

Machine Learning Accelerated Prediction of Self-Trapped Excitons in Double Halide Perovskites

Baian Chen, Rui Chen,* and Bolong Huang*

Broadband emission induced by self-trapped excitons (STEs) in double halide perovskites (DHPs) has received continuous attention in recent years. However, the comprehensive understanding of the STEs formation mechanism is still in its early stage. The corresponding roles of different B-site cations also remain unclear in these advanced materials. The lack of an effective STEs database for DHPs hinders the efficient discovery of potential optoelectronic materials with strong STEs. Herein, a systematic STEs database is built for DHPs through density functional theory (DFT) calculations and proposed a highly efficient predictive machine learning (ML) model of the Huang–Rhys factor S for the first time. Results reveal the different contributions of two B-site metal cations to the formation of STEs in DHPs, which helps to understand the in-depth nature of STEs. Based on the accurate predictions of the effective phonon frequency ω_{LO} , it is further realized that the prediction of S without conducting the time-consuming phonon property calculations of DHPs offers new opportunities for exploring the STEs. Combining DFT calculations and ML techniques, this study supplies an effective approach to efficiently discover the potential novel optoelectronic materials, which provides important guidance for the future exploration of promising solid-state phosphors.

1. Introduction

Halide perovskites have received intensive attention recently because of their advanced optoelectronic properties and extensive applications.^[1–13] In particular, the observed broadband white-light emissions in halide perovskites have been widely investigated due to their great potential in solid-state broad photoluminescence.^[14–18] Conventional broadband phosphors usually rely on extrinsic dopants and surface sites.^[19,20] In contrast, current methods produce white light by mixing multiple kinds of phosphors or light-emitting diodes. The broadband emission achieved by halide perovskites covers the most range of visible light to avoid the inevitable color changes and energy loss induced by the mixing of phosphors, which improves the general efficiency and quality of solid-state photoluminescence. Therefore, the stable and efficient broadband emissions in halide perovskites are of great importance for the development of single-source emitters and desirable


for the next generation of solid-state photoluminescence. Currently, the origin of these intriguing broadband emissions in halide perovskites is widely attributed to the inner self-trapped excitons (STEs).^[21–23] The STEs effect indicates the phenomenon that excitons are confined to the lattices by the intense electron-phonon coupling inside the materials. This self-trapping process will induce corresponding lattice distortions and result in the self-trapped states within the bandgap. These newly formed states have lower energy levels than free excitons and their emissions will process a broad full width at half maximum (FWHM) together with a large Stokes shift.^[24] The dimensionless Huang–Rhys factor S initially describes the interaction intensities of the electrons bound to the localized centers with the phonons in crystals.^[25] Such an electron–phonon coupling effect induces a corresponding lattice relaxation, which refers to the changes in the atomic equilibrium configurations between the ground state and the excited states. This process is accompanied by a unique optical transition, which matches the observed characteristics of the STEs emissions. Accordingly, the dimensionless Huang–Rhys factor S has been widely used as a reliable index to evaluate the strength of STEs effects in various materials.^[26–28]

The soft lattice property and impressive photovoltaic performance of conventional metal halide perovskites are conducive to the formation of STEs. However, these emerging materials also suffer serious challenges, such as instabilities at high

B. Chen, B. Huang
Department of Applied Biology and Chemical Technology
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong SAR 999077, China
E-mail: bhuang@polyu.edu.hk

B. Chen, R. Chen
Department of Electrical and Electronic Engineering
Southern University of Science and Technology
Shenzhen 518055, China
E-mail: chenr@sustech.edu.cn

B. Huang
Research Centre for Carbon-Strategic Catalysis
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong SAR 999077, China

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/aesr.202300134>.

© 2023 The Authors. Advanced Energy and Sustainability Research published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aesr.202300134

humidity, unsustainable performances at high temperatures, the toxicity of lead, etc.^[29–32] Compared to traditional halide perovskites, double halide perovskites (DHPs) have a combination of monovalent and trivalent cations on the original B-sites. This unique feature realizes the lead-free property in DHPs to improve their compositional compatibility. In addition, the alternating arrangement of the monovalent and trivalent B-site cations breaks the original spatial symmetry of the lattice array, improving the structural tunability and electronic dimensionality of DHPs.^[33] Therefore, DHPs are better candidates compared to traditional halide perovskites for the further development of STEs materials. In recent years, there have been increasing reports pointing out the observed strong STEs effect in DHPs.^[34–36] For example, Luo et al. have successfully obtained efficient broadband emission in DHPs by alloying the B-site cations. Their optimally alloyed Cs₂(Ag_{0.60}Na_{0.40})InCl₆ with 0.04% bismuth doping produced a durable white light with a high quantum efficiency of 86 ± 5%.^[34] In addition, Han et al. also realized an efficient broadband emission in DHPs nanocrystals with high stability by doping Ag⁺ into the Cs₂NaInCl₆ lattice to convert the original dark STEs into bright states.^[35] Wang et al. also observed the distinctive STEs emissions in a series of doped DHPs and they confirmed the corresponding lattice distortions in an excited state by using the femtosecond transient absorption spectroscopy.^[36] Although these previous studies have demonstrated the great potential of DHPs as STEs materials, they still lack a systematic summary of the STEs formation mechanisms and the corresponding roles of different B-site cations in the DHPs. Ke et al. obtained broad STEs emissions in their doped DHPs systems and revealed that the final dominant emission is attributed to the distinct energy-transfer channel from the dopant guest to the double-perovskite host.^[37] However, the doping mechanism to produce the obvious broadband emissions in DHPs is still obscure due to the limited reported samples. Another difficulty impeding the further development of STEs materials is the lack of high-throughput capabilities to screen large-scale STEs candidates. Even though the fully decoded-STE mechanisms are still absent, the high-throughput screening of large-scale STEs materials offers an effective approach to quickly identify the potential promising STEs candidates and inspire the design of next-generation STEs devices to break the current efficiency bottlenecks.

As a powerful data-driven tool to solve problems in materials sciences, machine learning (ML) is capable of processing large-scale datasets, predicting unknown properties, and screening target materials.^[38–48] Applying ML techniques for large-scale screening of STEs candidates in DHPs is expected to be a promising approach to realize the fast discovery of advanced STEs materials. As a powerful ML regression algorithm, the gradient boosting regression (GBR) algorithm method is able to achieve a reasonable balance between strong predictive capabilities and acceptable interpretability, which have been extensively applied in the screening of various materials.^[49–53] Technically, the GBR method is an ensemble algorithm that combines multiple weak prediction models to get an overall better performance.^[54] However, there is no valid and comprehensive database on STEs properties in DHPs supplied by experiments or theoretical calculations. In our previous work, we have evaluated the STEs properties in DHPs Cs₂B¹B²Cl₆ (B¹ = Na⁺, K⁺;

B² = Al³⁺, Ga³⁺, In³⁺) and confirmed that the Jahn–Teller distortion in a strong excitonic as well as electron–phonon coupling environment is the main source of STEs generation.^[55] Our proposed method provides an opportunity to evaluate STEs in DHPs based on DFT calculations and build a reliable database for the further development of STEs materials.

To overcome the challenges, we present a systematic STEs database of DHPs in this work based on the DFT calculations and propose an accurate ML model regarding the prediction of Huang–Rhys factor *S* for the first time. As shown in **Figure 1**, the whole workflow of this work is divided into three phases: dataset preparation, model selection, and result analysis. In the dataset preparation phase, we first construct 72 DHPs structures for DFT calculations based on the atomic substitutions of the B-site metal cations. Then, we perform geometry optimizations on these structures based on strict criteria and match them with a series of basic features (also called as descriptors). Various physical parameters of these structures including the effective mass, dielectric constants, and phonon frequencies are also carefully calculated as important parts of the database. According to these fundamental parameters, we further investigate the excitonic and electron–phonon coupling characteristics of our constructed DHPs structures. After excluding the structures with a failed estimation of the excited state energy, we successfully calculate the Huang–Rhys factor *S* values of the left 60 suitable structures. Accordingly, we combine the DFT results and the specific features to establish a complete dataset, which contains the STEs information of these materials. Through the GBR algorithm and cross-validation method to train and evaluate the target ML models, respectively, we finally screen out the best-performing ML model from the iterations to predict our target objects. In the end, the performance of these trained models and the potential connections between the physical parameters and the labeled features are investigated through the output Pearson correlation factors (CFs), which is one of the most used approaches to measure the strength and direction of a linear relationship between two variables.^[56] Notably, we have successfully achieved efficient and accurate predictions of *S* values based on our improved trained ML models with the predicted ω_{LO} . In particular, we predicted the *S* value of 56.05 for the unknown Cs₂CuSbCl₆ as a promising STEs material, supporting the practical predictive capabilities of our proposed method. Combining the DFT calculations and the ML techniques, this research provides a solid approach to screen potential novel STEs materials with high efficiency, which creates the possibility for the design of next-generation optoelectronic materials.

2. Results and Discussion

DHPs with formula A₂B¹B²X₆ generally contain metal or organic cations on the A-site positions, monovalent and trivalent metal cations on the B-site positions, and halogen anions on the X-site positions. To investigate the roles of different B-site cations on the STEs properties in DHPs, we fix the A-site cations as Cs⁺ and X-site anions as Cl[−] in this work for the DHPs models. For the selection criteria of B¹ and B² cations, we strictly choose the elements with common valence states of +1 (Group 1 and 11) and +3 (Group 3, 13, and 15), respectively. Elements with

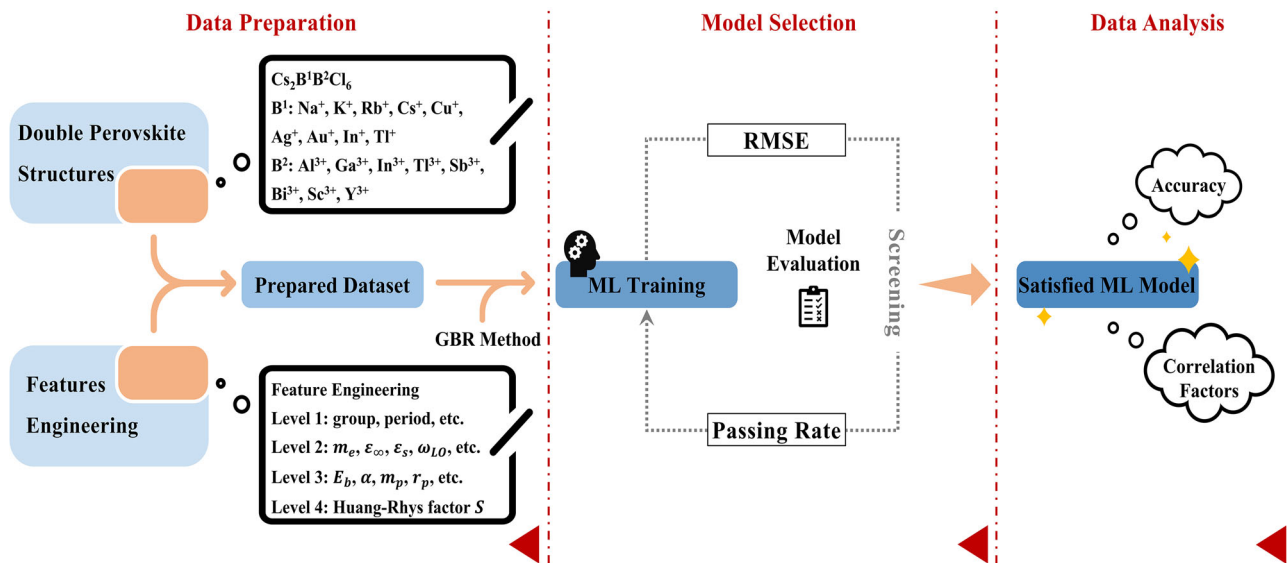


Figure 1. Schematic of the workflow combined with the DFT calculations and the ML techniques to train the satisfied ML models for the STEs properties in DHPs (Cs₂B¹B²Cl₆).

nonmetallic or rarely display the valence state +1 and +3 are not considered in this work because they are not applicable to the practical synthesis of perovskites. As a result, 72 possible DHPs structures Cs₂B¹B²Cl₆ are constructed based on the different combinations of B¹ and B² cations. In detail, as shown in **Figure 2a**, B¹ = Na⁺, K⁺, Rb⁺, Cs⁺, Cu⁺, Ag⁺, Au⁺, In⁺, Tl⁺; B² = Al³⁺, Ga³⁺, In³⁺, Tl³⁺, Sb³⁺, Bi³⁺, Sc³⁺, Y³⁺. The elements of In and Tl have both +1 and +3 as the common valence states, so we also consider their possibilities in the positions of B¹ even though they belong to Group 13. The unit cell of our constructed DHPs Cs₂B¹B²Cl₆ is shown in **Figure 2b**, of which the lattice belongs to the triclinic system, and the two different kinds of octahedrons are aligned alternately. The green octahedron and blue octahedron represent the B¹Cl₆, and B²Cl₆, respectively. The overall structure of DHPs is demonstrated as a corner-sharing three-dimensional array of metal-halogen octahedrons. The A-site cations are in the space between the octahedrons, while the B¹ and B² cations alternately occupy the center of these octahedrons. Different combinations of the B¹ and B² cations change the size ratio and spatial configuration of these octahedrons, resulting in various electronic states distribution and lattice vibration modes. All these key changes significantly alter the formation of STEs. Based on DFT calculations, we have first calculated the bandgap of all the DHPs (**Figure 2c**), which have shown a good consistency with the previous reported works (Table S1, Supporting Information). As we observed, the introduction of Sc and Y atoms increases the bandgap values because of their high energy level of antibonding orbitals with halogen atoms. In contrast, the introduction of Cu, Ag, and Au atoms leads to an obvious metallization of these materials. There are seven samples (Cs₂CuGaCl₆, Cs₂CuInCl₆, Cs₂CuTlCl₆, Cs₂AgTlCl₆, Cs₂AuGaCl₆, Cs₂AuInCl₆, and Cs₂AuTlCl₆) classified as conductors and will not be included in our dataset for STEs.

To construct an effective database of STEs information in DHPs, it is necessary to select and calculate quantifiable physical parameters for describing the STEs properties in these materials, which requires an in-depth understanding of the formation mechanism of STEs. In general, four essential factors determine the formation of STEs in DHPs: intense excitonic effects, unique atomic compositions, appropriate lattice configurations, and strong electron–phonon couplings. The atomic composition and lattice configuration are basic information of the specific lattice, which are usually described by several simple parameters such as atomic numbers, bond lengths, lattice volumes, etc. In this work, we briefly use the lattice volume to quantify the difference between these similar DHPs lattice configurations. For the atomic composition, considering the different combinations of the B-site cations, we select 8 fundamental features to accurately describe the difference between B¹ and B² cations for each structure including 1) group; 2) period; 3) atomic number; 4) atomic mass; 5) atomic radius; 6) electronegativity; 7) first ionization energy; and 8) their bond length with halogen atoms. The lattice volume and bond length data are obtained from the DFT calculations, and the other seven fundamental atomic information are collected from the online database of the International Union of Pure and Applied Chemistry (IUPAC).^[57] These primary data form the first part of our STEs database, which is defined as “Level 1 data”. Besides these data, we also need some quantifiable parameters to characterize the electronic, dielectric, and vibrational properties of these DHPs structures for the STEs effect. Accordingly, we select and calculate six common features for these materials: the effective mass of electrons (m_e) and holes (m_h), the reduced mass (μ), the static (ε_s) and optical (ε_∞) dielectric constants, and the effective phonon frequency (ω_{LO}). The m_e, m_h, ε_s, and ε_∞ are directly obtained from the DFT calculations while the μ and ω_{LO} are derived from the corresponding formulas based on the

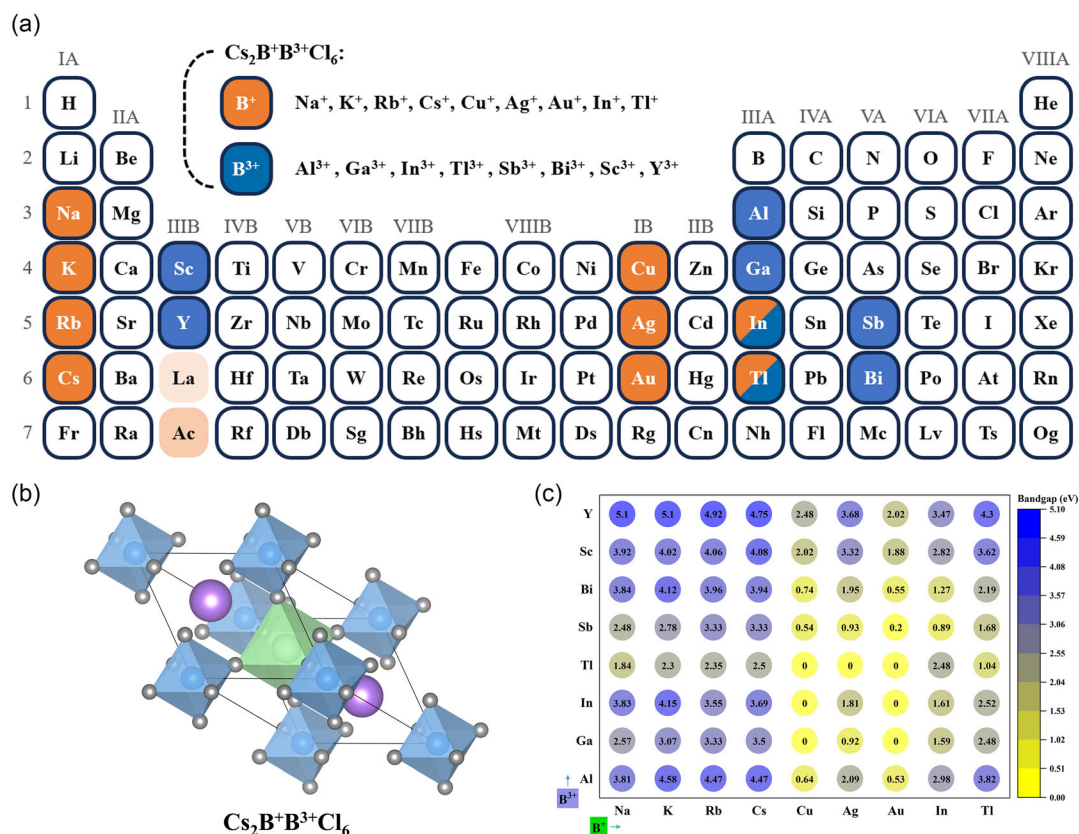


Figure 2. a) Schematic diagram of periodic table showing the selection of different cations in double halide perovskites; b) Lattice structure of the constructed double halide perovskite model for the DFT calculations. Black lines confine the triclinic lattice. The green and blue color represents the octahedron of B⁺ and B³⁺, respectively; c) Calculated bandgap values of various double halide perovskites based on different combinations of B-site metal cations.

acquisition of necessary parameters in advance. These data form the second part of our STEs database, which is defined as “Level 2 data”.

In the end, we attempt to calculate the excitonic properties and electron–phonon coupling effect in these structures. In general, a photogenerated exciton is understood as a stable lone electron–hole pair confined by the Coulomb force in the materials. From this aspect, the strength and size of one exciton are described by the binding energy and radius, respectively. By applying the hydrogenic Rydberg model, we calculate the exciton binding energy E_b and the exciton Bohr radius r_b of these DHPs materials to evaluate their excitonic characteristics.^[58,59] Our results of the E_b and r_b range are presented in Figure 3a, which supply important database for the ML predictions. As for the other issue, the electron–phonon coupling effect is described by the polaron theory. In detail, the moving electrons in ionic crystals may interact with the long-wave polar optic vibrations and keep inevitable coupling effect with their neighbor lattices, which are considered as the origin of the self-trapped states.^[60,61] After applying the Frohlich coupling model and the Feynman polaron theory, we choose three crucial parameters to describe the strength of the electron–phonon coupling effect in these DHPs structures: Frohlich coupling constant (α), polaron radius (r_p), and polaron mass (m_p).^[62,63] These parameters in different structures are

presented in Figure 3b, where the horizontal coordinate represents the coupling constant, the circle size represents the polaron radius, and the color depth represents the polaron mass. All these five parameters (E_b , r_b , α , r_p , and m_p) form the third part of our STEs database, which is defined as “Level 3 data”.

To train an effective ML model regarding the prediction of S values in DHPs, the S values of known materials are needed to generate the corresponding GBR model. Therefore, we calculate the corresponding lattice deformation energy E_d of these DHPs structures between the ground state and the first excited state. We further obtain the required S values by calculating $S = \frac{E_d}{\hbar\omega_{LO}}$, and the S values are defined as “Level 4 data”. Here, five DHPs structures (Cs₂InSbCl₆, Cs₂AgGaCl₆, Cs₂CuAlCl₆, Cs₂CuSbCl₆, Cs₂CuBiCl₆) fail to output a reasonable E_d value due to their negligible variations between the equilibrium atomic coordinates of the two different states. Therefore, they are excluded from the target STEs database of DHPs. Detailed computation methods related to all the above parameters are supplied in Supporting Information or referred to our previous work.^[55] After careful evaluations of all the parameters, we finally built a systematic STEs database of DHPs containing 60 valid samples for the first time.

As we mentioned above, the initial lattice characteristics of one identified DHPs structure should play a decisive role

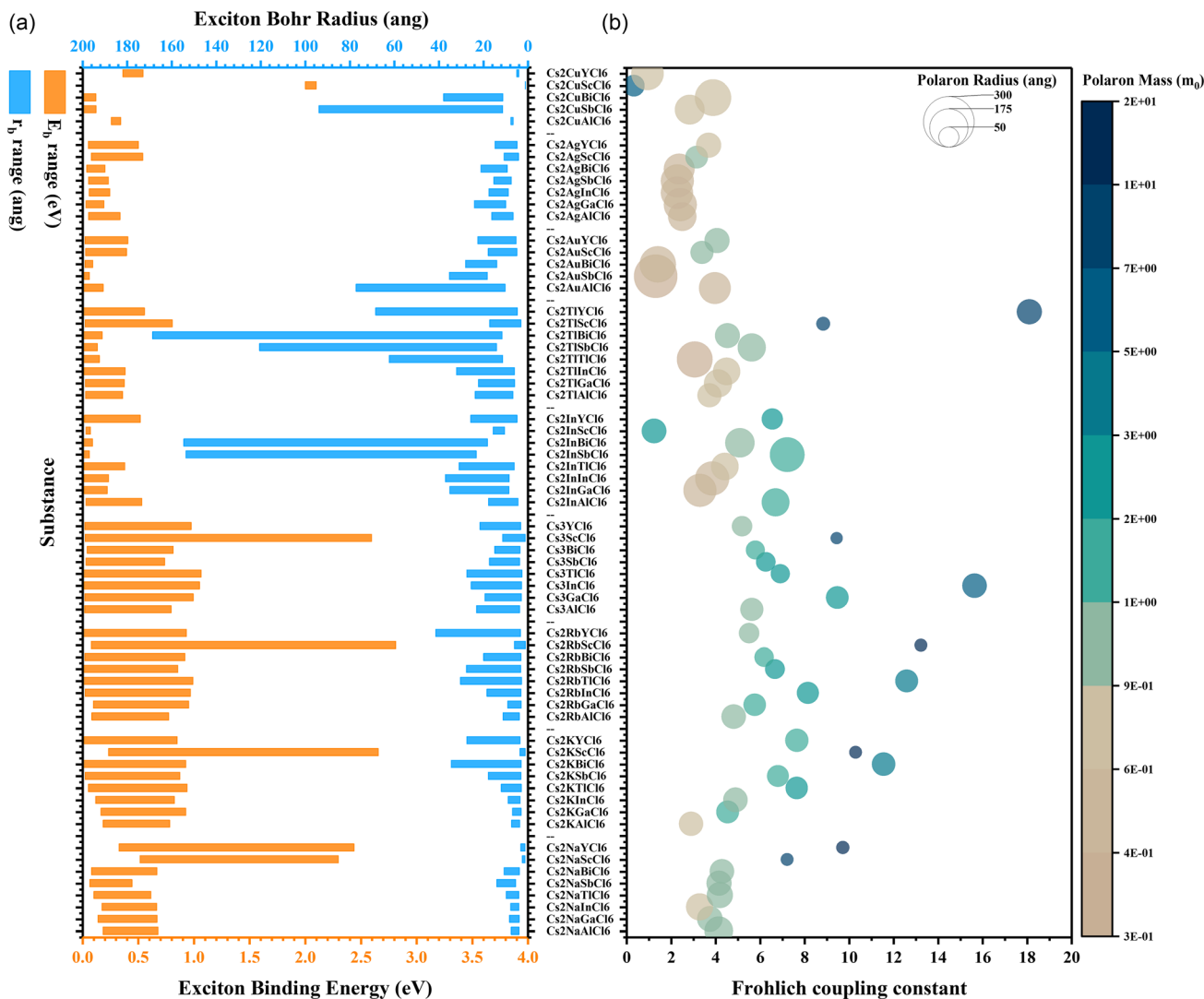


Figure 3. Calculated a) excitonic and b) polaronic characteristics of different DHPs.

in its fundamental physical nature such as the dielectric and electron–phonon coupling properties, which satisfies the law of causality. To obtain satisfactory ML models for accurate predictions of the expected properties of the known DHPs materials, we perform the supervised learning algorithms on the current databases. This approach alleviates the heavy loading of DFT calculations and bridges an efficient channel between the basic information and the complicated parameters, which creates the possibility of high-throughput screening of large-scale STEs materials. In this work, we choose the GBR method as the ML algorithm to tackle our STEs regression task. Based on our self-prepared database, we first train and screen six ML models to predict the six crucial indicative parameters (μ , ϵ_∞ , ϵ_s , ω_{LO} , α , and r_p) of DHPs. Detailed screen procedures refer to the corresponding descriptions and Figure S1, Supporting Information. During this process, we use the Level 1 data as the input labeled features to generate the prediction ML models for these six parameters. **Figure 4** presents the comparison of the electronic properties of our DHPs structures between ML predictions and DFT

calculations. Here, we use two indexes to evaluate the predictive performance of the trained ML models: root mean square error (RMSE) and passing rate (PR). The RMSE index measures the average difference between the ML-predicted results and the DFT-calculated results, which is expressed by the formula $RMSE = \sqrt{\frac{\sum_{i=1}^T (\hat{y}_i - y_i)^2}{T}}$. For the other index PR, it is defined specifically in this work to evaluate the proportion of the forecast results within the acceptable error range (25% in this work) to the total predictions, which is expressed as $PR(25\%) = \frac{NUM(\frac{|y_{pre} - y_{true}|}{y_{true}} < 25\%)}{NUM(y_{true})}$. As shown in Figure 4a–f, our ML-trained models for the six parameters (μ , ϵ_∞ , ϵ_s , ω_{LO} , α , and r_p) of the DHPs structures exhibit strong predictive ability with 100% PR and low RMSE of 0.01, 0.01, 0.17, 4.50, 0.17, and 1.50, respectively. Such high prediction accuracies confirm that these higher-level data are precisely predicted by our trained ML models based on the fundamental lattice information. In addition, we output the CFs between the expected parameters and the labeled features after training the ML models. The absolute value

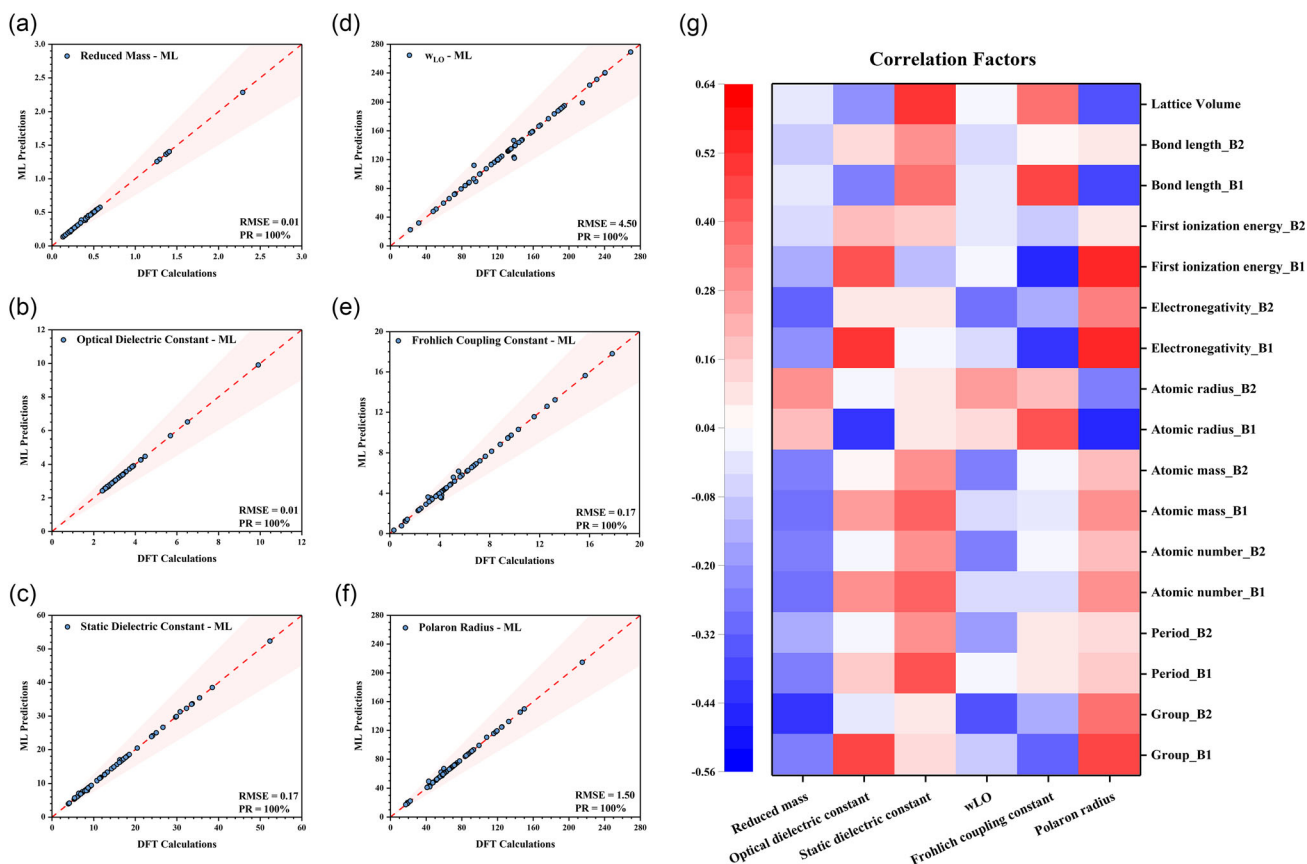


Figure 4. Comparison of six physical parameters between the ML predictions and the DFT calculations on DHPs. a) Reduced mass μ , b) optical dielectric constant ϵ_{∞} , c) static dielectric constant ϵ_s , d) effective phonon frequency ω_{LO} , (e) Frohlich coupling constant α , and f) polaron radius r_p . g) The heat map of the correlation factor matrices between these physical parameters and the selected descriptors.

of CFs represents the strength of their correlation, and the positive (or negative) factors indicate a positive (or negative) correlation, which helps us further explore the roles of different B-site cations in the generation process of STEs. As shown in Figure 4g, different features have different degrees of influence on the physical properties of the target. In this work, we propose a simple criterion, where the features with an absolute value of CF greater than 0.4 are regarded as strongly correlated features. Moreover, to distinguish the influence degree of the B¹ and B² cations on different target physical parameters, we calculate the sum of the absolute values of the CFs related to B¹ and B² cations, denoted as CF(B¹) and CF(B²), respectively. The integrated results of CF(B¹) and CF(B²) are listed in Table 1 for a brief comparison.

For the reduced mass μ , the group of B² is regarded as a strongly correlated feature with a high CF value of -0.51 .

Table 1. The comparison of the absolute value sum of the CFs related to B¹ and B² cations.

| | μ | ϵ_{∞} | ϵ_s | ω_{LO} | α | r_p | S (with ω_{LO}) |
|---------------------|-------|---------------------|--------------|---------------|----------|-------|-------------------------|
| CF(B ¹) | 2.02 | 3.25 | 2.20 | 0.49 | 2.61 | 3.61 | 0.78 |
| CF(B ²) | 2.20 | 0.61 | 1.84 | 1.90 | 0.86 | 1.82 | 1.21 |

The corresponding CF(B¹) is calculated as 2.02, which is slightly lower than the value of CF(B²), 2.20. Since reduced mass is derived from the effective mass of electrons and holes, this result implies that B² cations have more influence on the transfer ability of electrons in DHPs than B¹ cations. This also means that B² cations are more likely to become localization centers of electrons or holes inside the DHPs, providing suitable conditions for potential STEs formation. For the optical dielectric constant, ϵ_{∞} , the strongly correlated features are the group of B¹, the atomic radius of B¹, the electronegativity of B¹, and the first ionization energy of B¹, with CF values as 0.54, -0.51 , 0.58, and 0.51, respectively. The corresponding CF(B¹) is 3.25, and the CF(B²) is 0.61. For static dielectric constant ϵ_s , the strongly correlated features become the period of B¹, the atomic number of B¹, the atomic mass of B¹, the bond length of B¹, and the lattice volume, with CF values as 0.48, 0.45, 0.45, 0.41, and 0.56, respectively. The corresponding CF(B¹) is 2.20, and the CF(B²) is 1.84. These important data reflect that the B¹ cations are more responsible for the determination of the internal dielectric properties in DHPs, which will further affect the strength of the electron-phonon coupling effect. For the effective phonon frequency ω_{LO} , the group of B² is the only strongly correlated feature with a CF value of -0.42 . The CF(B¹) value of this parameter is calculated as 0.49, while the CF(B²) is 1.90. Clearly, the evident

difference between the $CF(B^1)$ and $CF(B^2)$ of ω_{LO} represents the major roles of B^2 cations in determining the phonon patterns of DHPs when compared to B^1 cations. This supports that the substitution or doping of B^2 cations induces the enhanced interactions between electrons and lattice vibrations by altering the phonon patterns, which contributes to the formation of STEs in DHPs. For the Frohlich coupling constant α , the strongly

correlated features are the atomic radius of B^1 , the electronegativity of B^1 , the first ionization energy of B^1 , the bond length of B^1 , and the lattice volume, with CF values as 0.50, -0.49 , -0.53 , 0.53 , and 0.43 , respectively. The corresponding $CF(B^1)$ is 2.61, and $CF(B^2)$ is 0.86. For the polaron radius r_p , there are more strongly correlated features including the group of B^1 and B^2 , the atomic radius of B^1 , the electronegativity of B^1 , the first

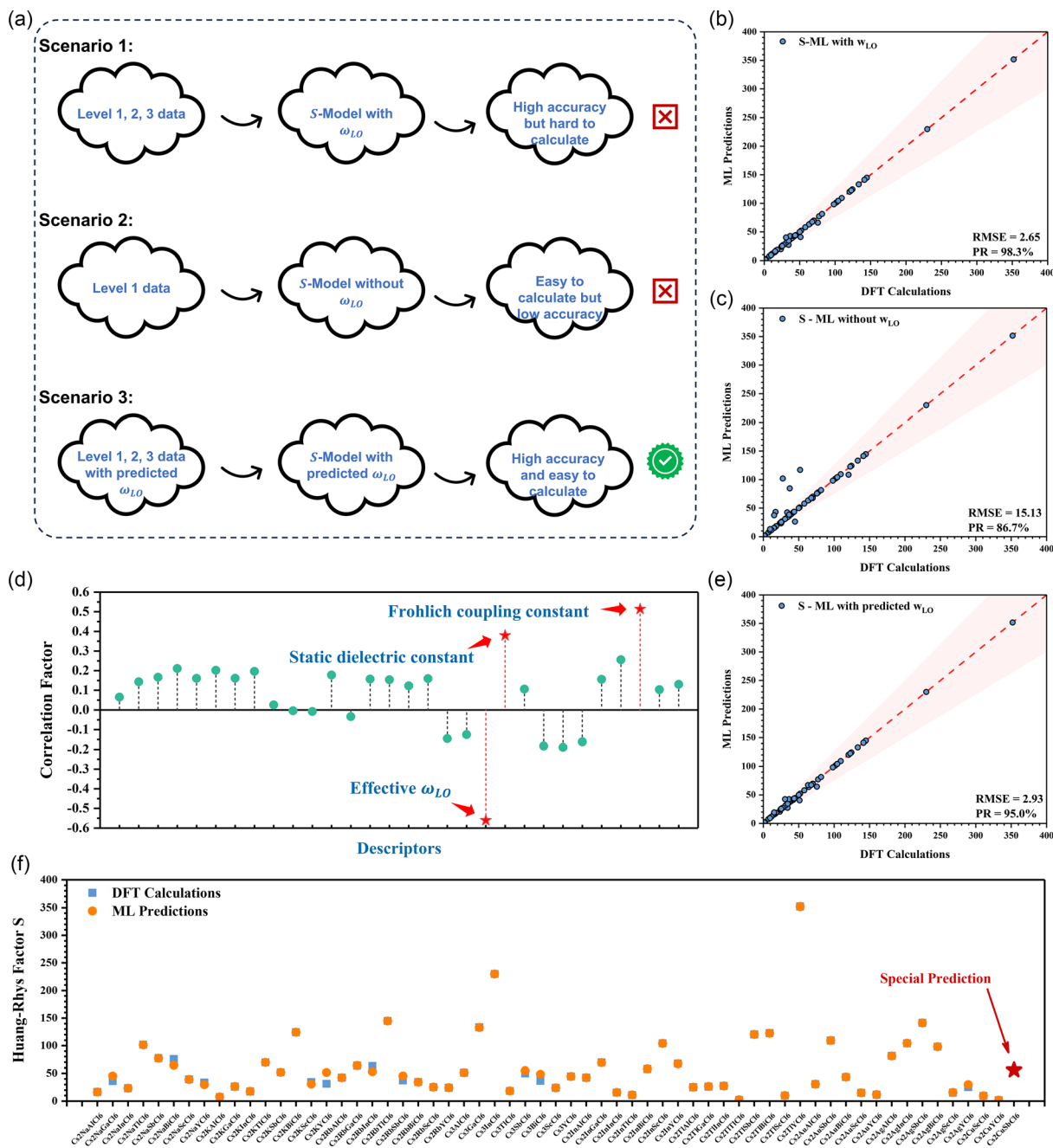


Figure 5. a) The schematic diagram indicates the strategies to realize the accurate prediction of S values in different scenarios. b–d) Comparison of the predictive ability between the different S-ML models: (b) with DFT-calculated ω_{LO} , c) without ω_{LO} , and d) with ML-predicted ω_{LO} . e) The correlation factors between Huang–Rhys factor S values and different descriptors in our trained S-model. f) The comparison of the S values between the DFT calculations and the ML predictions based on the S-ML model with predicted ω_{LO} . The red star represents the special predicted S value (56.05) for $Cs_2CuSbCl_6$.

ionization energy of B^1 , the bond length of B^1 , and the lattice volume, which exhibit CF values of 0.54, 0.42, -0.56 , 0.63, 0.61, -0.46 , and -0.41 , respectively. The corresponding $CF(B^1)$ is 3.61, and $CF(B^2)$ is 1.82. The relatively higher $CF(B^1)$ values of these two parameters indicate that the B^1 cations count more on the polaronic effect in DHPs when compared to B^2 cations, indicating that the B^1 cations play a more important role in tuning the strength of electron–phonon coupling in DHPs.

To characterize the STEs effects in our DHPs structures, the calculation of Huang–Rhys factor S becomes a necessary step. However, this process requires considerable computation demand due to the calculations of phonon properties and lattice deformation energies. Our above efforts have verified the feasibility of the implemented ML techniques in predicting complex parameters. To reduce the required computational loadings for the Huang–Rhys factor S , we further train an effective ML model to predict such values based on our prepared database. According to different input data, we have proposed three different scenarios (Figure 5a). In this way, the Huang–Rhys factor S values are expected to be directly obtained to realize the efficient and high-throughput screening of STEs candidates. For the S-ML model training, we start with inputting all the obtained data (Level 1, 2, and 3 data) as descriptors to verify our understanding of the STEs formation mechanisms. It is noted that an excellent S-ML model with strong predictive ability is achieved (Figure 5b), which has a high PR of 98.3% and a low RMSE of 2.65. Although this result shows accurate predictions, such an ideal S-ML model fed with the complete physical parameters (Level 1, 2, and 3 data) is not practical, which corresponds to Scenario 1 in Figure 5a. This is because the reuse of the ideal S-ML model also requires the acquisition of these comprehensive data in advance, where the time-consuming calculations of phonon properties are still challenging. The goal of the S-ML model is to achieve accurate predictions of S values with only the input of rough physical parameters of the DHPs. We next train a S-ML model with only inputting the Level 1 data as the labeled features, corresponding to Scenario 2 in Figure 5a. As a result, this trained S-ML model exhibits a poor predictive ability with a low PR of 86.7% and a high RMSE of 15.13 (Figure 5c). According to the performance of these two S-ML models, we understand the competitive relationship between their accurate prediction results and rough input data. To maintain a balance between the high predictive accuracy and the low computation amount of our S-ML model, the parameters strongly correlated to the S values should be preserved as the input labeled features during the training process. As shown in Figure 5d and Table 1, we present the CFs between S values and different descriptors based on the S-ML model with ω_{LO} , where the corresponding CF is 0.78 and 1.21 for B^1 and B^2 , respectively. Although the overall $CF(B^2)$ is slightly higher than $CF(B^1)$, neither B^1 nor B^2 cations show a strongly correlated feature. There are three parameters that are important for the S-prediction: static dielectric constant ϵ_s , Frohlich coupling constant α , and effective phonon frequency ω_{LO} . For the static dielectric constant ϵ_s , we calculate this parameter of DHPs with little DFT computational costs. The Frohlich coupling constant α is easily derived if the prerequisite parameter ω_{LO} has been obtained in advance. Thus, the effective phonon frequency ω_{LO} is the only parameter worth concerning, which usually requires considerable computation loadings. Fortunately, we

have trained the ω_{LO} -ML model as we discussed above, and the ML-predicted results are confirmed to be highly consistent with the DFT-calculated data. As the Scenario 3 in Figure 5a, we consider generating predicted ω_{LO} from our valid ω_{LO} -ML model, and calculate other related parameters based on the predicted ω_{LO} . Then we combine these processed parameters and the original rough data to form a new input set for the S-ML model. In this way, we are able to train a new S-ML model with ML-predicted ω_{LO} rather than DFT-calculated ω_{LO} , which predicts the S values of DHPs without the calculations of their phonon properties. The trained S-ML model with predicted ω_{LO} has a satisfactory predictive performance with a high PR of 95% and a low RMSE of 2.93 (Figure 5e). This performance is comparable to the S-ML model with calculated ω_{LO} and much better than the S-ML model without ω_{LO} . To verify the actual predictive ability of the S-ML model with predicted ω_{LO} , we list the comparison results of S values between the DFT calculations and the ML predictions for each structure (Figure 5f). The high consistency of these results further supports the validation of our methodology. Notably, we propose the predicted S value of an unknown DHPs material $Cs_2CuSbCl_6$ as 56.05 based on our trained S-ML model with predicted ω_{LO} , which indicates the potential of our proposed method in exploring novel perovskite materials.

3. Conclusion

In this work, we have constructed an effective ML model for the prediction of the STEs effect in DHPs by combining DFT calculations and ML techniques based on 72 DHPs structures $Cs_2B^1B^2Cl_6$ with the atomic substitutions on the B-site cations. The bandgap values of these structures reflect that the existence of cation Cu^+ , Ag^+ , and Au^+ will lead to an obvious metallization of these semiconductors. Considering the formation mechanism of STEs in DHPs, four types of data are required in our database: atomic composition, lattice configuration, excitonic effect, and electron–phonon coupling property. For atomic compositions, we select eight features to describe the differences in various combinations of the B^1 and B^2 cations for these structures including group, period, atomic number, atomic mass, atomic radius, electronegativity, first ionization energy, and their bond length with halogen atoms. For lattice configuration, we briefly use the lattice volume to quantify the difference between these similar lattices. For the excitonic effect, we calculate the exciton binding energy E_b and the exciton Bohr radius r_b of these structures by applying the hydrogenic Rydberg model. Three crucial parameters Frohlich coupling constant α , polaron radius r_p , and polaron mass m_p are calculated to describe the strength of the electron-phonon coupling effect. Most importantly, the Huang–Rhys factor S values are calculated as the evaluation indexes of the STEs effects in these DHPs. Through the careful evaluation of different parameters, we have built an effective database containing 60 valid samples of the STEs properties in DHPs for the first time. Based on the supervised ML with the GBR algorithm, we train and screen six predictive ML models for the corresponding important physical parameters, which all exhibit strong predictive capabilities with full PR and low RMSE. The calculated CFs of these ML models reflect that B^1 cations contribute to the tuning of electron–phonon coupling strength

while B² cations are mainly responsible for electronic transfer and phonon patterns. In the end, we train a predictive S-ML model based on our prepared database (Level 1, 2, and 3 data), where the introduction of ω_{LO} displays accurate predictions. In particular, the trained S-ML model with predicted ω_{LO} displays the best balance between high predictive accuracy and low computation loadings, leading to a satisfactory predictive performance with a high PR of 95% and a low RMSE of 2.93. Moreover, the unknown *S* value of the DHPs material Cs₂CuSbCl₆ is also predicted as 56.05 by the proposed S-ML model, supporting the feasibility of our model for exploring novel perovskite structures. Therefore, the construction of the dataset and proposed ML approach are expected to accelerate the screening of potential novel STEs materials in the future, which provides opportunities for the high-throughput design of next-generation optoelectronic materials.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

The authors gratefully acknowledge the support from the National Key R&D Program of China (2021YFA1501101), the National Natural Science Foundation of China/Research Grant Council of Hong Kong Joint Research Scheme (N_PolyU502/21), the National Natural Science Foundation of China/Research Grants Council (RGC) of Hong Kong Collaborative Research Scheme (CRS_PolyU504_22), the funding for Projects of Strategic Importance of the Hong Kong Polytechnic University (Project Code: 1-ZE2V), Shenzhen Fundamental Research Scheme-General Program (JCYJ20220531090807017), the Natural Science Foundation of Guangdong Province (2023A1515012219), and Departmental General Research Fund (Project Code: ZVUL) from the Hong Kong Polytechnic University. The authors also thank the support from the Research Centre for Carbon-Strategic Catalysis (RC-CSC), Research Institute for Smart Energy (RISE), and Research Institute for Intelligent Wearable Systems (RI-IWEAR) of the Hong Kong Polytechnic University.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

density functional theory calculations, double halide perovskites, electron-phonon coupling, machine learning, self-trapped excitons

Received: July 11, 2023

Revised: August 8, 2023

Published online: August 30, 2023

- [1] M. M. Lee, J. Teuscher, T. Miyasaka, T. N. Murakami, H. J. Snaith, *Science* **2012**, *338*, 643.
- [2] Z. Ning, X. Gong, R. Comin, G. Walters, F. Fan, O. Voznyy, E. Yassitepe, A. Buin, S. Hoogland, E. H. Sargent, *Nature* **2015**, *523*, 324.
- [3] L. Lu, M. Sun, T. Wu, Q. Lu, B. Chen, B. Huang, *Nanoscale Adv.* **2022**, *4*, 680.
- [4] J. Burschka, N. Pellet, S. Moon, R. Humphry-Baker, P. Gao, M. K. Nazeeruddin, M. Grätzel, *Nature* **2013**, *499*, 316.
- [5] F. Thouin, D. A. Valverde-Chávez, C. Quarti, D. Cortecchia, I. Bargigia, D. Beljonne, A. Petrozza, C. Silva, A. R. Srimath Kandada, *Nat. Mater.* **2019**, *18*, 349.
- [6] M. Wei, F. P. G. de Arquer, G. Walters, Z. Yang, L. N. Quan, Y. Kim, R. Sabatini, R. Quintero-Bermudez, L. Gao, J. Z. Fan, F. Fan, A. Gold-Parker, M. F. Toney, E. H. Sargent, *Nat. Energy* **2019**, *4*, 197.
- [7] A. M. Sanni, S. N. Lavan, Z.-F. Liu, A. S. Rury, *J. Phys. Chem. C* **2021**, *125*, 28004.
- [8] S. Bhaumik, A. Bruno, S. Mhaisalkar, *RSC Adv.* **2020**, *10*, 13431.
- [9] A. Yangui, S. Pillet, E.-E. Bendeif, A. Lussion, S. Triki, Y. Abid, K. Boukhehdaden, *ACS Photonics* **2018**, *5*, 1599.
- [10] B. Cui, Y. Han, B. Huang, Y. Zhao, X. Wu, L. Liu, G. Cao, Q. Du, N. Liu, W. Zou, M. Sun, L. Wang, X. Liu, J. Wang, H. Zhou, Q. Chen, *Nat. Commun.* **2019**, *10*, 5190.
- [11] W.-J. Yin, T. Shi, Y. Yan, *Adv. Mater.* **2014**, *26*, 4653.
- [12] L. C. Chen, J. C. Chen, C. C. Chen, C. G. Wu, *Nanoscale Res. Lett.* **2015**, *10*, 312.
- [13] S. Luo, W. A. Daoud, *J. Mater. Chem. A* **2015**, *3*, 8992.
- [14] E. R. Dohner, A. Jaffe, L. R. Bradshaw, H. I. Karunadasa, *J. Am. Chem. Soc.* **2014**, *136*, 13154.
- [15] M. D. Smith, H. I. Karunadasa, *Acc. Chem. Res.* **2018**, *51*, 619.
- [16] M. D. Smith, A. Jaffe, E. R. Dohner, A. M. Lindenberg, H. I. Karunadasa, *Chem. Sci.* **2017**, *8*, 4497.
- [17] D. Cortecchia, S. Neutzner, A. R. S. Kandada, E. Mosconi, D. Meggiolaro, F. De Angelis, C. Soci, A. Petrozza, *J. Am. Chem. Soc.* **2017**, *139*, 39.
- [18] E. R. Dohner, E. T. Hoke, H. I. Karunadasa, *J. Am. Chem. Soc.* **2014**, *136*, 1718.
- [19] S. Ye, F. Xiao, Y. X. Pan, Y. Y. Ma, Q. Y. Zhang, *Mater. Sci. Eng.: R: Rep.* **2010**, *71*, 1.
- [20] M. J. Bowers, J. R. McBride, S. J. Rosenthal, *J. Am. Chem. Soc.* **2005**, *127*, 15378.
- [21] S. R. Li, J. J. Luo, J. Liu, J. Tang, *J. Phys. Chem. Lett.* **2019**, *10*, 1999.
- [22] J. C. Yu, J. T. Kong, W. Hao, X. T. Guo, H. J. He, W. R. Leow, Z. Y. Liu, P. Q. Cai, G. D. Qian, S. Z. Li, X. Y. Chen, X. D. Chen, *Adv. Mater.* **2019**, *31*, 9.
- [23] B. Wu, W. Ning, Q. Xu, M. Manjappa, M. Feng, S. Ye, J. Fu, S. Lie, T. Yin, F. Wang, T. W. Goh, P. C. Harikesh, Y. K. E. Tay, Z. X. Shen, F. Huang, R. Singh, G. Zhou, F. Gao, T. C. Sum, *Sci. Adv.* **2021**, *7*, eabd3160.
- [24] J. Li, H. Wang, D. Li, *Front. Optoelectron.* **2020**, *13*, 225.
- [25] K. Huang, A. Rhys, *Proc. R. Soc. London, Ser. A* **1950**, *204*, 406.
- [26] B.-B. Zhang, F. Wang, B. Xiao, Y. Xu, K. Gao, W. Jie, *J. Mater. Chem. C* **2020**, *8*, 12108.
- [27] X. Wang, W. Meng, W. Liao, J. Wang, R.-G. Xiong, Y. Yan, *J. Phys. Chem. Lett.* **2019**, *10*, 501.
- [28] W. Li, L. Hu, J. Ma, C. Jiang, S. Zhang, Y. Chen, J. Hu, X. Liu, T. Wu, D. Li, *ACS Photonics* **2022**, *9*, 2008.
- [29] F. Hao, C. C. Stoumpos, D. H. Cao, R. P. H. Chang, M. G. Kanatzidis, *Nat. Photonics* **2014**, *8*, 489.
- [30] N. K. Noel, S. D. Stranks, A. Abate, C. Wehrenfennig, S. Guarnera, A.-A. Haghighirad, A. Sadhanala, G. E. Eperon, S. K. Pathak, M. B. Johnston, A. Petrozza, L. M. Herz, H. J. Snaith, *Energy Environ. Sci.* **2014**, *7*, 3061.

