

DeePKS Model for Halide Perovskites with the Accuracy of a Hybrid Functional

Qi Ou,* Ping Tuo, Wenfei Li, Xiaoxu Wang, Yixiao Chen, and Linfeng Zhang*



Cite This: *J. Phys. Chem. C* 2023, 127, 18755–18764



Read Online

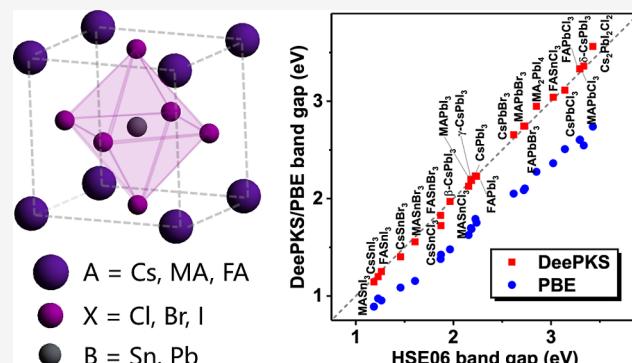
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Accurate prediction of the electronic structure properties of halide perovskites plays a significant role in the design of highly efficient and stable solar cells. While density functional theory (DFT) within the generalized gradient approximation (GGA) offers reliable prediction in terms of lattice constants and potential energy surface for halide perovskites, it severely underestimates the band gap due to the lack of a non-local exact exchange term, which exists in computationally expensive hybrid functionals. In this work, a universal Deep Kohn–Sham (DeePKS) model based on a neural network is trained so as to enable electronic structure calculations with the accuracy of the hybrid functional HSE06 and efficiency comparable to the GGA functional for a plethora of halide perovskites, i.e., ABX_3 ($A = \text{FA, MA, Cs}$; $B = \text{Sn, Pb}$; $X = \text{Cl, Br, I}$), as well as two types of Ruddlesden–Popper perovskites. Forces, band gaps, and density of states (DOS) predicted by our DeePKS model for all aforementioned perovskites are in good agreement with the HSE06 results, with significantly improved efficiency. In addition, even though the spin–orbit coupling (SOC) effect has not been taken into consideration during the training process, DeePKS + SOC offers a highly consistent band gap and DOS as compared to HSE06 + SOC for Pb-containing systems. We believe that such a DeePKS model can be readily applied for an accurate yet efficient prediction of various properties for the family of halide perovskites.



INTRODUCTION

Owing to the unsustainable nature of fossil-based energy sources, new forms of energy that are both renewable and environmentally friendly have been comprehensively investigated worldwide.^{1–5} Solar cells, or photovoltaics, are regarded as one of the promising alternatives to traditional fossil fuel resources.^{6–9} While the silicon-based solar cells, i.e., the first- and second-generation photovoltaics, exhibit good performance in terms of stability and efficiency,^{10–13} the cost of ultrahigh-pure metallic silicon somewhat hinders their application. Third-generation photovoltaics, including organic and dye-sensitized solar cells, have been developed later on, with significantly reduced processing costs.^{14–17} Nevertheless, the power conversion efficiency (PCE) of the third-generation photovoltaics remains less satisfactory compared with the silicon-based modules,^{18–20} which limits their chances for commercialization.

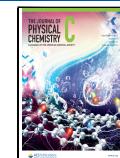
First synthesized in the 1890s,²¹ halide perovskites (ABX_3 , $X = \text{halogen anion}$) have shown immense potential to become a low-cost alternative to the currently commercialized photovoltaic technologies owing to their desired semiconducting properties and economic fabrication.^{20,22–25} In addition to the relatively high PCE reached by halide perovskites, one of their favored traits that is rarely possessed by other conventional semiconductors is the tunable absorption edge wavelength

(band gap),^{26–28} which can be straightforwardly realized by varying the ratios of different halide ions. An accurate theoretical description of the electronic structure for halide perovskites is a long-standing goal for the realistic design of photovoltaics with higher efficiency and better stability. Density functional theory (DFT) with the generalized gradient approximation (GGA) combined with D3 dispersion correction has been extensively applied in previous studies, offering a balanced trade-off between accuracy and efficiency in predicting properties such as lattice constants and potential energy surfaces for various perovskites.^{29–34} Nevertheless, such methodology suffers a severe underestimation of the band gaps, which can only be correctly captured via either the hybrid functional that incorporates a portion of exact exchange from Hartree–Fock theory or the many-body perturbation theory such as the GW method.^{34–38} Unfortunately, both these

Received: July 12, 2023

Revised: August 21, 2023

Published: September 11, 2023



methods are computationally much more expensive compared to GGA and even become prohibitive for large systems.

With the development of massive computational resources and advanced machine learning (ML) algorithms, efficient yet accurate descriptions of the electronic structure for large systems have been made possible. Proposed in 2020, the Deep Kohn–Sham (DeePKS) approach introduces a general framework for generating highly accurate self-consistent energy functionals with remarkably reduced computational cost,^{39–41} in which the energy difference between the expensive high-level method and the cheap low-level method is fitted by a neural network. While DeePKS has been comprehensively validated on molecular systems and specific condensed phase systems,^{40,41} its generalizability to multiple or a family of systems, though theoretically feasible, remains unexplored. Indeed, it would be essential to construct a general DeePKS model that can offer highly accurate electronic structure results with low computational cost for systems like halide perovskites, given their undeniable significance in photovoltaic and other fields.

In this work, our goal is to establish a general DeePKS model that can compute the electronic structure properties with the accuracy of the hybrid functional HSE06⁴² and the efficiency comparable to the non-hybrid GGA functional for various halide perovskites, i.e., ABX_3 ($A = \text{FA}, \text{MA}, \text{Cs}$; $B = \text{Sn}, \text{Pb}$; $X = \text{Cl}, \text{Br}, \text{I}$) as well as the organic–inorganic hybrid alternatives and the Ruddlesden–Popper (RP) perovskites. Based on an iterative training process with 460 configurations over seven types of halide perovskites, we demonstrate that the resulting DeePKS model is capable of reproducing closely matched forces, stress, band gaps, and density of states (DOS) near the Fermi energy as compared to HSE06 for any constitution of the aforementioned halide perovskites (including the RP structures) as well as the hybrid ones. In addition, even though the spin–orbit coupling (SOC) effect has not been taken into consideration during the training process, our computational results show that DeePKS + SOC offers a highly consistent band gap and DOS as compared to HSE06 + SOC for Pb-containing systems, in which the SOC effect is necessary for a quantitatively correct description of the band gap.⁴³ The equation of state (EoS) and corresponding band gap predictions with respect to varied sizes of the unit cell are also validated. We believe such DeePKS model can be readily applied for an accurate yet efficient prediction of various properties for the family of halide perovskites and potentially benefit the discovery and design of better photovoltaics based on such materials.

METHODS

The DeePKS algorithm for periodic systems has been previously implemented in an open-source DFT package, ABACUS on the basis of numerical atomic orbitals (NAO),^{44,45} of which the theoretical details can be found in ref 41. Extended labels for the DeePKS training process implemented for this work are stress and band gap, which enable the converged DeePKS model to reproduce the stress and DOS near the Fermi energy given by the target method. The stress given by the DeePKS model can be calculated via

$$\begin{aligned} \sigma_{\alpha\beta}^{\text{DeePKS}}[\{\phi|\omega\}] &= \sigma_{\alpha\beta}^{\text{baseline}}[\{\phi|\omega\}] - \frac{\partial E^\delta[\{\phi|\omega\}, \omega]}{\partial \epsilon_{\alpha\beta}} \\ &= \sigma_{\alpha\beta}^{\text{baseline}}[\{\phi|\omega\}] \\ &\quad - \sum_{Inlm'm'} \frac{\partial E_\delta}{\partial D_{nlm'm'}^I} \sum_i \frac{d}{d\epsilon_{\alpha\beta}} [f_i \langle \phi_i | \alpha_{nlm}^I \rangle \\ &\quad \langle \alpha_{nlm}^I | \phi_i \rangle] \end{aligned} \quad (1)$$

where f_i is the occupation number of orbital ϕ_i . Note that the stress label for the DeePKS training process is of the energy unit, i.e., without dividing by the volume of the unit cell, and only the upper triangle parts of the stress tensor serve as labels. For band gap labels, the energy differences between the highest valence band (HVB) and the lowest conducting band (LCB) at each k point

$$\epsilon_{g,k} := \epsilon_{\text{LCB},k} - \epsilon_{\text{HVB},k} \quad (2)$$

are served as labels for the training process, aiming to provide an accurate description of the DOS near Fermi energy. Total energies, forces, stresses, and band gaps of each k point of the target method are included in the loss function $L(\omega)$, and the optimization problem now becomes

$$\begin{aligned} \min_\omega L(\omega), L(\omega) &= |E^{\text{target}} - E^{\text{DeePKS}}[\{\phi|\omega\}]|^2 \\ &\quad + \lambda_1 |\mathbf{F}^{\text{target}} - \mathbf{F}^{\text{DeePKS}}[\{\phi|\omega\}]|^2 \\ &\quad + \lambda_2 |\epsilon_g^{\text{target}} - \epsilon_g^{\text{DeePKS}}[\{\phi|\omega\}]|^2 \\ &\quad + \lambda_3 |\sigma^{\text{target}} - \sigma^{\text{DeePKS}}[\{\phi|\omega\}]|^2 \end{aligned} \quad (3)$$

where λ_1 , λ_2 , and λ_3 correspond to the weighting factors of force, band gap, and stress in the loss function and are set to be 1.0, 0.1, and 1.0 in this work, respectively.

It should be noted that the goal of this work is to contrive a general DeePKS model for various halide perovskites instead of a single system. This poses challenges to the energy fitting process by the neural network owing to the fact that the absolute energy difference between the base and the target method contributed from each element of multiple systems might be significantly variant. To circumvent this obstacle, we apply the method of linear least squares to calculate the element-specific contribution to the absolute energy difference between the base and the target method. Specifically, the sum of squares to be minimized in our case is

$$S = \sum_s \left[\Delta E_s - \sum_I \epsilon_I N_{I,s} \right]^2 \quad (4)$$

where ΔE_s is the energy difference between the base and target method of system s , and $N_{I,s}$ is the number of atoms of element I contained in system s . The minimization of S gives ϵ_I , i.e., the contribution of element I to the energy difference between the base and target method, and the absolute energy difference resulting from ϵ_I for each system s , i.e., $\sum_I \epsilon_I N_{I,s}$ is then deducted from the loss function. ϵ_I of each element involved in the investigated systems is listed in Table S1.

The training sets of the DeePKS model consist of FAPbI_3 , MAPbI_3 , CsPbI_3 , MAPbBr_3 , MASnCl_3 , CsPbCl_3 , and CsSnBr_3 , and the number of configurations of each system is given in Table S2. Configurations of the aforementioned XPbI_3 family

are randomly (subject to uniform distribution) picked from previously reported DeePMD training data,⁴⁶ while those of other systems are randomly picked from corresponding DP-GEN jobs.⁴⁷ The test sets consist of four systems that are different from the training sets, i.e., MAPbCl₃, MASnBr₃, CsSnI₃, and FA_{0.125}Cs_{0.875}PbI₃. Configurations of FA_{0.125}Cs_{0.875}PbI₃ are randomly picked from previously reported DeePMD training data⁴⁶ while those of the other three are randomly picked from the corresponding DP-GEN jobs.

The target functional applied in this work for label generation is the hybrid GGA functional HSE06. Self-consistent field (SCF) jobs for all configurations in the training test sets are carried out with the HSE06 functional and projector-augmented wave (PAW) method in the Vienna ab initio simulation package (VASP).^{48–50} The energy cutoff is set to be 500 eV, and the allowed spacing between k points (“kspacing”) is set to be 0.1 Bohr⁻¹. Such an energy cutoff has been previously validated in refs 46 and 51 for XPbI₃ (X = Cs, FA, MA) systems, and the kspacing value for the HSE06 jobs is slightly larger compared to that applied in ref 46 for GGA calculation (i.e., 0.16 Å⁻¹) due to the high computational cost of the exact exchange term. We will demonstrate in the next section that the DeePKS model trained with a relatively sparse k -point label can successfully recover the energy, force, and DOS of the target method with a much denser k -point grid. During the DeePKS iterative training process, the single-point SCF calculations are performed in ABACUS^{44,45} with double-zeta polarized (DZP) NAO basis,^{52,53} 100 Ry energy cutoff, 0.1 Bohr⁻¹ kspacing value, and the SG15 optimized norm-conserving Vanderbilt (ONCV)⁵⁴ pseudopotentials. The base functional applied in this work is the Perdew–Burke–Ernzerhof (PBE) GGA functional.⁵⁵ More details of the DeePKS training process can be found in the Supporting Information.

RESULTS AND DISCUSSION

Accuracy Cross-Checking of the DeePKS Model. Mean absolute errors (MAEs) of energies, forces, band gaps for multi- k points, and stresses of the training set given by the converged DeePKS model are listed in Table S2, while those for the test set are listed in Table S3. Note that the absolute energies shift significantly between the HSE06 calculation and the DeePKS/PBE calculation due to the fact that PAW is employed for HSE06 jobs, while ONCV pseudopotentials are applied in DeePKS/PBE. We exclude the influence of such a “constant” shift by defining the mean absolute relative error (MARE) of the energy as

$$E_{\text{MARE}} = \frac{1}{N} \sum_{i=1}^N |E_i^{\text{HSE06}} - [E_i^{\text{DeePKS/PBE}} - \bar{E}^{\text{DeePKS/PBE}} + \bar{E}^{\text{HSE06}}]| \quad (5)$$

where \bar{E} refers to the average energy value of a training system given by the corresponding method. It can be seen that compared to PBE results, the MAE/MARE of the four labels given by DeePKS is significantly reduced, especially that of band gaps and stresses, and similar accuracy is achieved via DeePKS for the test sets. As shown in Table S3, for the four tested systems, a pronounced decrease of the MAE is observed for the band gap and stress (0.053 and 0.123 eV, respectively) compared to the PBE results (0.825 and 0.657 eV,

respectively), while the MAE/MARE for energy and force (2.2 meV/atom, 0.045 eV/Å, respectively) is also decreased by around a factor of 2.

It should be noted that for halide perovskite systems, PBE has been validated as a reliable functional to describe the potential energy surfaces.^{29–34,46} The key problem encountered by PBE is indeed the prediction of the band gap for these systems, which is essential for photovoltaic properties. HSE06, albeit computationally demanding, is able to offer a significantly improved description of the band gap for perovskites. The successful recovery of the HSE06 band gap by DeePKS evinces its capability to offer a notably more accurate description of the electronic structure properties as compared to PBE. The reason that the stress MAE of DeePKS remains relatively large (>0.1 eV) is attributed to the severe deviation between the PBE-predicted stress and the HSE06-predicted one (with an overall MAE larger than 1.3 eV for the training set). That being said, the MAE of the stress predicted by the DeePKS model decreases remarkably as compared to that given by PBE, indicating the strength of the training process.

Property Prediction on Various Halide Perovskite Systems. We extensively test our DeePKS model on 18 different types of non-hybrid perovskites (mostly cubic phase unless otherwise specified), two RP perovskites, i.e., Cs₂PbI₂Cl₂⁵⁶ and MA₂PbI₄,⁵⁷ and the hybrid organic–inorganic perovskite FA_aCs_{1.0-a}PbI₃ with varied a values. Cell relaxation is performed for all non-hybrid perovskites, while ionic relaxation is performed for hybrid perovskites. All relaxation calculations are carried out with PBE functional and D3_0 van der Waals (vdW) dispersion correction in ABACUS. The DZP NAO basis set with ONCV pseudopotentials is applied, and the energy cutoff is set to 100 Ry. For all cubic-phase systems, the Brillouin zone is sampled with a gamma-centered 9 × 9 × 9 k -point grid, while for all other systems, the k -spacing value is set to be 0.07 Bohr⁻¹. DeePKS/PBE and HSE06 single-point SCF calculations are performed on these optimized structures with the same basis sets and energy cutoff as those applied in the iterative training and label generation processes, while the k -point settings are consistent with the aforementioned relaxation calculations. Detailed input parameters of HSE06 (carried out in VASP) and DeePKS/PBE (carried out in ABACUS) are provided in the Supporting Information.

The total CPU times of the SCF calculations performed via HSE06, DeePKS, and PBE for these tested perovskites are given in Table S4. It can be seen that the CPU time of DeePKS is of the same order of magnitude (though twice to three times as long) as compared to that of PBE for most tested systems, while it is around 2 orders of magnitude faster as compared to that of HSE06 based on our input parameters and parallelization settings. In addition to the omission of constructing the Hartree–Fock exact exchange part, such huge savings in computational time of DeePKS with respect to HSE06 also stem from the higher efficiency of the NAO basis as compared to the plane-wave basis for large systems.

Force and Stress. While the absolute energies between DeePKS and HSE06 are not directly comparable owing to the discrepancy between SG15 ONCV pseudopotential applied in DeePKS and the PAW method applied in HSE06, we first investigate the accuracy of our DeePKS model in calculating the forces of the selected systems. Forces are tested on all systems but cubic-phase inorganic perovskites, of which the

atomic forces remain zero due to their ideal cubic symmetry ($Pm\text{-}3m$ space group). The MAE of the force predicted by DeePKS and PBE as compared to the HSE06 result for each tested system is listed in **Table 1**. Forces along three directions

Table 1. Force MAE for All Tested Halide Perovskites Given by DeePKS and PBE with Respect to the HSE06 Results^a

system	DeePKS	PBE
$\beta\text{-CsPbI}_3$	0.0020	0.0054
$\gamma\text{-CsPbI}_3$	0.0025	0.0053
$\delta\text{-CsPbI}_3$	0.0086	0.0268
MAPbCl ₃	0.0245	0.0765
MAPbBr ₃	0.0211	0.0779
MAPbI ₃	0.0210	0.0800
MASnCl ₃	0.0265	0.0888
MASnBr ₃	0.0219	0.0833
MASnI ₃	0.0317	0.0895
FAPbCl ₃	0.0535	0.0897
FAPbBr ₃	0.0381	0.0887
FAPbI ₃	0.0408	0.0927
FASnCl ₃	0.0453	0.0897
FASnBr ₃	0.0370	0.1053
FASnI ₃	0.0447	0.0942
MA ₂ PbI ₄	0.0419	0.0984
FA _{0.125} Cs _{0.875} PbI ₃	0.0291	0.0445
FA _{0.25} Cs _{0.75} PbI ₃	0.0281	0.0540
FA _{0.375} Cs _{0.625} PbI ₃	0.0313	0.0620
FA _{0.5} Cs _{0.5} PbI ₃	0.0347	0.0782
FA _{0.625} Cs _{0.375} PbI ₃	0.0383	0.0817
FA _{0.75} Cs _{0.25} PbI ₃	0.0372	0.0909
FA _{0.875} Cs _{0.125} PbI ₃	0.0434	0.0905
overall MAE	0.0306	0.0737

^aAll tested ABX₃ systems are cubic phase except for those indicated by Greek letters. Numbers are shown in the unit of eV/Å.

of hybrid perovskite systems given by DeePKS and PBE are plotted with respect to the HSE06 values in **Figure 1**. It can be seen that while the PBE forces are already close to the HSE06 results, forces calculated by DeePKS are notably more concentrated on the diagonal. The overall MAE of the force given by DeePKS for all tested systems is around 0.031 eV/Å, while that given by PBE is around 0.074 eV/Å, indicating an

improved accuracy in the description of the potential energy surface with sharply higher efficiency as mentioned above.

Compared to the force label, the target stress given by HSE06 is more difficult to reproduce by DeePKS due to the notable deviation between the PBE and HSE06 stresses as mentioned above. Here, we calculate the stress for those hybrid perovskites via DeePKS and PBE, and results are compared with those given by HSE06 as shown in **Figure 2** (while explicit numbers can be found in **Table S5**). It can be seen that the diagonal element of the stress calculated by PBE is significantly larger than the HSE06 counterparts, while the off-diagonal elements are severely underestimated. The stress calculated by DeePKS is qualitatively in line with the HSE06 result, with the remarkably reduced MAE for both diagonal and off-diagonal elements of the stress, i.e., 0.394 and 0.168 KBar, respectively (as compared to 5.435 and 1.556 KBar given by PBE).

Band Gaps and Density of States. Band gaps predicted by DeePKS for all tested systems are compared with the results of HSE06 in **Table 2**. Values computed via the PBE functional are also listed for reference. As shown in **Table 2**, the DeePKS predicted band gaps for all tested perovskites (including the hybrid systems, the non-perovskite phase, i.e., $\delta\text{-CsPbI}_3$, and the RP perovskites) are in excellent agreement with the HSE06 results, with an overall MAE of 0.0350 eV. The well-known underestimation of the band gap values suffered by the PBE functional (indicated by the large overall MAE, i.e., 0.5222 eV) is effectively surmounted by the DeePKS model. We visualize the band gap results of DeePKS and PBE with respect to the HSE06 benchmarks in **Figure 3**. While the PBE-predicted values lie significantly lower, the DeePKS results are concentrated on the diagonal line, demonstrating a remarkably improved accuracy for the band gap description obtained by the DeePKS model.

Next, we examine the DOS of various halide perovskites described by DeePKS. Even though the DOS does not directly serve as one of the training labels to be fitted by the neural network, we can see from **Figures 4, 5, 6** and **7** that for all tested halide perovskites, the overall DOSs near the Fermi energy predicted by DeePKS precisely match those given by HSE06, while those of PBE exhibit significantly smaller band gaps. Such an accurate description of the DOS given by DeePKS lies in the fact that bands predicted by PBE near the Fermi energy possess an overall reasonable structure (as speculated by the similar shape of the DOS predicted by HSE06 and PBE) but with a much narrower gap. By fitting the

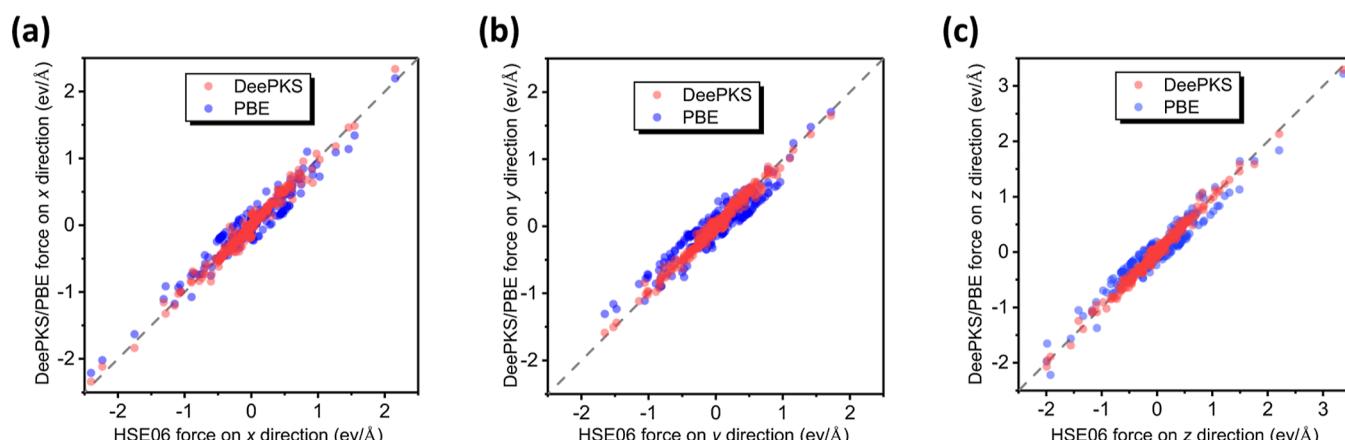


Figure 1. Forces along the three directions of hybrid perovskite systems given by DeePKS and PBE with respect to the HSE06 results.

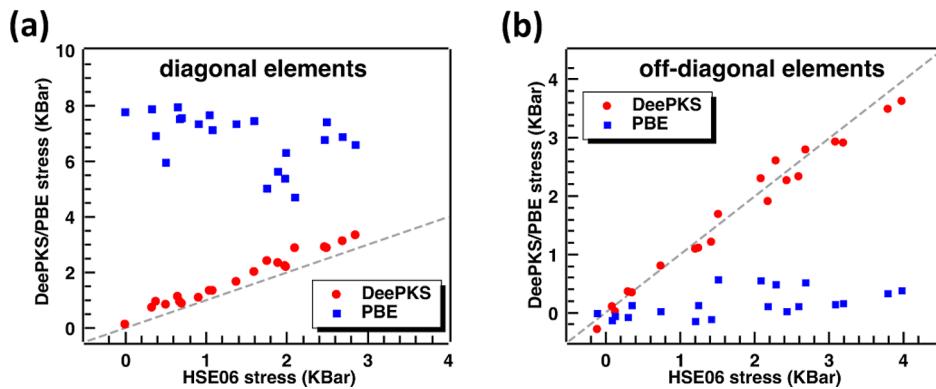


Figure 2. Stresses (diagonal and off-diagonal elements) of hybrid perovskite systems given by DeePKS and PBE with respect to the HSE06 results.

Table 2. Band Gaps for All Tested Halide Perovskites Given by HSE06, DeePKS, and PBE^a

system	HSE06	DeePKS	DeePKS Abs. Err	PBE	PBE Abs. Err
CsPbCl ₃	3.1423	3.1135	0.0288	2.5079	0.6344
CsPbBr ₃	2.6174	2.6551	0.0377	2.0502	0.5672
CsPbI ₃	2.2254	2.2331	0.0077	1.7906	0.4348
β -CsPbI ₃	1.9650	1.9703	0.0053	1.4782	0.4868
γ -CsPbI ₃	2.1854	2.1861	0.0007	1.6873	0.4981
δ -CsPbI ₃	3.3353	3.3603	0.0250	2.5454	0.7899
Cs ₂ PbI ₂ Cl ₂	3.4287	3.5629	0.1342	2.7385	0.6902
CsSnCl ₃	1.8727	1.7226	0.1501	1.4237	0.4490
CsSnBr ₃	1.4584	1.3750	0.0834	1.0856	0.3728
CsSnI ₃	1.2288	1.1979	0.0309	0.9730	0.2558
MAPbCl ₃	3.2972	3.3318	0.0346	2.6055	0.6917
MAPbBr ₃	2.7211	2.7439	0.0228	2.0861	0.6350
MAPbI ₃	2.1789	2.1981	0.0192	1.6966	0.4823
MA ₂ PbI ₄	2.8506	2.9467	0.0961	2.2748	0.5758
MASnCl ₃	2.1558	2.1264	0.0294	1.6259	0.5299
MASnBr ₃	1.6072	1.5557	0.0515	1.1542	0.4530
MASnI ₃	1.1858	1.1443	0.0415	0.8897	0.2961
FAPbCl ₃	3.2938	3.3319	0.0381	2.6024	0.6914
FAPbBr ₃	2.7325	2.7443	0.0118	2.1036	0.6289
FAPbI ₃	2.2374	2.2284	0.0090	1.7521	0.4853
FASnCl ₃	3.0226	3.0412	0.0186	2.3624	0.6602
FASnBr ₃	1.8691	1.8273	0.0418	1.3795	0.4896
FASnI ₃	1.2622	1.2502	0.0120	0.9536	0.3086
FA _{0.125} Cs _{0.875} PbI ₃	2.3968	2.3805	0.0163	1.8556	0.5412
FA _{0.25} Cs _{0.75} PbI ₃	2.3929	2.3952	0.0023	1.8499	0.5430
FA _{0.375} Cs _{0.625} PbI ₃	2.0071	1.9875	0.0196	1.5313	0.4758
FA _{0.5} Cs _{0.5} PbI ₃	2.3547	2.3428	0.0119	1.8072	0.5475
FA _{0.625} Cs _{0.375} PbI ₃	2.0222	2.0061	0.0161	1.5499	0.4723
FA _{0.75} Cs _{0.25} PbI ₃	2.0145	1.9997	0.0148	1.5411	0.4734
FA _{0.875} Cs _{0.125} PbI ₃	2.0808	2.0429	0.0379	1.5760	0.5048
overall MAE	0.0350		0.5222		

^aAbsolute errors of the DeePKS and PBE results with respect to the HSE06 results are listed, as well as the overall MAEs for all tested systems. All tested ABX₃ systems are cubic phases except for those indicated by Greek letters. Numbers are shown in the unit of eV.

band gap for each k point, such deficiency of PBE can be largely mitigated.

Spin–Orbit Coupling Effect on Pb-Containing Systems. It is well known that the SOC effect plays an important role in determining the magnitude of the band gaps in perovskite systems,^{36,58} especially for Pb-containing perovskite.^{43,59} The inclusion of the SOC effect may significantly

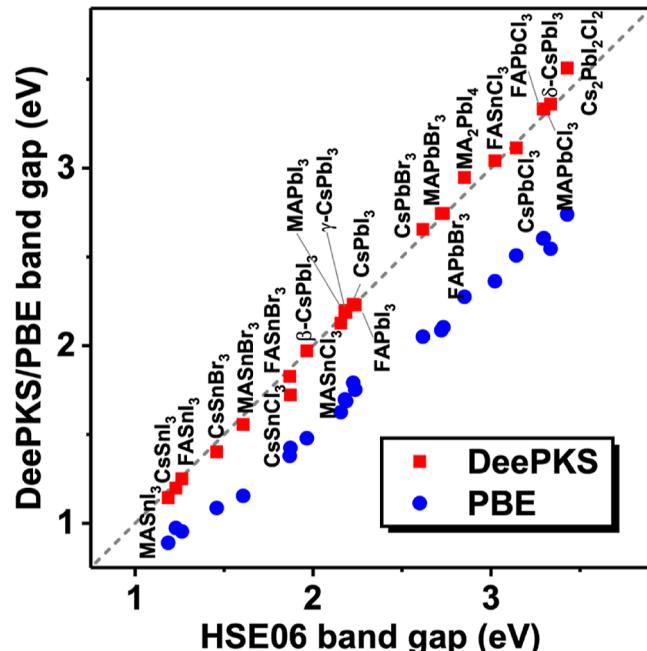


Figure 3. Band gaps predicted by DeePKS and PBE with respect to the HSE06 results for all tested systems. All tested ABX₃ systems are cubic phase except for those indicated by Greek phase.

narrow the band gap for APbX₃ and thus become essential for accurate property prediction in practice.^{43,60} Taking CsPbX₃ as an example, we investigate the performance of the DeePKS model with the inclusion of the SOC effect. To do that, the full relativistic (FR) version of the SG15 ONCV pseudopotentials⁶¹ is applied for all involved elements. As shown in Table 3 and Figure 8, while the SOC effect has not been taken into account during our iterative DeePKS training process, the DeePKS + SOC offers closely matched band gap values as well as the DOS for three tested systems as compared to HSE06 + SOC. Indeed, the relativistic effect is mainly handled in the pseudopotential part, either via the PAW method (for HSE06 functional performed in VASP) or via the norm-conserving pseudopotentials (for DeePKS model performed in ABACUS). Even though the scalar relativistic approximation is applied in PAW for the valence electron⁶² while the FR effect is employed in ONCV pseudopotentials, the influence of such discrepancy is subtle for our tested systems, as demonstrated by the almost identical DOS given by PBE + SOC calculations performed in VASP with PAW and that in ABACUS with NAO and the FR

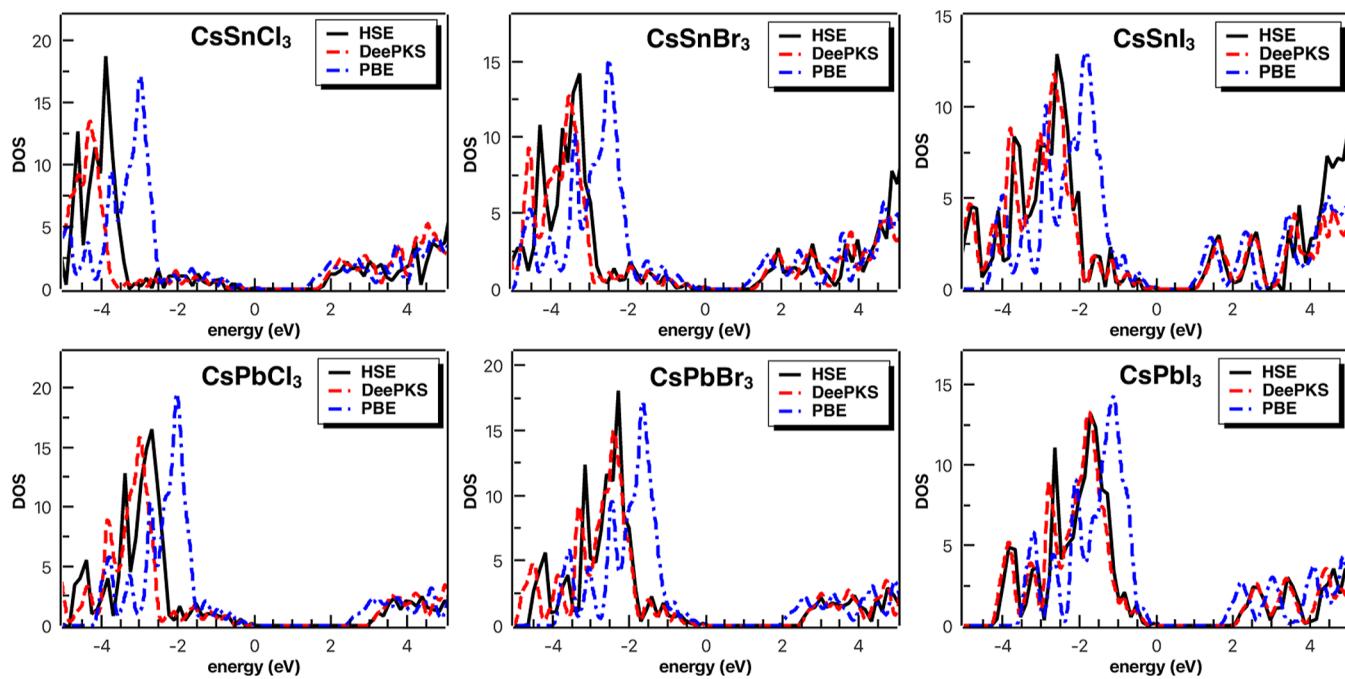


Figure 4. DOS for cubic-phase inorganic halide perovskites, i.e., CsBX_3 ($B = \text{Sn}, \text{Pb}; X = \text{Cl}, \text{Br}, \text{I}$).

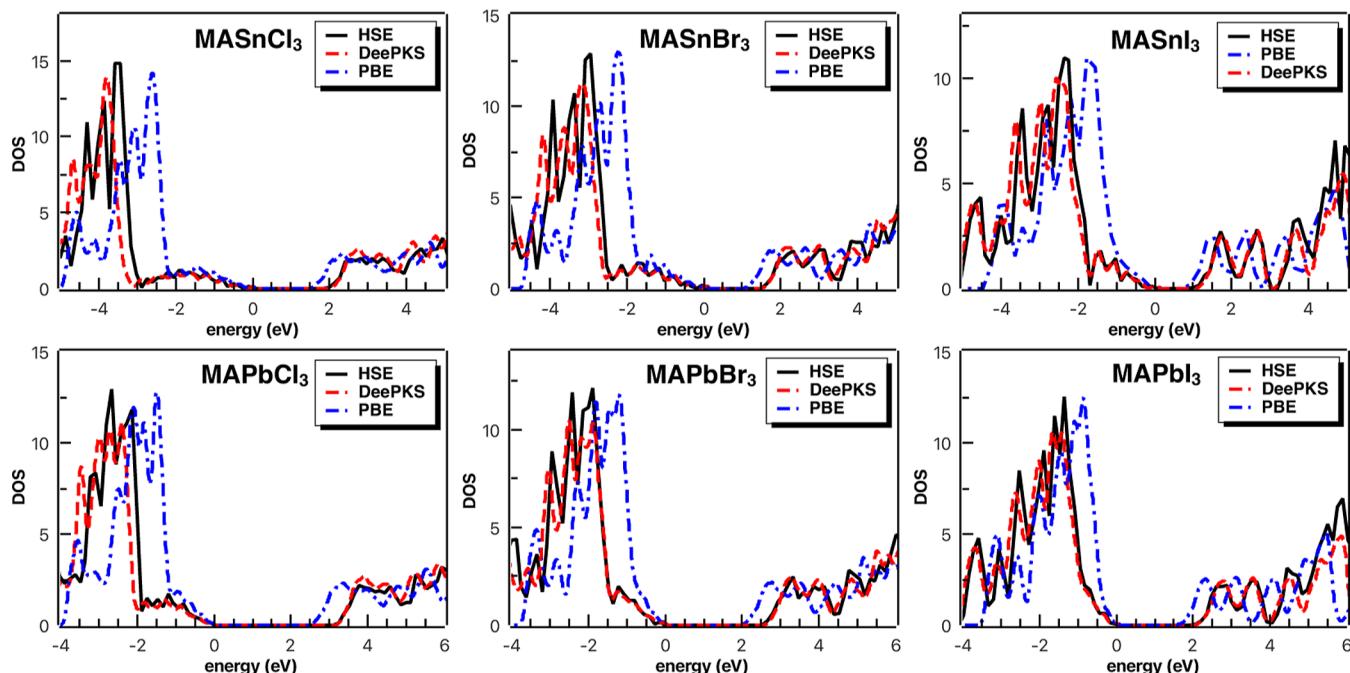


Figure 5. DOS for cubic-phase organic halide perovskites, i.e., MABX_3 ($B = \text{Sn}, \text{Pb}; X = \text{Cl}, \text{Br}, \text{I}$).

version of ONCV pseudopotentials (Figure S1). Therefore, the HSE06 + SOC results can be successfully reproduced by DeePKS + SOC as long as the DeePKS model provides a close enough description of the electronic structure as compared to the HSE06 functional, which has been extensively evinced in the former sections.

Equation of State for Cubic Inorganic Perovskites.

Finally, we examine the capability of our DeePKS model in predicting mechanical properties by performing the EoS calculation for three cubic-phase inorganic perovskites, i.e., CsPbCl_3 , CsPbBr_3 , and CsPbI_3 . A cell relaxation is first performed with the DeePKS model and PBE functional (with

the same input settings as those applied in the former single-point calculation). The resulting optimized volumes are then scaled from 0.9 to 1.1 with a step of 0.01 to compute the total energy at different volumes. Due to the extremely high computational cost of HSE06 functional, we skip the relaxation step by applying the DeePKS-optimized structure at each volume and perform a single-point SCF calculation via HSE06 functional. Since the absolute energies given by ONCV pseudopotentials and the PAW method are not directly comparable, we plot the relative energies with respect to the lowest energy for each functional. As shown in the upper panel of Figure 9, the DeePKS computed EoS for these systems is in

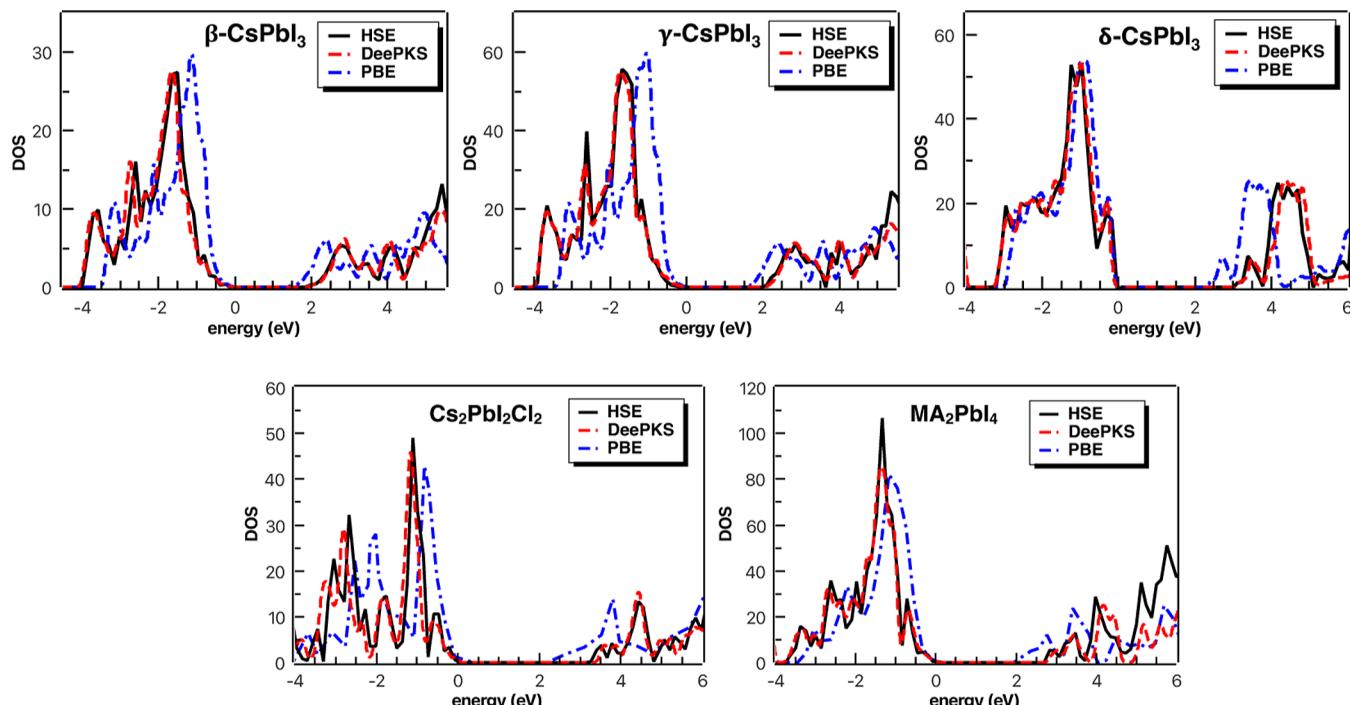


Figure 6. DOS for non-cubic-phase CsPbI_3 and two Ruddlesden–Popper perovskites, i.e., $\text{Cs}_2\text{PbI}_2\text{Cl}_2$ and MA_2PbI_4 .

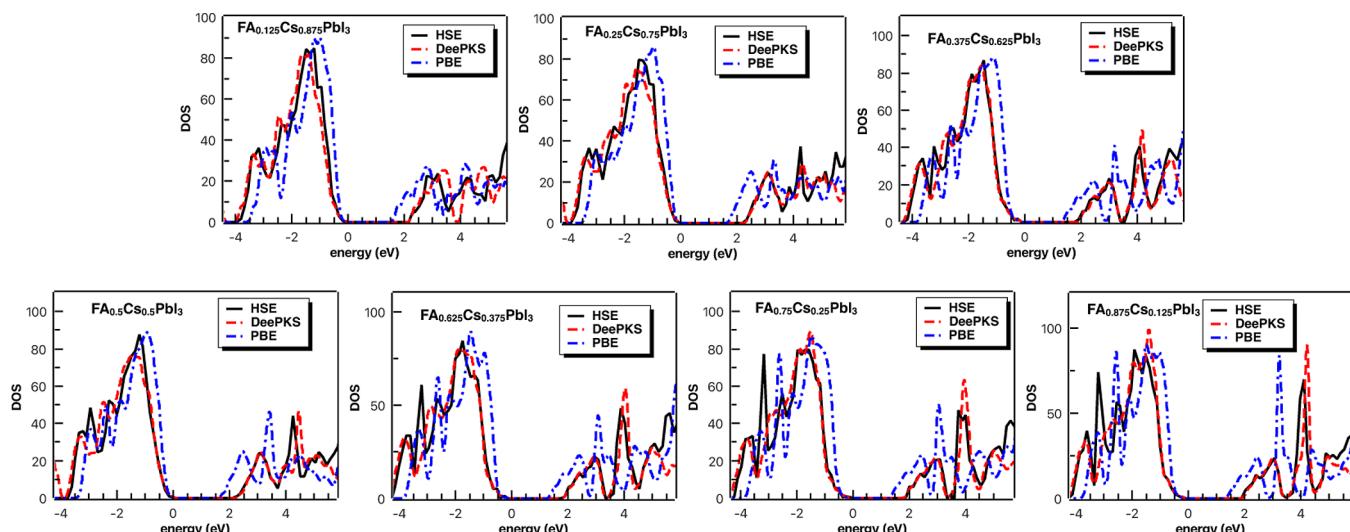


Figure 7. DOS for hybrid organic–inorganic halide perovskites, i.e., $\text{FA}_a\text{Cs}_{1.0-a}\text{PbI}_3$.

Table 3. Band Gaps for Cubic-Phase CsPbX_3 Given by HSE06, DeePKS, and PBE with the SOC Effect Taken into Account^a

system	HSE06	DeePKS	DeePKS Abs. Err	PBE	PBE Abs. Err
CsPbCl_3	2.1187	2.0777	0.0410	1.5382	0.5805
CsPbBr_3	1.6298	1.6295	0.0003	1.1261	0.5037
CsPbI_3	1.2333	1.2456	0.0123	0.9353	0.2980

^aAbsolute errors of the DeePKS and PBE results with respect to HSE06 results are listed. Numbers are shown in the unit of eV.

line with the HSE06 counterpart, while PBE predicts a larger volume of the equilibrium structure. Change of the band gap values with respect to the size of the unit cell is also explored, as shown in the lower panel of Figure 9. Closely matched band

gaps are obtained via DeePKS and HSE06, indicating that the band gap tuning by adjusting the pressure can be accurately predicted via the DeePKS model.

CONCLUSIONS

In this work, with 460 configurations over seven types of halide perovskites, we have established a general DeePKS model that is applicable to a variety of combinations, i.e., ABX_3 ($\text{A} = \text{FA}$, MA , Cs ; $\text{B} = \text{Sn}$, Pb ; $\text{X} = \text{Cl}$, Br , I) as well as the organic–inorganic hybrid perovskites and two RP perovskites ($\text{Cs}_2\text{PbI}_2\text{Cl}_2$ and MA_2PbI_4), with the accuracy of HSE06 in predicting electronic structure properties such as force, stress, band gap, and DOS, and the efficiency comparable to that of PBE. For all tested systems, especially hybrid perovskites with large unit cells, DeePKS offers orders of magnitude savings in

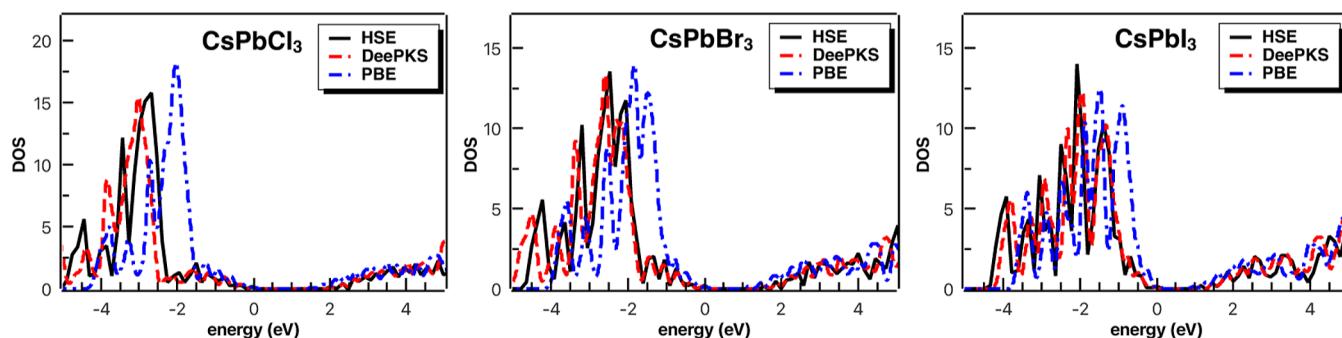


Figure 8. DOS for Pb-containing cubic-phase inorganic halide perovskites, i.e., CsPbX_3 ($\text{X} = \text{Cl}, \text{Br}, \text{I}$) with the SOC effect taken into account.

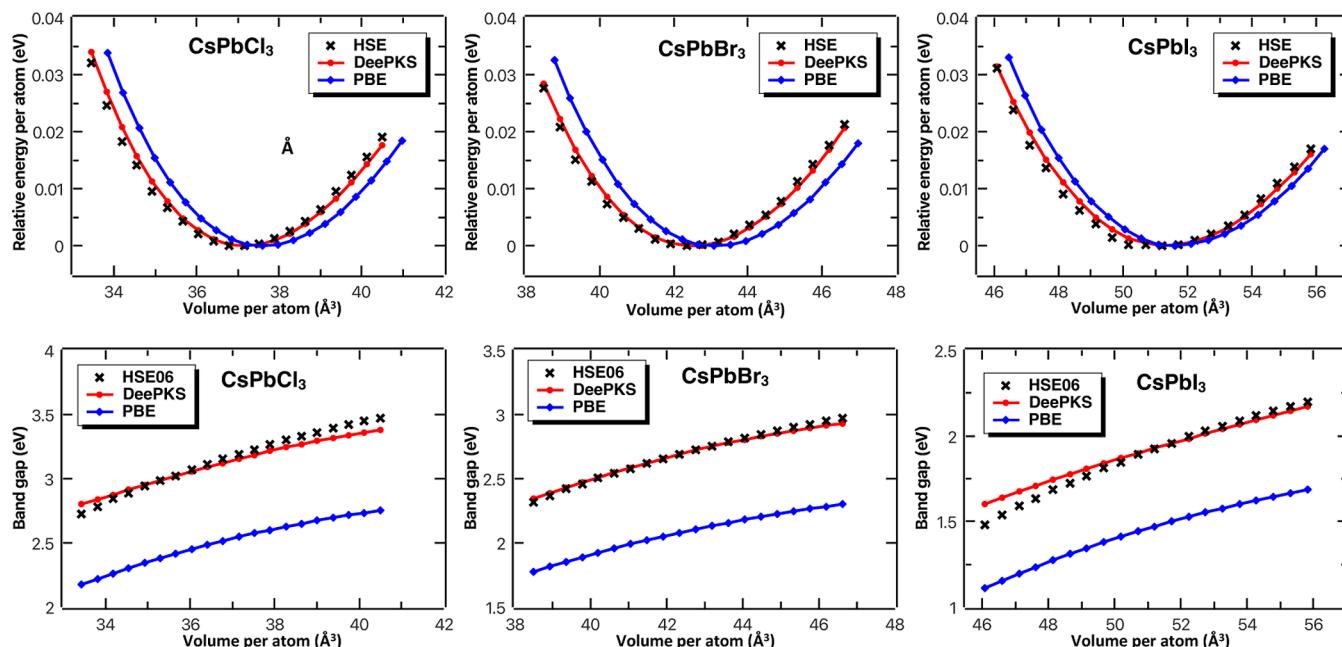


Figure 9. EoS (upper panel) and band gap values (lower panel) with respect to the size of the unit cell of CsPbBr_3 , CsPbCl_3 , CsPbBr_3 , and CsPbI_3 computed with HSE06, DeePKS, and PBE.

computational cost as compared to HSE06. While the MAEs of the energy and force predicted by DeePKS with respect to HSE06 are reduced by a factor of 2 as compared to those of PBE, the MAEs for the stress and band gap are reduced by at least 1 order of magnitude. Specifically, the MAE for the band gap given by DeePKS over 30 tested systems is less than 0.035 eV, demonstrating a promising alleviation of the underestimation issue suffered by PBE, of which the MAE is larger than 0.5 eV. The DOS predicted by DeePKS also closely matches with the target method in the vicinity of Fermi energy.

For Pb-containing perovskites, the SOC effect needs to be taken into account so as to offer a quantitatively correct description for the band gap and DOS. We have shown that DeePKS + SOC is able to offer a highly consistent band gap and DOS as compared to HSE06 + SOC for Pb-containing systems even though the SOC effect is not considered during the iterative training process. In addition, mechanical properties given by DeePKS have also been validated for inorganic halide perovskites by performing the EoS calculation, with highly consistent band gaps with respect to the size of the unit cell as compared to HSE06. In summary, with the closely matched electronic structure description with respect to the HSE06 functional and remarkably reduced computational cost,

our DeePKS model can be extensively applied to precisely predict a variety of properties and/or generate highly accurate labels for machine learning-based potential energy models for the family of halide perovskites, facilitating the discovery and design of better photovoltaics based on such systems.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcc.3c04703>.

Details of SCF calculation (including input parameters and run-time statistics), DeePKS iterative training procedure, and MAEs of various electronic structure properties for the training and test sets given by the DeePKS model ([PDF](#))

AUTHOR INFORMATION

Corresponding Authors

Qi Ou – AI for Science Institute, Beijing 100080, P. R. China;
orcid.org/0000-0002-6400-7522; Email: ou_qi@163.com

Linfeng Zhang – AI for Science Institute, Beijing 100080, P. R. China; DP Technology, Beijing 100080, P. R. China;

orcid.org/0000-0002-8470-5846;
Email: linfeng.zhang.zlf@gmail.com

Authors

Ping Tuo – AI for Science Institute, Beijing 100080, P. R. China
Wenfei Li – AI for Science Institute, Beijing 100080, P. R. China
Xiaoxu Wang – DP Technology, Beijing 100080, P. R. China
Yixiao Chen – Program in Applied and Computational Mathematics, Princeton University, Princeton, New Jersey 08544, United States; orcid.org/0000-0001-8201-5887

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.jpcc.3c04703>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Dr. Han Wang and Dr. Mohan Chen for their helpful conversations and computational guidance. The authors gratefully acknowledge the funding support from AI for Science Institute, Beijing (AISI) and DP Technology Corporation. The computing resource of this work was provided by the Bohrium Cloud Platform (<https://bohrium.dp.tech>), which is supported by DP Technology. The work of Y.C. was supported by the Computational Chemical Sciences Center: Chemistry in Solution and at Interfaces (CSI) funded by DOE Award DE-SC0019394 and a gift from iFlytek to Princeton University.

REFERENCES

- (1) Chu, S.; Majumdar, A. Opportunities and Challenges for a Sustainable Energy Future. *Nature* **2012**, *488*, 294–303.
- (2) Deng, Y.; Xu, J.; Liu, Y.; Mancl, K. Biogas as a Sustainable Energy Source in China: Regional Development Strategy Application and Decision Making. *Renewable Sustainable Energy Rev.* **2014**, *35*, 294–303.
- (3) Chu, S.; Cui, Y.; Liu, N. The Path towards Sustainable Energy. *Nat. Mater.* **2016**, *16*, 16–22.
- (4) Parida, B.; Iniyani, S.; Goic, R. A Review of Solar Photovoltaic Technologies. *Renewable Sustainable Energy Rev.* **2011**, *15*, 1625–1636.
- (5) Riede, M.; Spoltore, D.; Leo, K. Organic Solar Cells—The Path to Commercial Success. *Adv. Energy Mater.* **2021**, *11*, 2002653.
- (6) Nayak, P. K.; Mahesh, S.; Snaith, H. J.; Cahen, D. Photovoltaic Solar Cell Technologies: Analysing the State of the Art. *Nat. Rev. Mater.* **2019**, *4*, 269–285.
- (7) Gong, J.; Liang, J.; Sumathy, K. Review on dye-sensitized solar cells (DSSCs): Fundamental concepts and novel materials. *Renewable Sustainable Energy Rev.* **2012**, *16*, 5848–5860.
- (8) Shah, A.; Torres, P.; Tscharner, R.; Wyrsch, N.; Keppner, H. Photovoltaic Technology: The Case for Thin-Film Solar Cells. *Science* **1999**, *285*, 692–698.
- (9) Polman, A.; Knight, M.; Garnett, E. C.; Ehrler, B.; Sinke, W. C. Photovoltaic materials: Present efficiencies and future challenges. *Science* **2016**, *352*, aad4424.
- (10) Andreani, L. C.; Bozzola, A.; Kowalczewski, P.; Liscidini, M.; Redorici, L. Silicon solar cells: toward the efficiency limits. *Adv. Phys.: X* **2019**, *4*, 1548305.
- (11) Sun, Z.; Chen, X.; He, Y.; Li, J.; Wang, J.; Yan, H.; Zhang, Y. Toward Efficiency Limits of Crystalline Silicon Solar Cells: Recent Progress in High-Efficiency Silicon Heterojunction Solar Cells. *Adv. Energy Mater.* **2022**, *12*, 2200015.
- (12) Battaglia, C.; Cuevas, A.; De Wolf, S. High-efficiency crystalline silicon solar cells: status and perspectives. *Energy Environ. Sci.* **2016**, *9*, 1552–1576.
- (13) Peng, K.-Q.; Lee, S.-T. Silicon Nanowires for Photovoltaic Solar Energy Conversion. *Adv. Mater.* **2011**, *23*, 198–215.
- (14) Jin, J.; Wang, Q.; Ma, K.; Shen, W.; Belfiore, L. A.; Bao, X.; Tang, J. Recent Developments of Polymer Solar Cells with Photovoltaic Performance over 17%. *Adv. Funct. Mater.* **2023**, *33*, 2213324.
- (15) Wang, X.; Zhao, B.; Kan, W.; Xie, Y.; Pan, K. Review on Low-Cost Counter Electrode Materials for Dye-Sensitized Solar Cells: Effective Strategy to Improve Photovoltaic Performance. *Adv. Mater. Interfaces* **2022**, *9*, 2101229.
- (16) Datt, R.; Bishnoi, S.; Lee, H. K. H.; Arya, S.; Gupta, S.; Gupta, V.; Tsoi, W. C. Down-conversion materials for organic solar cells: Progress, challenges, and perspectives. *Aggregate* **2022**, *3*, No. e185.
- (17) Zhang, G.; Lin, F. R.; Qi, F.; Heumüller, T.; Distler, A.; Egelhaaf, H.-J.; Li, N.; Chow, P. C. Y.; Brabec, C. J.; Jen, A. K.-Y.; et al. Renewed Prospects for Organic Photovoltaics. *Chem. Rev.* **2022**, *122*, 14180–14274.
- (18) Lin, Y.; Zhan, X. Oligomer Molecules for Efficient Organic Photovoltaics. *Acc. Chem. Res.* **2016**, *49*, 175–183.
- (19) Bernardo, G.; Lopes, T.; Lidzey, D. G.; Mendes, A. Progress in Upscaling Organic Photovoltaic Devices. *Adv. Energy Mater.* **2021**, *11*, 2100342.
- (20) Jena, A. K.; Kulkarni, A.; Miyasaka, T. Halide Perovskite Photovoltaics: Background, Status, and Future Prospects. *Chem. Rev.* **2019**, *119*, 3036–3103.
- (21) Wells, H. L. Über die Cäsium- und Kalium-Bleihalogenide. *Z. Anorg. Allg. Chem.* **1893**, *3*, 195–210.
- (22) Xu, F.; Zhang, M.; Li, Z.; Yang, X.; Zhu, R. Challenges and Perspectives toward Future Wide-Bandgap Mixed-Halide Perovskite Photovoltaics. *Adv. Energy Mater.* **2023**, *13*, 2203911.
- (23) Wang, R.; Huang, T.; Xue, J.; Tong, J.; Zhu, K.; Yang, Y. Prospects for Metal Halide Perovskite-Based Tandem Solar Cells. *Nat. Photonics* **2021**, *15*, 411–425.
- (24) Chen, Y.; Liu, X.; Wang, T.; Zhao, Y. Highly Stable Inorganic Lead Halide Perovskite toward Efficient Photovoltaics. *Acc. Chem. Res.* **2021**, *54*, 3452–3461.
- (25) Yang, W.-F.; Igbari, F.; Lou, Y.-H.; Wang, Z.-K.; Liao, L.-S. Tin Halide Perovskites: Progress and Challenges. *Adv. Energy Mater.* **2020**, *10*, 1902584.
- (26) Tong, Y.; Najar, A.; Wang, L.; Liu, L.; Du, M.; Yang, J.; Li, J.; Wang, K.; Liu, S. F. Wide-Bandgap Organic-Inorganic Lead Halide Perovskite Solar Cells. *Adv. Sci.* **2022**, *9*, 2105085.
- (27) Gu, S.; Lin, R.; Han, Q.; Gao, Y.; Tan, H.; Zhu, J. Tin and Mixed Lead-Tin Halide Perovskite Solar Cells: Progress and their Application in Tandem Solar Cells. *Adv. Mater.* **2020**, *32*, 1907392.
- (28) Rigter, S. A.; Quinn, X. L.; Kumar, R. E.; Fenning, D. P.; Massonnet, P.; Ellis, S. R.; Heeren, R. M. A.; Svane, K. L.; Walsh, A.; Garnett, E. C. Passivation Properties and Formation Mechanism of Amorphous Halide Perovskite Thin Films. *Adv. Funct. Mater.* **2021**, *31*, 2010330.
- (29) He, R.; Wu, H.; Zhang, L.; Wang, X.; Fu, F.; Liu, S.; Zhong, Z. Structural phase transitions in SrTiO_3 from deep potential molecular dynamics. *Phys. Rev. B* **2022**, *105*, 064104.
- (30) Jain, D.; Chaube, S.; Khullar, P.; Goverapet Srinivasan, S.; Rai, B. Bulk and surface DFT investigations of inorganic halide perovskites screened using machine learning and materials property databases. *Phys. Chem. Chem. Phys.* **2019**, *21*, 19423–19436.
- (31) Grote, C.; Berger, R. F. Strain Tuning of Tin–Halide and Lead–Halide Perovskites: A First-Principles Atomic and Electronic Structure Study. *J. Phys. Chem. C* **2015**, *119*, 22832–22837.
- (32) ten Brinck, S.; Infante, I. Surface Termination, Morphology, and Bright Photoluminescence of Cesium Lead Halide Perovskite Nanocrystals. *ACS Energy Lett.* **2016**, *1*, 1266–1272.
- (33) Kashtiban, R. J.; Patrick, C. E.; Ramasse, Q.; Walton, R. I.; Sloan, J. Picoperovskites: The Smallest Conceivable Isolated Halide

- Perovskite Structures Formed within Carbon Nanotubes. *Adv. Mater.* **2023**, *35*, 2208575.
- (34) Xiao, Z.; Yan, Y. Progress in Theoretical Study of Metal Halide Perovskite Solar Cell Materials. *Adv. Energy Mater.* **2017**, *7*, 1701136.
- (35) Brivio, F.; Butler, K. T.; Walsh, A.; van Schilfgaarde, M. Relativistic quasiparticle self-consistent electronic structure of hybrid halide perovskite photovoltaic absorbers. *Phys. Rev. B* **2014**, *89*, 155204.
- (36) Even, J.; Pedesseau, L.; Jancu, J.-M.; Katan, C. Importance of Spin–Orbit Coupling in Hybrid Organic/Inorganic Perovskites for Photovoltaic Applications. *J. Phys. Chem. Lett.* **2013**, *4*, 2999–3005.
- (37) Wiktor, J.; Rothlisberger, U.; Pasquarello, A. Predictive Determination of Band Gaps of Inorganic Halide Perovskites. *J. Phys. Chem. Lett.* **2017**, *8*, 5507–5512.
- (38) Kaiser, W.; Carignano, M.; Alothman, A. A.; Mosconi, E.; Kachmar, A.; Goddard, W. A. I.; De Angelis, F. First-Principles Molecular Dynamics in Metal-Halide Perovskites: Contrasting Generalized Gradient Approximation and Hybrid Functionals. *J. Phys. Chem. Lett.* **2021**, *12*, 11886–11893.
- (39) Chen, Y.; Zhang, L.; Wang, H.; E, W. Ground State Energy Functional with Hartree–Fock Efficiency and Chemical Accuracy. *J. Phys. Chem. A* **2020**, *124*, 7155–7165.
- (40) Chen, Y.; Zhang, L.; Wang, H.; E, W. DeePKS: A Comprehensive Data-Driven Approach toward Chemically Accurate Density Functional Theory. *J. Chem. Theory Comput.* **2020**, *17*, 170–181.
- (41) Li, W.; Ou, Q.; Chen, Y.; Cao, Y.; Liu, R.; Zhang, C.; Zheng, D.; Cai, C.; Wu, X.; Wang, H.; et al. DeePKS + ABACUS as a Bridge between Expensive Quantum Mechanical Models and Machine Learning Potentials. *J. Phys. Chem. A* **2022**, *126*, 9154–9164.
- (42) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Hybrid Functionals Based on a Screened Coulomb Potential. *J. Chem. Phys.* **2003**, *118*, 8207–8215.
- (43) Das, T.; Di Liberto, G.; Pacchioni, G. Density Functional Theory Estimate of Halide Perovskite Band Gap: When Spin Orbit Coupling Helps. *J. Phys. Chem. C* **2022**, *126*, 2184–2198.
- (44) 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.
- (45) Chen, M.; Guo, G.-C.; He, L. Systematically Improvable Optimized Atomic Basis Sets for *Ab Initio* Calculations. *J. Phys. Condens. Matter* **2010**, *22*, 445501.
- (46) Tuo, P.; Li, L.; Wang, X.; Chen, J.; Zhong, Z.; Xu, B.; Dai, F.-Z. Spontaneous Hybrid Nano-Domain Behavior of the Organic–Inorganic Hybrid Perovskites. *Adv. Funct. Mater.* **2023**, *33*, 2301663.
- (47) Zhang, Y.; Wang, H.; Chen, W.; Zeng, J.; Zhang, L.; Wang, H.; EDP-GEN, W. DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models. *Comput. Phys. Commun.* **2020**, *253*, 107206.
- (48) Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **1996**, *6*, 15–50.
- (49) Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **1996**, *54*, 11169–11186.
- (50) Kresse, G.; Hafner, J. Ab initio molecular dynamics for open-shell transition metals. *Phys. Rev. B* **1993**, *48*, 13115–13118.
- (51) Yang, W.; Li, J.; Chen, X.; Feng, Y.; Wu, C.; Gates, I. D.; Gao, Z.; Ding, X.; Yao, J.; Li, H. Exploring the Effects of Ionic Defects on the Stability of CsPbI₃ with a Deep Learning Potential. *ChemPhysChem* **2022**, *23*, No. e202100841.
- (52) Jin, G.; Zheng, D.; He, L. Calculation of Berry Curvature Using Non-Orthogonal Atomic Orbitals. *J. Phys. Condens. Matter* **2021**, *33*, 325503.
- (53) Lin, P.; Ren, X.; He, L. Strategy for Constructing Compact Numerical Atomic Orbital Basis Sets by Incorporating the Gradients of Reference Wavefunctions. *Phys. Rev. B* **2021**, *103*, 235131.
- (54) Hamann, D. R. Optimized Norm-Conserving Vanderbilt Pseudopotentials. *Phys. Rev. B* **2013**, *88*, 085117.
- (55) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (56) Li, J.; Yu, Q.; He, Y.; Stoumpos, C. C.; Niu, G.; Trimarchi, G. G.; Guo, H.; Dong, G.; Wang, D.; Wang, L.; Kanatzidis, M. G. Cs₂PbI₂Cl₂, All-Inorganic Two-Dimensional Ruddlesden–Popper Mixed Halide Perovskite with Optoelectronic Response. *J. Am. Chem. Soc.* **2018**, *140*, 11085–11090.
- (57) Gebhardt, J.; Kim, Y.; Rappe, A. M. Influence of the Dimensionality and Organic Cation on Crystal and Electronic Structure of Organometallic Halide Perovskites. *J. Phys. Chem. C* **2017**, *121*, 6569–6574.
- (58) Zhang, L.; Zhang, X.; Lu, G. Intramolecular Band Alignment and Spin–Orbit Coupling in Two-Dimensional Halide Perovskites. *J. Phys. Chem. Lett.* **2020**, *11*, 6982–6989.
- (59) Yumoto, G.; Hirori, H.; Sekiguchi, F.; Sato, R.; Saruyama, M.; Teranishi, T.; Kanemitsu, Y. Strong Spin–Orbit Coupling Inducing Autler–Townes Effect in Lead Halide Perovskite Nanocrystals. *Nat. Commun.* **2021**, *12*, 3026.
- (60) Chatterjee, S.; Payne, J.; Irvine, J. T. S.; Pal, A. J. Bandgap bowing in a zero-dimensional hybrid halide perovskite derivative: spin–orbit coupling versus lattice strain. *J. Mater. Chem. A* **2020**, *8*, 4416–4427.
- (61) Scherpelz, P.; Govoni, M.; Hamada, I.; Galli, G. Implementation and Validation of Fully Relativistic GW Calculations: Spin–Orbit Coupling in Molecules, Nanocrystals, and Solids. *J. Chem. Theory Comput.* **2016**, *12*, 3523–3544.
- (62) Hafner, J. Ab-initio simulations of materials using VASP: Density-functional theory and beyond. *J. Comput. Chem.* **2008**, *29*, 2044–2078.