters s_r and a_r and the design of the loss function.

In Eq. (5), x denotes the input parameter vector sampled from a multivariate normal distribution N with mean m and generation-dependent covariance matrix C. Here, g refers to the generation index, and k indicates the dimensionality of the parameter space. The scalar $\sigma^{(g)}$ represents the global step size, i.e., the overall standard deviation governing the sampling process. In the present context, the optimization variables x correspond to the parameters of the basis set, specifically the potential sphere radius s_r and hard sphere radius a_r for all atoms, along with their associated loss function values

L, which will be introduced in Subsection 2.3. These quantities are iteratively updated following the CMA-ES algorithm. A schematic illustration of the optimization workflow, including all relevant parameters, is shown in Fig. 1(b).

2.3. Loss function

Before training the parameters, it is necessary to obtain benchmark band structures for comparison. These reference bands can be derived either from experimental measurements or from high-accuracy plane-wave DFT calculations. In this work, the band structure obtained from plane-wave-based VASP calculations serves as the reference. The loss function L, which quantifies the discrepancy between the NMTO and VASP band structures using a manually defined error weighting scheme, is defined as

$$L = \frac{1}{n_k m} \sum_{k=1}^{n_k} \sum_{n=1}^{m} \left| \rho_{kn} E_{kn}\,(\mathrm{NMTO}) - \rho_{kn} E_{kn}\,(\mathrm{VASP}) \right|. \quad (6)$$

Here, k and n denote the indices of the K-points and energy bands, respectively. Since high-symmetry points in the Brillouin zone encode significant structural information, the band energies E_{kn} are selected along paths connecting these symmetry points, which depend on the crystal structure under consideration. In practice, physical properties such as electronic transport, magnetism, and optical response are often dominated by states near the Fermi surface. Accordingly, a weight factor ρ_{kn} is introduced to emphasize contributions from energy levels close to the Fermi level. For example, a larger value of ρ_{kn} indicates that the corresponding part of the band structure has a greater impact on the total loss, thus guiding the direction of MTO training. Conversely, a smaller ρ_{kn} reduces the influence of that band region on the total loss, causing it to be effectively neglected during MTO training. This typically requires prior physical insight and iterative adjustment to achieve the most reliable match between NMTO and reference data.

3. Computation details

The relevant structural parameters used in this work are listed in Table 1. First-principles calculations were performed within the framework of density functional theory (DFT) using established computational approaches.

On the one hand, calculations based on the projector augmented-wave (PAW) method[33,34] and a plane-wave basis set, as implemented in VASP, were performed with the Perdew–Burke–Ernzerhof (PBE)[35] exchange–correlation functional within the generalized gradient approximation (GGA). Computational details are provided below to ensure reproducibility. For Si, a plane-wave energy cutoff of 500 eV and a $9\times9\times9$ Monkhorst–Pack K-point mesh are employed. For GaAs, a cutoff energy of 600 eV and a $9\times9\times9$ K-point mesh are used. For GaP, a cutoff energy of 600 eV and a $9\times9\times9$ K-point mesh are used. For the 1T-phase bulk $CrTe_2$,[36] a cutoff energy of 500 eV and a $7\times7\times7$ K-point mesh are adopted. For CrI_3,[37] a cutoff energy of 500 eV and a $9\times9\times9$ K-point mesh are used. A Gaussian smearing scheme with a width of 0.05 eV is applied for electronic structure calculations in both semiconducting and metallic systems.

Table 1. Structural parameters of all materials used in this paper; units of the lattice constants a, b, and c are Å.

Material	a	b	c	Space group
Si	5.429	5.429	5.429	$Fd\bar{3}m$
GaAs	5.664	5.664	5.664	$Fd\bar{3}m$
$CrTe_2$	3.788	3.788	6.095	$P\bar{3}m1$
CrI_3	6.867	6.867	19.807	$R\bar{3}$
GaP	5.492	5.492	5.492	$Fd\bar{3}m$

On the other hand, DFT calculations employing the MTO method, as implemented in NMTO, use the von Barth–Hedin local exchange–correlation potential within the local density approximation (LDA).[38] For Si, 3s, 3p, 3d orbitals are selected for the Si atom and a $9\times9\times9$ Monkhorst–Pack K-point mesh is employed. For GaAs, 4s, 4p, 3d orbitals are selected for the Ga atom and 4s, 4p, 4d orbitals for the As atom, with a $9\times9\times9$ K-point mesh. For GaP, 4s, 4p, 3d orbitals are selected for the Ga atom and 3s, 3p, 3d orbitals for the P atom, with a $9\times9\times9$ K-point mesh. For $CrTe_2$,[36] 4s, 4p, 3d orbitals are selected for the Cr atom and 5s, 5p, 5d, 4f orbitals for the Te atom, with a $7\times7\times7$ K-point mesh. For CrI_3,[36] 4s, 4p, 3d orbitals are selected for the Cr atom and 5s, 5p, 5d, 4f orbitals for the I atom, with a $9\times9\times9$ K-point mesh.

4. Result

4.1. Band structure improvement by machine-learning method

We first demonstrate the robustness of our algorithm using Si as a representative example, as shown in Fig. 2(a). The original Si structure exhibits unoccupied regions within its spatial framework. Upon introducing empty spheres, however, the interior of Si becomes substantially filled. All subsequent band structure results presented are based on calculations incorporating these empty spheres. The "default parameters" refer to those automatically generated by the NMTO method, while the "converged parameters" are obtained through an exploration-based algorithm. The results in Fig. 2(b) demonstrate the reliability of our MTO method, showing that for simple materials, it can achieve an accuracy comparable to that of plane-wave methods. Furthermore, Fig. 2(c) highlights the strong exploration capability of our algorithm, revealing an improvement of approximately 0.2 eV near the Fermi level

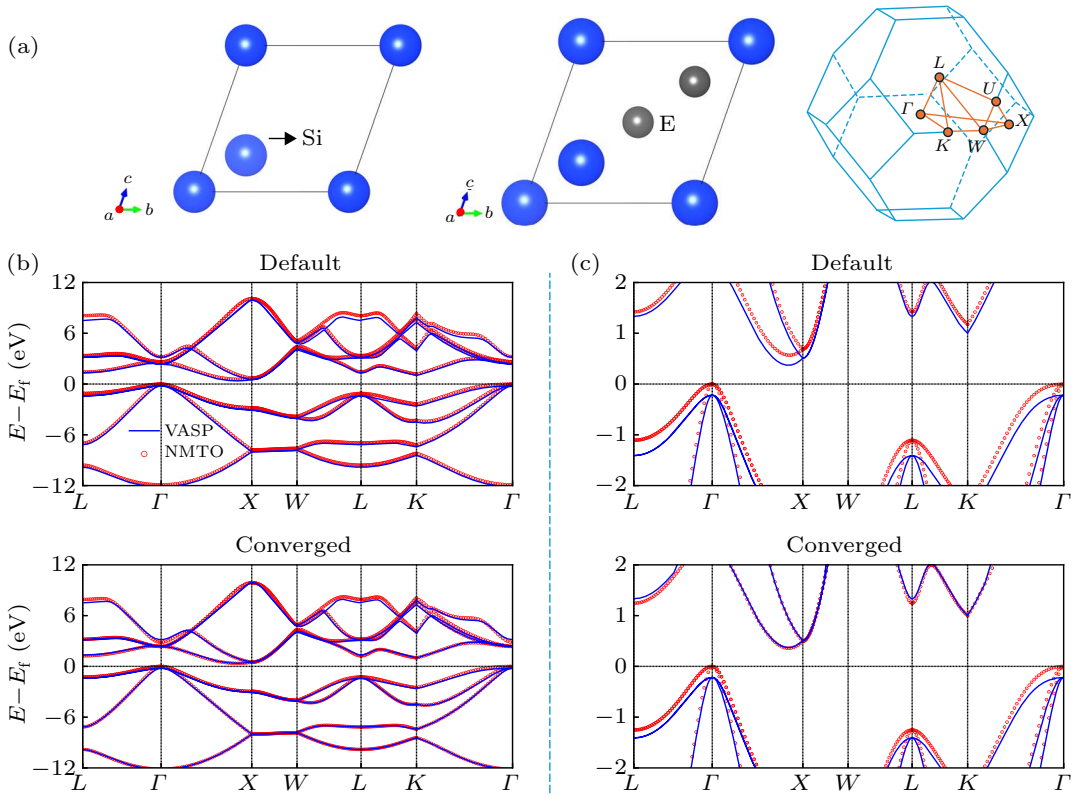


Fig. 2. (a) The structure of Si: without empty spheres and with empty spheres, including the three-dimensional Brillouin zone of the face-centered cubic lattice. (b) Band comparison over a larger energy range, and (c) near the Fermi level, before and after optimization.

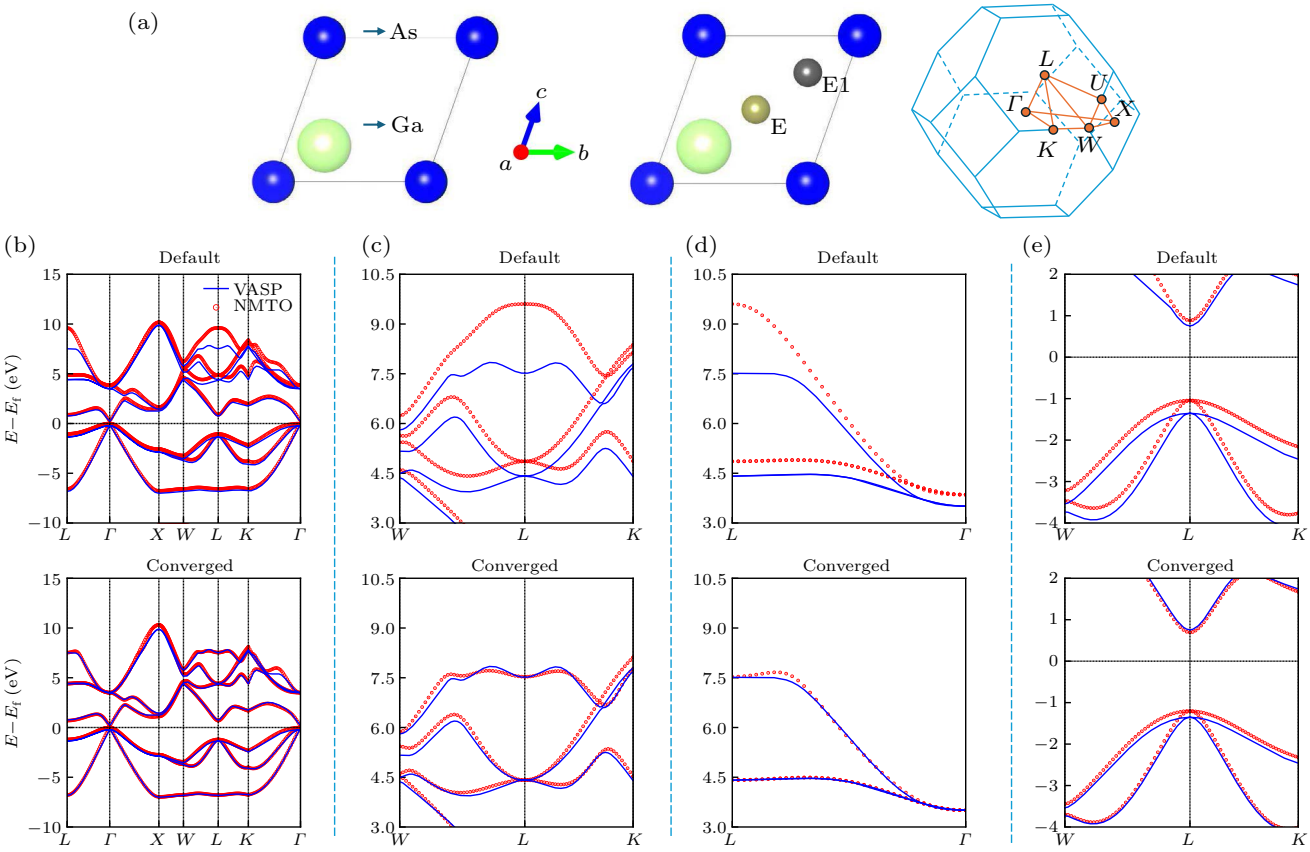


Fig. 3. (a) The structure of GaAs: without empty spheres and with empty spheres, including the three-dimensional Brillouin zone of the face-centered cubic lattice. (b) Band comparison before and after optimization along the high-symmetry path L–Γ–X–W–L–K–Γ. The extracted energy ranges for the band comparison are displayed in panels (c) and (e) for the path W–L–K, and in panel (d) for the path L–Γ.

even when the overall band structure is already reasonably accurate. Possible reasons for this improvement will be discussed below. As shown in Table 2, under the default parameters, the hard sphere and potential sphere radii of the empty spheres and Si are identical. After optimization, however, these radii are adjusted, leading to an improvement in the corresponding band structures, as previously discussed. We attribute this to the fact that the default parameters primarily consider spatial filling when introducing empty spheres, whereas the converged parameters additionally account for the different types between Si atoms and empty spheres. Consequently, the calculated band structures are further improved.

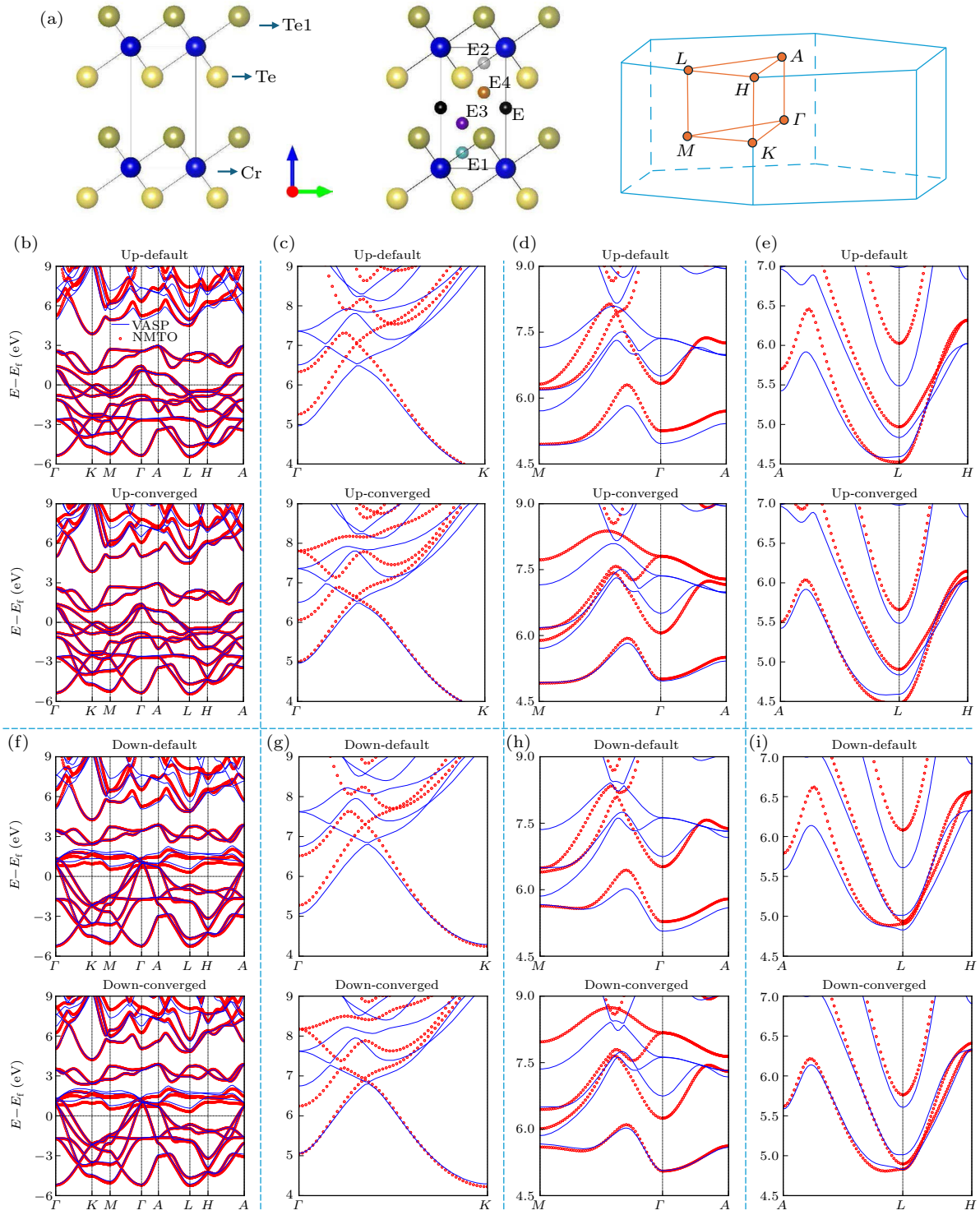


Fig. 4. (a) The structure of $CrTe_2$: without empty spheres and with empty spheres, including the three-dimensional Brillouin zone of the hexagonal lattice. Spin-up and spin-down results are shown. For spin-up, the results along the high-symmetry path Γ–K–M–Γ–A–L–H–A are presented in panel (b) within the energy range of -6 to 9 eV. The extracted results for occupied states are shown along the paths Γ–K (c), M–Γ–A (d), and A–L–H (e). For spin-down, the same analysis is performed, with results extracted along the paths Γ–K (g), M–Γ–A (h), and A–L–H (i), as shown in panel (f) along the path Γ–K–M–Γ–A–L–H–A.

Table 2. Default and converged parameters of Si in NMTO in units of Bohr.

Sphere	Default a_r	Converged a_r	Default s_r	Converged s_r
Si	1.776	1.541	3.110	3.082
E	1.776	1.645	3.110	3.486

Next, we extended our study to the conventional material GaAs. The structures before and after the addition of empty spheres are shown in Fig. 3(a). All subsequent band structure results are based on the structures with empty spheres. As shown in Table 3, in the default parameter set, the hard sphere and potential sphere radii of As and E are identical, as are those of Ga and E1. This parameter setting gives less

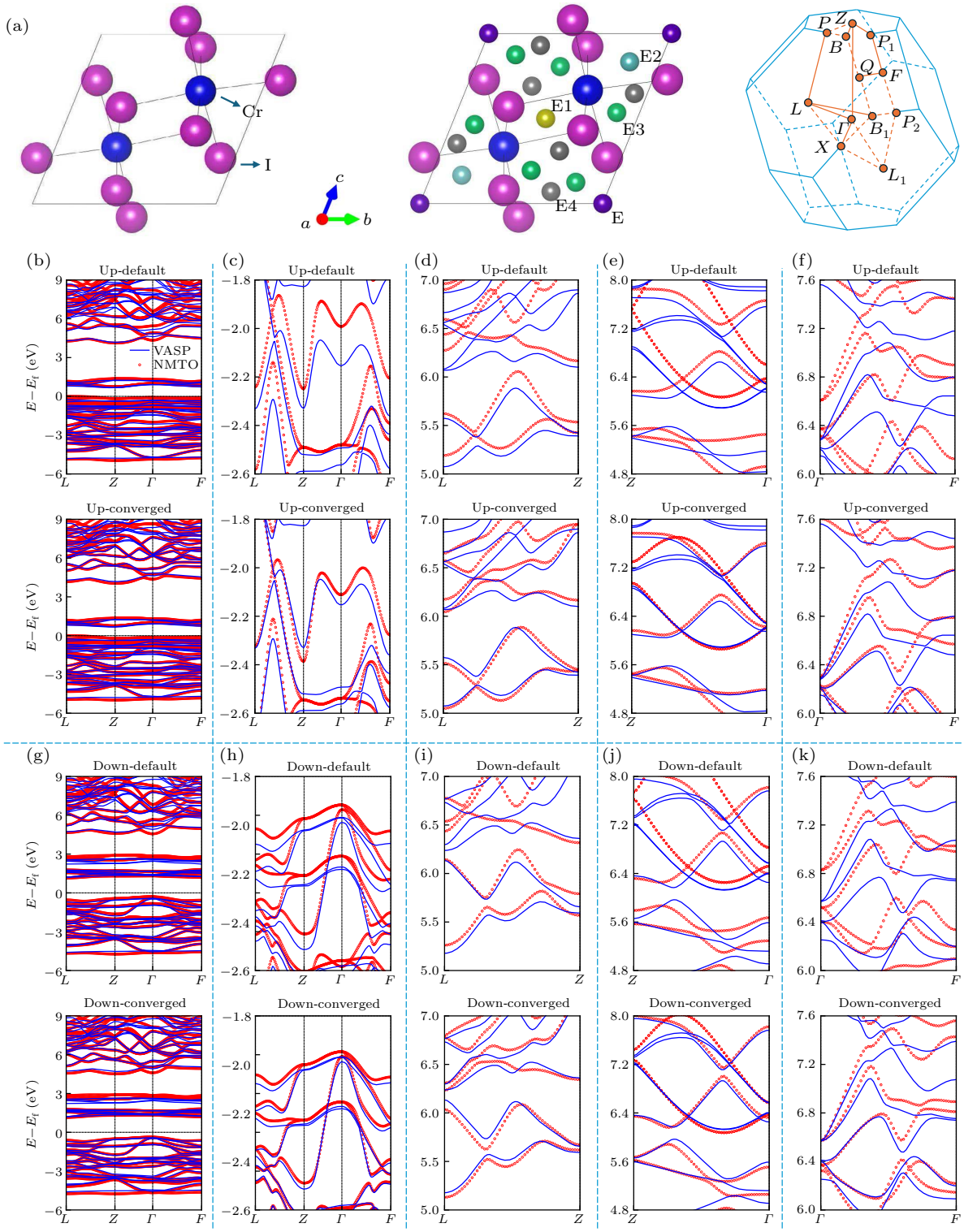


Fig. 5. (a) The structure of CrI_3: without empty spheres and with empty spheres, including the three-dimensional Brillouin zone of the rhombohedral lattice. Spin-up and spin-down results are presented. For spin-up, the overall band structure is shown in (b), the occupied states along the high-symmetry path L–Z–Γ–F are shown in panel (c) within the energy range of 6 to 9 eV, and the unoccupied states are extracted from the paths $L-Z$ (d), Z–Γ (e), and Γ–F (f). For spin-down, the overall band structure is shown in (g), the occupied states are shown in panel (h) along the path L–Z–Γ–F, and the unoccupied states are extracted from the paths L–Z (i), Z–Γ (j), and Γ–F (k).

consideration to the differences in chemical species between atoms and empty spheres. In contrast, the converged parameters we obtained assign distinct radii to each element of empty spheres and atomic spheres. As shown in Fig. 3(b), the NMTO method with default parameters can, in general, reproduce the plane-wave results for GaAs. However, as illustrated in Figs. 3(c) and 3(d), the accuracy is insufficient at the high-symmetry L point. When our algorithm further accounts for the differences between atomic species by adjusting the empty sphere parameters, the band structure at the L point is significantly improved, correcting a 2 eV error in the 3–10.5 eV range and achieving a general improvement of approximately 0.2 eV near the Fermi level (Fig. 3(e)).

Table 3. Default and converged parameters of GaAs in NMTO in units of Bohr.

Sphere	Default a_r	Converged a_r	Default s_r	Converged s_r
As	1.930	1.303	3.334	3.421
Ga	1.773	1.780	3.155	3.118
E	1.930	2.290	3.334	3.659
E1	1.773	2.337	3.155	3.219

We also investigated the van der Waals material $CrTe_2$. The structural models before and after the addition of empty spheres are shown in Fig. 4(a), respectively. As summarized in Table 4, under the default parameter settings, the two types of Te atoms share the same sphere parameters, the Cr atoms and E spheres share another set of parameters, and the remaining empty spheres (E1, E2, E3, and E4) have identical parameters. In contrast, after optimization, all sphere parameters are distinguished. Figures 4(b) and 4(f) present the overall band structures for the up- and down-channels, respectively. We find that the converged parameters enable the calculation of previously missing bands along the high-symmetry point Γ, which were absent when using the default parameters, as illustrated in Figs. 4(c), 4(d), 4(g), and 4(h). Furthermore, as shown in Figs. 4(e) and 4(i), the converged parameters lead to an improvement of approximately 0.5 eV in the band energies at the L point.

Table 4. Default and converged parameters of $CrTe_2$ in NMTO in units of Bohr.

Sphere	Default a_r	Converged a_r	Default s_r	Convergeds_r
Te	2.199	2.504	3.727	3.478
Te1	2.199	2.075	3.727	3.641
Cr	1.828	2.134	3.325	3.431
E	1.828	2.259	3.325	2.949
E1	1.116	1.536	1.954	2.133
E2	1.116	1.139	1.954	2.527
E3	1.116	1.077	1.954	2.293
E4	1.116	0.754	1.954	1.963

Finally, we investigated CrI_3, another van der Waals material. The structure before and after introducing empty spheres is shown in Fig. 5(a), and the corresponding sphere parameters are summarized in Table 5. Figures 5(b) and 5(g) present the overall band structures for the spin-up and spin-down channels, respectively. As shown in Figs. 5(c) and 5(h), the converged parameters lead to a band improvement of approximately 0.2 eV for the occupied states at the Γ point. Moreover, along the other high-symmetry paths, the conduction bands exhibit corrected curvatures when using the converged parameters, even though some discrepancies in the band gap may still remain.

Table 5. Default and converged parameters of CrI_3 in NMTO in units of Bohr.

Sphere	Default a_r	Converged a_r	Default s_r	Converged s_r
Cr	2.281	2.463	3.814	3.947
I	1.836	2.031	3.394	3.154
E	2.056	1.803	3.780	3.563
E1	1.999	1.301	3.682	2.793
E2	1.985	1.994	3.656	3.555
E3	1.927	1.688	3.609	2.487
E4	1.234	1.157	2.341	2.209

4.2. Generalization of machine-learning method

To further assess the generalization of our machine learning method, we applied the parameters trained on the original systems to structures subject to strain and perturbation. In Fig. 6(a), $CrTe_2$ is considered under two modifications: a uniaxial tensile strain of 1% along the a-direction, and a random displacement of all atoms by 0.05 Å. In comparison with the band structure of the original system (Figs. 6(b) and 6(e)), we find that for both the strained (Figs. 6(c) and 6(f)) and perturbed (Figs. 6(d) and 6(g)) structures, our method with converged parameters continues to reproduce the bands at the Γ point that are absent in the default parameter calculations.

Beyond structural perturbations, we further tested the generalization of the optimized parameters to systems with elemental substitution. Specifically, we considered substituting elements from GaAs to GaP, which share the same crystal structure. Since the atomic sphere parameters depend on the lattice constant, we scale these parameters according to the lattice constant, with specific values detailed in Table 6. The structures before and after the addition of empty spheres are shown in Fig. 7(a). The overall band structure comparison is presented in Fig. 7(b). As shown in Figs. 7(c) and 7(d), the converged parameters achieve a band improvement of approximately 2 eV at the L point, demonstrating that the optimized parameters from GaAs can be successfully generalized to the GaP system.

Table 6. Default and converged parameters of GaP in NMTO in units of Bohr. The converged a_r and s_r values are derived by scaling the GaAs convergence parameters.

Sphere	Default a_r	Converged a_r	Default s_r	Converged s_r
P	1.851	1.265	3.192	3.316
Ga	1.743	1.728	3.101	3.022
E	1.851	2.223	3.192	3.120
E1	1.743	2.268	3.101	3.547

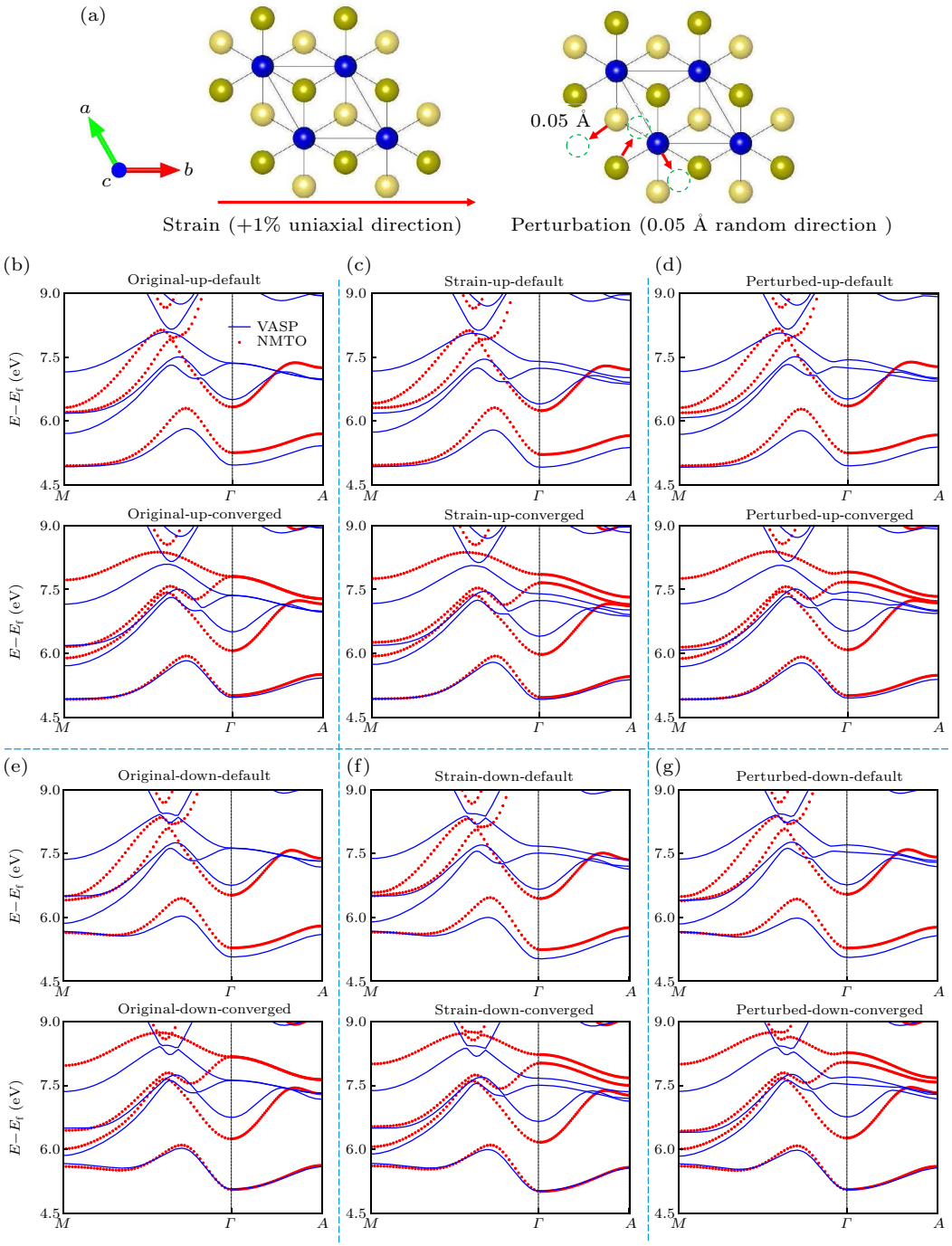


Fig. 6. (a) The stretched structure of $CrTe_2$, with a uniaxial tensile strain of 1% along the a-axis, and the perturbed structure of $CrTe_2$, with each atom subjected to a random displacement of 0.05 Å. Along the high-symmetry path *M*–Γ–*A* in the 4.5–9 eV range, the band structures for the original structure with default and converged parameters are shown in panels (b) and (e), respectively. Results for the strained structure are presented in panels (c) and (f), while those for the perturbed structure are displayed in panels (d) and (g).

We also examined the influence of the default and converged parameters on the lattice constant for GaAs and fitted the results using the Birch-Murnaghan[39] equation of state as follows:

$$E(V) = E_0 + \frac{9V_0B_0}{16}\left\{\left[\left(\frac{V_0}{V}\right)^{2/3} - 1\right]^3 B_0' + \left[\left(\frac{V_0}{V}\right)^{2/3} - 1\right]^2 \left[6 - 4\left(\frac{V_0}{V}\right)^{2/3}\right]\right\}, \quad (7)$$

where E_0 denotes the total energy at the equilibrium volume V_0, B_0 is the bulk modulus, and B_0' is the first derivative of the bulk modulus with respect to pressure. In Fig. 8, we show the calculated and fitted results using both the default and converged parameters. The horizontal axis represents the lattice volume, and the vertical axis shows the total energy shifted by a constant. The lattice constant obtained from the fit with default parameters is 5.59 Å, while that with converged parameters is 5.61 Å, which is closer to the experimental value of 5.64 Å.

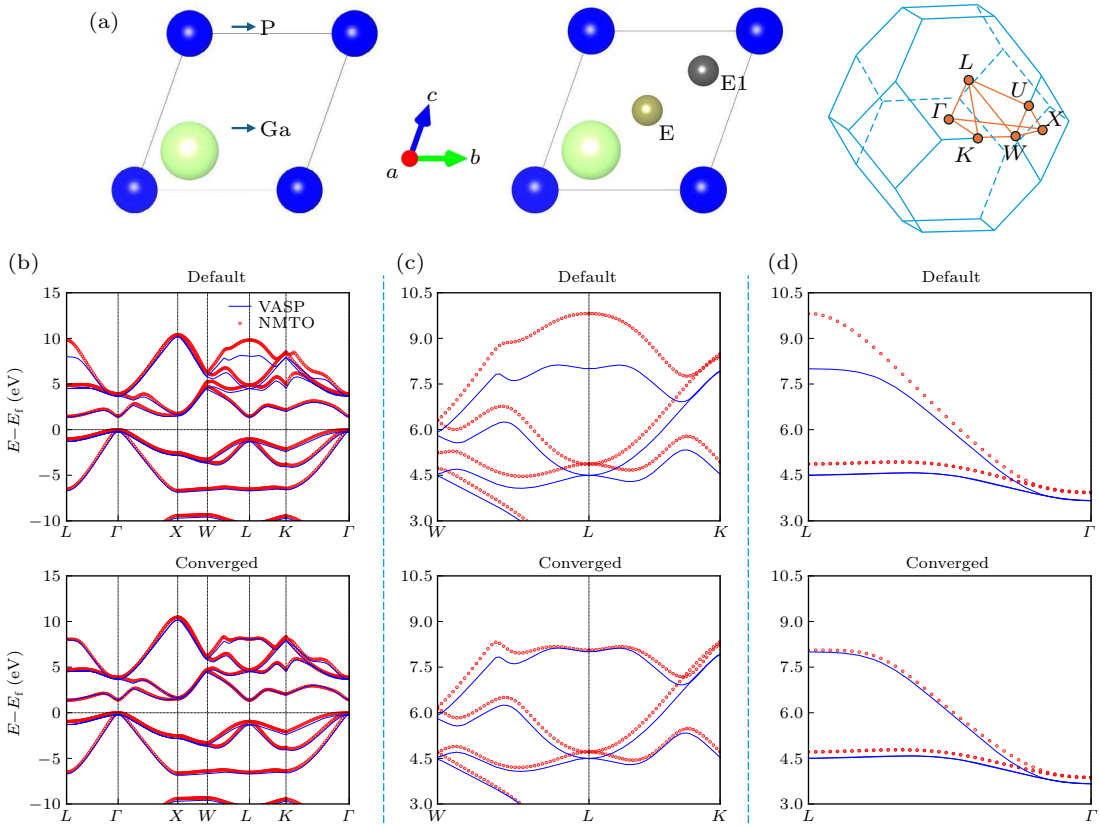


Fig. 7. (a) The structure of GaP: without empty spheres and with empty spheres, including the three-dimensional Brillouin zone of the face-centered cubic lattice. (b) Band comparison diagram before and after optimization along the high-symmetry path L–Γ–X–W–L–K–Γ. The extracted energy ranges for the band comparison are displayed in panel (c) for the path W–L–K, and in panel (d) for the path L–Γ.

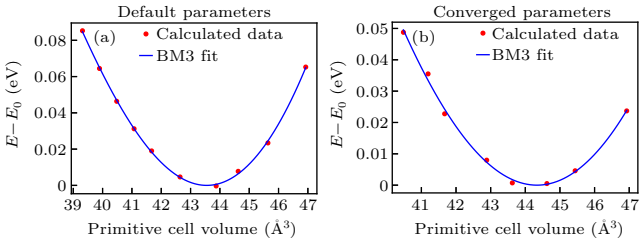


Fig. 8. Lattice constant optimization of GaAs: (a) calculation with default parameters, and (b) with converged parameters.

Taken together, these studies involving structural perturbation, strain, elemental variation, and lattice constant optimization sufficiently demonstrate that the optimized parameters obtained from our machine learning approach possess strong generalization capability.

5. Conclusion

In this work, we employed a machine learning approach, specifically the CMA-ES stochastic algorithm, to optimize the selection of NMTO parameters. This approach enables the MTO method to achieve computational accuracy comparable to that of the plane-wave calculation method. Our study investigates four materials and demonstrates overall improvements near the Fermi level, which is the most critical region. Notably, significant enhancements are observed in the unoccupied states, which are also crucial for various physical processes.[40–42] Furthermore, we demonstrate the generalization capability of the machine learning approach by employing perturbation, strain, elemental substitution, and lattice constant optimization. Looking ahead, we anticipate the continued reliability of this machine learning method and plan to extend it to more complex materials and transport research, further maximizing the computational capabilities of the MTO method. Furthermore, this approach can also be applied to the parameter tuning of other localized basis sets if desired.

Acknowledgment

This project was supported by the National Key Research and Development Program of China (Grant Nos. 2023YFA1406600 and 2021YFA1202200).

References

[1] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
[2] Burke K and Wagne L O 2013 *Int. J. Quantum Chem.* **113** 96
[3] Perdew J P, Ruzsinszky A, Tao J, Staroverov V N, Scuseria G E and Csonka G I 2005 *J. Chem. Phys.* **123** 062201
[4] Chachiyo T 2016 *J. Chem. Phys.* **145** 021101
[5] Jitropas U and Hsu C H 2017 *Jpn. J. Appl. Phys.* **56** 070313
[6] Perdew J P, Chevary J A, Vosko S H, Jackson K A, Pederson M R, Singh D J and Fiolhais C 1992 *Phys. Rev. B* **46** 6671
[7] Becke A D 1988 *Phys. Rev. A* **38** 3098
[8] Langreth D C and Mehl M J 1983 *Phys. Rev. B* **28** 1809
[9] Topp W C and Hopfield J J 1973 *Phys. Rev. B* **7** 1295

[10] Kresse G and Furthmüller J 1996 *Phys. Rev. B* **54** 11169
[11] Kresse G and Hafner J 1993 *Phys. Rev. B* **47** 558
[12] Blaha P, Schwarz K, Tran F, Laskowski R, Madsen G K H and Marks L D 2020 *J. Chem. Phys.* **152** 074101
[13] Soler J M, Artacho E, Gale J D, García A, Junquera J, Ordejón P and Sánchez-Portal D 2002 *J. Phys.: Condens. Matter* **14** 2745
[14] Artacho E, Anglada E, Diéguez O, Gale J D, García A, Junquera J, Martin R M, Ordejón P, Pruneda J M, Sánchez-Portal D and Soler J M 2008 *J. Phys.: Condens. Matter* **20** 064208
[15] García A, Papior N, Akhtar A, *et al*. 2020 *J. Chem. Phys.* **152** 204108
[16] Ozaki T 2003 *Phys. Rev. B* **67** 155108
[17] Ozaki T and Kino H 2004 *Phys. Rev. B* **69** 195113
[18] Ozaki T and Kino H 2005 *Phys. Rev. B* **72** 045121
[19] Chen M H, Guo G C and He L X, 2010 *J. Phys. Condens. Matter* **22** 44550
[20] Li P F, Liu X H, Chen M H, Lin P Z, Ren X G, Lin L, Yang C and He L X 2016 *Comput. Mater. Sci.* **112** 503
[21] Kühne T D, Annuzzi M, Ben M D, *et al*. 2020 *J. Chem. Phys.* **152** 194103
[22] Andersen O K 2012 2012 *Correlated Electrons: From Models to Materials* 2nd edn. (Germany: Forschungszentrum Jülich) pp. 55–84
[23] Zurek E, Jepsen O and Andersen O K 2005 arXiv:cond-mat/0504374 [cond-mat.str-el]
[24] Andersen O K 1975 *Phys. Rev. B* **12** 3060
[25] Andersen O K, Jepsen O and Krier G 1994 *Lectures on Methods of Electronic Structure Calculations* (Singapore: World Scientific Publishing Co) pp. 63–124
[26] Ando T 1991 *Phys. Rev. B* **44** 8017
[27] Xia K, Zwierzycki M, Talanana M, Kelly P J and Bauer G E W 2006 *Phys. Rev. B* **73** 064420
[28] Wesselink R J H, Gupta K, Yuan Z and Kelly P J 2019 *Phys. Rev. B* **99** 144409
[29] Starikov A A, Liu Y, Yuan Z and Kelly P J 2018 *Phys. Rev. B* **97** 214415
[30] Wang L, Min T and Xia K 2021 *Phys. Rev. B* **103** 054204
[31] Nomura M and Shibata M 2024 arXiv:2402.01373 [cs.NE]
[32] Akiba T, Sano S, Yanase T, Ohta T and Koyama M 2019 *in Proceedings of the* 25*th ACM SIGKDD Int. Conf. Knowledge Discovery & Data Mining* KDD'19, ACM, New York, NY, USA, pp. 2623–2631
[33] Blöchl P E 1994 *Phys. Rev. B* **50** 17953
[34] Kresse G and Joubert D 1999 *Phys. Rev. B* **59** 1758
[35] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
[36] Liu Y, Kwon S, de Coster G J, Lake R K and Neupane M R 2022 *Phys. Rev. Mater.* **6** 084004
[37] Mcguire M A, Dixit H, Cooper V R and Sales B C 2014 *Chem. Mater.* **27** 612
[38] von Barth U and Hedin L 1972 *J. Phys. C: Solid State Phys.* **5** 1629
[39] Katsura T and Tange Y 2019 *Minerals* **9** 745
[40] Dufek P, Schwarz K and Blaha P 1993 *Phys. Rev. B* **48** 12672
[41] Drittler B, Weinert M, Zeller R and Dederichs P H 1990 *Phys. Rev. B* **42** 9336
[42] Zhang Y C 2017 *J. Chem. Phys.* **146** 194705