



PCCP

**Data-driven analysis of the electronic-structure factors
controlling the work functions of perovskites oxides**

Journal:	<i>Physical Chemistry Chemical Physics</i>
Manuscript ID	CP-ART-10-2020-005595.R1
Article Type:	Paper
Date Submitted by the Author:	15-Feb-2021
Complete List of Authors:	Xiong, Yihuang; Pennsylvania State University, Chen, Weinan; The Pennsylvania State University, Department of Materials Science and Engineering Guo, Wenbo; The Pennsylvania State University, College of information Sciences and Technology Wei, Hua; The Pennsylvania State University, College of Information Sciences and Technology Dabo, Ismaila; The Pennsylvania State University, Materials Science and Engineering

SCHOLARONE™
Manuscripts

Cite this: DOI: 00.0000/xxxxxxxxxx

Data-driven analysis of the electronic-structure factors controlling the work functions of perovskite oxides

Yihuang Xiong,^a Weinan Chen,^a Wenbo Guo,^b Hua Wei,^b Ismaila Dabo^{a,c}

Received Date

Accepted Date

DOI: 00.0000/xxxxxxxxxx

Tuning the work functions of materials is of practical interest for maximizing the performance of microelectronic and (photo)electrochemical devices, as the efficiency of these systems depends on the ability to control electronic levels at surfaces and across interfaces. Perovskites are promising compounds to achieve such control. In this work, we examine the work functions of more than 1,000 perovskite oxide surfaces (ABO_3) by data-driven (machine-learning) analysis and identify the factors that determine their magnitude. While the work functions of BO_2 -terminated surfaces are sensitive to the energy of the hybridized oxygen p bands, the work functions of AO-terminated surfaces exhibit a much less trivial dependence with respect to the filling of the d bands of the B-site atom and of its electronic affinity. This study shows the utility of interpretable data-driven models in analyzing the work functions of cubic perovskites from a limited number of electronic-structure descriptors.

1 Introduction

The work function measures the energy of extracting an electron from a material. Understanding trends in the work function is technologically important to thermionics¹, optoelectronics, electrochemistry, and photocatalysis^{2–8}—with one primary example being the possibility to optimize the activity of a surface by tuning its electronic affinity⁹. Perovskites are a remarkably versatile class of materials that can be synthesized with controlled purity and relatively high yield^{10–12}. Due to the interplay between their structural, chemical, and electronic characteristics, perovskites are promising candidates for achieving sensitive control of the work function. Figure 1 compares the work functions of elemental metals¹³ with those of perovskite oxides; it is apparent that perovskites show a wide distribution of work functions, providing a rich compositional space for the design of *e.g.* thermionic converters (requiring low work functions)¹⁴ and photovoltaic hole collectors (requiring high work functions).

In this work, we develop a data-driven understanding of the work functions of perovskite oxides in their prototypical cubic symmetry. For comparison, Fig. 2 shows the work functions of

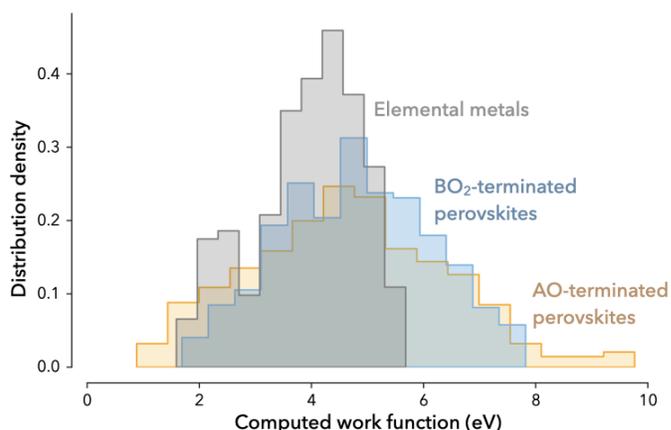


Fig. 1 Distribution of the computed work functions for elemental metals¹³ and cubic perovskites. Perovskites show a broad distribution of work functions.

^a Department of Materials Science and Engineering, and Materials Research Institute, The Pennsylvania State University, University Park, PA, USA, E-mail: YihuangXiong@psu.edu

^b College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA 16802, USA

^c Institutes of Energy and Environment, The Pennsylvania State University, University Park, PA 16802, USA

† Electronic Supplementary Information (ESI) available: A schematic of a regression tree in the random forests, feature selections based on correlation matrix and importance ranking, and projected density of states of selected perovskites. See DOI: 10.1039/cXCP00000x/

10 representative perovskites in the orthorhombic ($Pnma$) and cubic ($Pm\bar{3}m$) phases along their [001] surface facets. These results highlight a strong correlation between the work functions of these structures, indicating that the high-symmetry, cubic phase may provide a reliable basis to infer the work functions of low-symmetry, perovskite-related structures featuring octahedral rotations. Examining cubic structures is also relevant to high-entropy perovskites¹⁵ that tend to spontaneously adopt high symmetry^{15,16}. We thus present a detailed analysis of the dependence of the work functions of cubic perovskites as a function of composition and termination using extensive computational data.

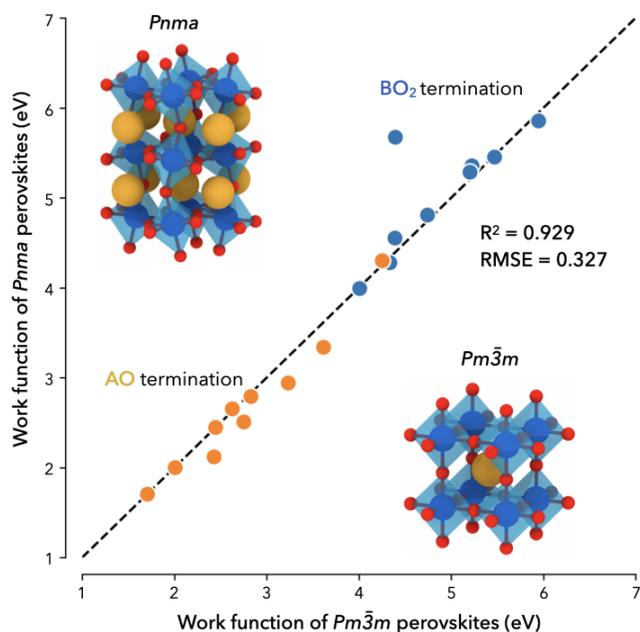


Fig. 2 Comparison of computed work functions along the [001] direction for a representative set of 10 perovskite oxides with space groups $Pm\bar{3}m$ and $Pnma$. The selected compositions are ABO_3 where $A = \text{Ca}$ or Sr and $B = \text{Ti}, \text{V}, \text{Cr}, \text{Mn}$ or Fe . The overall coefficient of determination (R^2) and root mean squared error are of 0.929 and 0.237 eV, respectively.

2 Computational method

2.1 Crystal structures and first-principles calculations

A perovskite crystal structure with formula ABO_3 is shown in Fig. 3. The B-site cation is octahedrally coordinated to oxygen, and, typically, a larger A-site cation adopts a twelve-fold coordination with the surrounding oxygen atoms. The cubic perovskite phase exists in nature (e.g. SrTiO_3 and SrVO_3), while many other lower-symmetry stable structures are also found. Compared to the ideal cubic structure, distortions such as octahedral rotations and cations displacements may occur, and some of them are responsible for functional properties such as ferroelectricity¹⁷. Nevertheless, we adopt the cubic phase as a simple template for statistical analysis, as discussed above and justified in Fig. 2.

Following Refs. 18 and 19, we select the constituent metal cations based on their propensity to form a stable cubic phase. The elements that are considered in this work are highlighted in Fig. 3. The A-site elements include the main-group metals, while the majority of the B-site elements belong to the transition metals series. Considering the alternating AO and BO_2 layers, we construct two types of interfaces along the [001] direction, as shown in Fig. 4. Using the optimized bulk structures, each slab geometry is built symmetrically with 9 ionic layers. The periodic slabs are separated by 14 Å of vacuum. Only the two outermost layers are allowed to move during geometry optimization.

All first-principles calculations are managed by the AiiDA high-throughput calculation infrastructure²⁰. The self-consistent-field calculations are performed at the semilocal Perdew–Burke–Ernzerhof (PBE) level²¹ using the PW code of the QUANTUM ESPRESSO distribution²². Ionic cores are represented by norm-

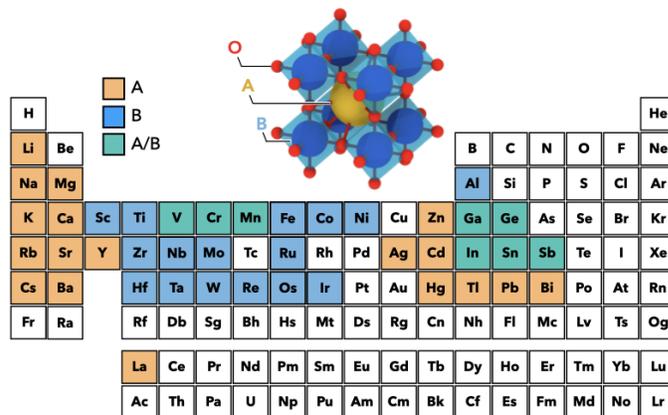


Fig. 3 A perovskite unit cell is composed of A cation in the center of the unit cell, and B cation octahedrally coordinated with the oxygen (top). The elemental compositions that are used to construct the perovskites are highlighted in the periodic table (bottom).

conserving pseudopotentials with kinetic-energy cutoffs of 100 Ry for the reciprocal-space expansion of the wave functions²³. Bulk structures are fully optimized through variable-cell optimization, while sampling the Brillouin zone with a Γ -centered Monkhorst–pack grid of $12 \times 12 \times 12$ ²⁴. For slab calculations, a Marzari–Vanderbilt cold smearing of 0.01 Ry²⁵ is employed to discretize the Brillouin zone with a reduced \mathbf{k} -points mesh of $6 \times 6 \times 1$. In addition, the ENVIRON module is applied to automatically align the Fermi level with respect to vacuum^{26–28}. The atomic positions are then fully optimized until the interatomic forces are smaller than 0.02 eV/Å.

Based on the optimized perovskite surface, we can calculate the work functions as

$$\Phi = \Phi^\circ - E_F, \quad (1)$$

where Φ° is the potential in vacuum and E_F is the Fermi energy.

Due to the semilocal PBE approximation, the calculated band gap and work function are expected to be underestimated^{13,29}. Even though previous work suggests that the work functions of metals calculated from PBE are consistent with experimental measurements^{13,30}, there is still debate about the accuracy of PBE work functions for perovskite oxides. More generally, the work functions of metal oxides can be strongly influenced by surface orientations, terminations, and defects. Studies by Ma *et al.* and Guo *et al.* showed that an accurate description of band gaps could lead to improved predictions of the work functions and band edges of semiconductors^{31,32}. Although predicting absolute work functions using the PBE functional may not be accurate, Ma *et al.*, and Chambers and Sushko showed that the PBE approximation is reliable to estimate differences in the work functions of AO and BO_2 -terminated surfaces^{32,33}. While beyond the scope of this work, it is expected that hybrid functionals such as Heyd–Scuseria–Ernzerhof (HSE)³⁴ could be more accurate for evaluating the work functions of perovskite metal oxides. Since the goal of this study is to understand trends between work functions and electronic descriptors, we argue that it is suitable to use the PBE functional.

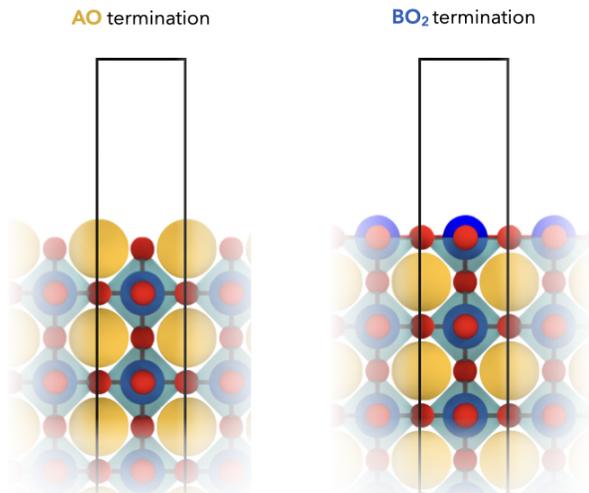


Fig. 4 Surface structures considered in this work: the AO and BO_2 terminations.

2.2 Machine-learning method and descriptor selection

On the basis of our computed datasets, we aim to identify the features that best describe the work functions of the perovskites. To achieve this, we employ a statistical learning method. We chose our model based on interpretability and performance. Here we use random forest regression³⁵, which is an ensemble statistical learning method that integrates a number of decision trees and that returns the average prediction of these trees³⁶. In specific terms, given a training set (\mathbf{X}, \mathbf{y}) where \mathbf{X} is the features and \mathbf{y} are the corresponding responses, the random forest model is trained by repeatedly sampling a subset \mathbf{x} of the training set to form the trees. The quality of the branch split is measured using the mean squared error (MSE) of the regression: $\text{MSE} = \frac{1}{N} \sum_{n=1}^N (y_n - \hat{y}_n)^2$.

For each testing sample \mathbf{x} , the prediction is obtained from the averaged prediction of the individual trees: $f(\mathbf{x}) = \frac{1}{M} \sum_{m=1}^M f_m(\mathbf{x})$, where M is the total number of trees and f_m stands for the prediction of each tree model using data \mathbf{x} . Random forests are known to be robust against overfitting, and have been widely applied for both regression and classification tasks³⁷. In addition, random forests offer means of interpreting the model using importance ranking and partial dependence analysis^{38,39}. To train the model for predicting the work function, we ‘fingerprint’ the interface structures in our database with a number of features that are physically meaningful and are expected to be correlated with the work functions. Some of the selected features have been shown previously to be critical to describe phase stability⁴⁰, thermal conductivity⁴¹, optical absorption^{42,43}, superconductivity⁴⁴, catalytic activity⁴⁵, and fuel-cell performance⁴⁶. In total 38 features are selected and summarized in Table 1.

All features, except \bar{E}_{2p}^O and $\chi_M^{\text{ABO}_3}$, are selected for both A and B elements. We computed the ionization potential IP_{calc} and the electron affinity EA_{calc} of the atoms using the energy of the half occupied Kohn-Sham orbital⁴⁹. Furthermore, the band center of orbital φ is the energy difference between the weighted center of the φ -projected band and the Fermi level in a crystal:

Table 1 Atomic descriptors that are selected in this work

Notation	Definition
$\text{IP}_{\text{expt}}, \text{EA}_{\text{expt}}, \text{IP}_{\text{calc}}, \text{EA}_{\text{calc}}$	Experimental and calculated ionization potential and electron affinity
χ_P	Pauling electronegativity
δ	Bonding covalency with oxygen
r_s, r_p, r_d	s , p and d valence orbital radii of the element ⁴⁷
$R_{\text{atom}}, \bar{R}_{\text{ion}}$	Atomic radii and averaged ionic radii
\mathcal{P}	Pettifor’s chemical scale ⁴⁸
Z	Atomic number
\mathcal{M}	Mendeleev number
\bar{E}_p	p band center in bulk perovskite
$\theta_d, \theta_{e_g}, \theta_{t_{2g}}$	Filling factor of d band, e_g and t_{2g} in bulk perovskite
\bar{E}_{2p}^O	Center of oxygen $2p$ band in bulk perovskite
$\chi_M^{\text{ABO}_3}$	Geometric mean of the electronegativity of the perovskite constituents on a Mulliken scale

$$\bar{E}_\varphi = \frac{\int_{-\infty}^{\infty} E \rho_\varphi(E) dE}{\int_{-\infty}^{\infty} \rho_\varphi(E) dE} - E_F, \quad (2)$$

and the filling factor of the φ band (similarly for e_g and t_{2g} bands) is calculated from

$$\theta_\varphi = \frac{\int_{-\infty}^{E_F} \rho_\varphi(E) dE}{\int_{-\infty}^{\infty} \rho_\varphi(E) dE}, \quad (3)$$

where ρ_φ stands for the projected density of states of the φ orbital. This projection is expressed as

$$\rho_\varphi(E) = \frac{1}{N_k} \sum_{nk,\sigma} \int |\langle \psi_{nk}^\sigma | \varphi \rangle|^2 \delta(E - E_{nk}^\sigma) dE, \quad (4)$$

where n , k and σ denote the band index, \mathbf{k} -points and the spin states of the wave function ψ , respectively.

We note that the inclusion of DFT features requires some initial bulk calculations. Constructing models using only readily available features such as elemental properties^{50–52} would overcome this requirement; however, these DFT features enable us to establish closer correlation between the work functions and electronic-structure properties, as further analyzed and discussed below.

3 Results and discussion

3.1 Random forest regression

We develop the random forest models using the SCIKIT-LEARN library⁵³. The dataset contains 1248 interface work functions and

is composed of an equal amount of AO and BO₂ surfaces. Two models are trained independently on the work functions of the AO and BO₂ terminations. Before training the models, we note that some of the selected features are correlated. Although such correlations would not impact the performance of the model, they could deteriorate its interpretability. This is because the correlated features carry similar information, thus the feature importance would be shared among them, causing a ‘dilution’ of the importance score across the feature group. Therefore, we carried out a reduction of the feature dimension using the Pearson correlation analysis, as detailed in Supplementary Figure S1. This process reduces the number of features from 38 to 21 by eliminating the most highly correlated ones. We start the analysis by using all the features to train the random forest regression models. For both AO and BO₂ terminations, we partition the dataset into 80% and 20% for training and test set. Using the training set, the hyperparameters that gives the lowest root mean squared error (RMSE) are selected. The RMSE is evaluated with fivefold cross-validations. The obtained model’s performance is then validated using only the test set. Such evaluation is repeated 40 times by shuffling the datasets to obtain an averaged total performance. By doing so, we can consistently evaluate the accuracy of the model.

We first aim to identify the features that are relevant to the work functions. This is achieved by examining their importance score. In specific terms, the importance of a feature measures how much the feature would impact the predictions. For example, we can calculate the importance of a feature by adding up the weighted variance reduction for all nodes that use this feature as the splitting feature, and then averaged over the trees in the trained forests. Based on the importance score, we perform recursive feature elimination^{43,44}, and then re-train the model each time to obtain a new importance ranking. To optimize the model’s performance when each feature is removed, the hyperparameter is re-selected using aforementioned process. The yielded model contains compatible hyperparameter and number of features. The averaged RMSE of the regression with respect to the number of features are shown in Supplementary Figures S2, along with detailed descriptions of the model constructing process. We find that the averaged RMSE of the work function of AO and BO₂ terminations are 0.468 eV and 0.531 eV, respectively. The averaged predicted work function Φ are plotted against the DFT values in Fig. 5 (a) and (b). The prediction accuracy is reasonable considering that the work functions span a range of 9 eV. For the six most predominant features that are identified in the models, we summarize their normalized feature importance in Fig. 6.

Based on the importance ranking, we find that the most relevant features for both surfaces show a consistent pattern despite the different surface structures. For the BO₂ terminations, the work function is strongly influenced by the the bulk 2p band center of oxygen \bar{E}_{2p}^O , which has an importance score of 0.44. Following that, two features that are related to the terminated element, namely r_d^B and \mathcal{P}^B , are found to be relatively important for the work functions of BO₂ terminations. This indicates that the work function of BO₂ is largely determined by its bulk properties. On the other hand, though \bar{E}_{2p}^O is still relevant to the AO work

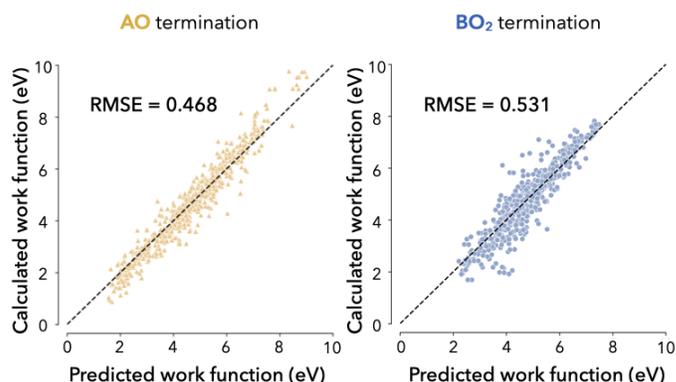


Fig. 5 Predicted versus computed work functions of AO and BO₂ terminations. The performance of the random forest regression models is evaluated by averaging the results from randomly shuffled datasets. The averaged root mean squared error for AO and BO₂ terminations are 0.468 eV and 0.531 eV, with standard deviations of 0.047 and 0.048, respectively.

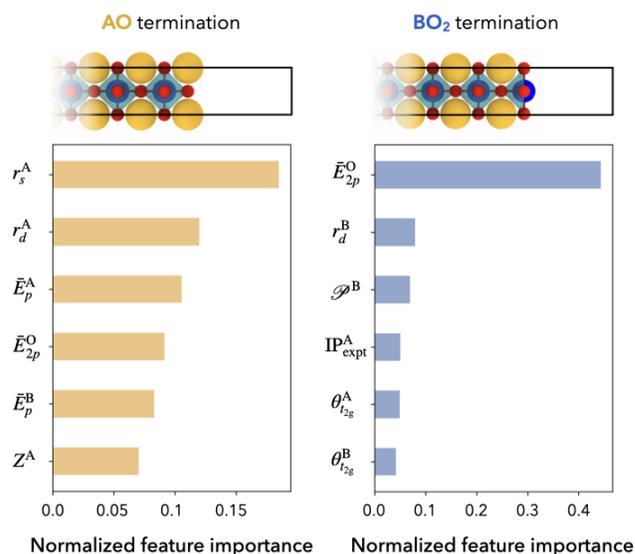


Fig. 6 Normalized feature importance of the 6 most relevant features for AO and BO₂ terminations.

functions, it only ranked as the fourth most important feature. The first three features are consistently correlated to the surface species of AO terminations; they are r_s^A , r_d^B , and \bar{E}_{2p}^O . This result shows that, in contrast to BO_2 surfaces, the surface contribution to the work function is more significant than the bulk properties for the AO terminations¹⁴. In general, the machine-learning model correctly recognizes the termination effect, where the valence orbital radii for A and B elements are predicted to be among the most essential features for the AO and BO_2 terminations. In addition, it is known that the work function is influenced by both bulk and surface properties^{14,54,55}. The machine-learning models correctly capture this dependence.

It is interesting to note that, for both AO and BO_2 terminations, the energy of the oxygen $2p$ orbital in the bulk phase plays a critical role. In fact, \bar{E}_{2p}^O is an important bulk electronic predictor that has been used to describe many electronic properties of perovskites, including vacancy formation energies⁵⁶, oxygen reduction reactivity of oxide fuel cell⁴⁶ and oxygen evolution reactivity^{57,58}. Specifically for work functions, Jacobs *et al.* have reported \bar{E}_{2p}^O as a critical descriptor by exploring 20 technologically relevant perovskite materials that are composed of Sr and La for the A atoms and $3d$ transition metals for the B atoms¹⁴. Here our data-driven approach corroborates that \bar{E}_{2p}^O remains an effective descriptor even for a wide range of metastable perovskites.

Although it helps identify the most significant features, the importance score only indicates how much the predictions are affected by the features, without explaining the specific relationship. To answer this question, we conduct a partial dependence analysis³⁹ for the two most predominant features. Partial dependence plots (PDP) illustrate the marginal effect (in the probabilistic sense) of the selected features on the predictions after integrating out the other variables. If we only focus on one specific feature \mathbf{x} , the interactions between \mathbf{x} and the response of the target can be estimated by marginalizing the predictions over all other features. This partial dependence function $\bar{f}_s(\mathbf{x})$ can be expressed as

$$\bar{f}_s(\mathbf{x}) = \frac{1}{N} \sum_{n=1}^N f(x_{1,n}, \dots, x_{s-1,n}, x, \dots, x_{M,n}), \quad (5)$$

where $\bar{f}_s(\mathbf{x})$ is approximated by averaging the output of the trained model for all features except the selected feature $x = x_s$ in the dataset. M is the total number of features in the model, and N is the total number of samples. Similarly to the previous analysis, the PDP is obtained by averaging the results using the shuffled datasets. In Fig. 7, we illustrate the PDP for both AO and BO_2 work functions with respect to \bar{O}_{2p} and the valence orbital radii r_s^A and r_d^B .

Despite different surface structures, we find that the general trend of how \bar{E}_{2p}^O influences Φ is universal, as shown in Fig. 7 (a): with a larger separation between \bar{E}_{2p}^O and the Fermi level (more negative band center of O $2p$ in bulk), the work function shows an approximately monotonic decrease. Interestingly, such correlation starts to break down for the perovskite interfaces with low work functions, where the work function reaches a plateau when the \bar{E}_{2p}^O is below -4 eV, especially for AO termination. To

explain these trends, we examine the density of states, and the correlations between \bar{E}_{2p}^O and the work function. In general, the low work function of a perovskite originates from low filling of the d bands, as shown in Fig. 8(a). One of the representative compounds of this class is SrVO_3 ¹⁴. Thus, as we move across the $3d$ transition metal series, the d bands are filled up with electrons and move down to hybridize with the O $2p$ band. This can also be understood by analyzing electron affinities. A more electronegative B site will create a more covalent bond with oxygen, thus leading to more pronounced band hybridization. A key characteristic for such hybridization is that, the band center of oxygen $2p$ is almost unchanged with respect to vacuum level [see Supplementary Fig. S5(a)]. This observation is consistent with previous literature¹⁴ and enables one to understand the linear correlation between the \bar{E}_{2p}^O and the work functions: with increasing d filling, the d bands hybridize with the O $2p$ bands and reduce the energy separation between the Fermi level and the O $2p$ band center. Since \bar{E}_{2p}^O remains almost constant with respect to the vacuum level (indicating moderate charge transfer between the inner and outer layers), a decrease in the Fermi level and an increase in the work function is observed.

Yet, we observe that the previously described correlation breaks down for deep \bar{E}_{2p}^O levels. To understand these deviations, we examined the compounds with \bar{E}_{2p}^O deeper than -4 eV and found that those perovskites primarily contain $5d$ elements, such as Ta, W and Re. By examining their projected density of states [see Supplementary Fig. S5(b)], we found that the key difference lies in the stability of the band center for O $2p$. In this case, it is observed that \bar{E}_{2p}^O is no longer constant (indicating charge transfer between the inner and outer layers), as depicted in Fig. 8(b) by the shift of \bar{E}_{2p}^O towards the vacuum energy level. This trend explains the loss of correlation between the \bar{E}_{2p}^O level and the work function of those compounds. We further study this trend by examining the partial dependence of the work function with respect to the \bar{E}_{2p}^O and r_s^A for the AO termination in Fig. 7(b). It is apparent that when \bar{E}_{2p}^O is above -4 eV, the isocontours align horizontally, which indeed confirms the strong correlation between \bar{E}_{2p}^O and the work functions. In contrast, the isocontours are mostly vertical when \bar{E}_{2p}^O is deeper than -4 eV.

Next, we turn our attention to the influence of the valence orbital radii on the work function. We first discuss the AO termination, and we highlight different groups of elements in Fig. 7 (c). We find that the work function can be parsed into three regions: (1) $r_{sp}^A < 1.0$ bohr, (2) 1.0 bohr $< r_s^A < 1.3$ bohr, and (3) $r_s^A > 1.3$ bohr. In fact, these three regions correspond to alkali/alkaline-earth metals, transition metals, and post-transition metals and metalloids. The valence orbital radii have been shown to capture the periodic trends⁴⁷, except for Li and Na due to their small radii. For elements belonging to the family of post-transition metals and metalloids, the work function tends to decrease with respect to the increase of the valence orbital radii. This can be explained in terms of the electronegativity: alkali/alkaline-earth metals (larger r_s^A) show lower electronegativity compared to that of post-transition metals and metalloids (smaller r_s^A), thereby yielding smaller work functions. In addition, because of our

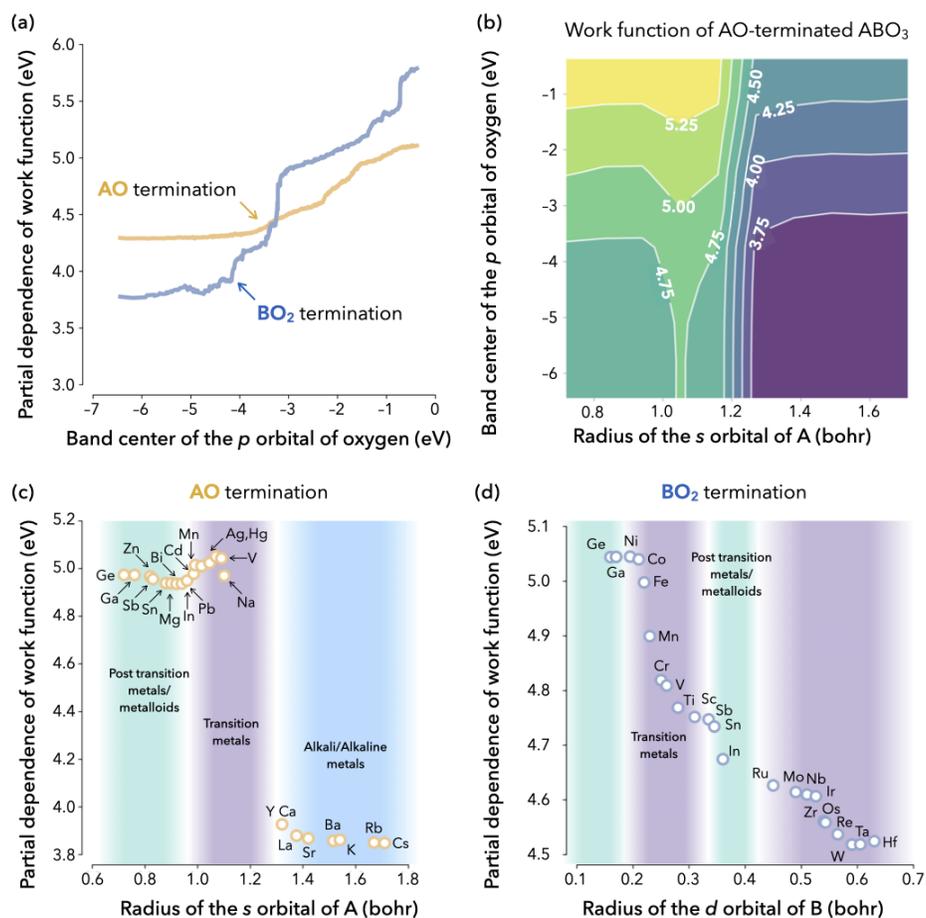


Fig. 7 Partial dependence plots for the AO and BO_2 terminations. (a) Dependence of the work function with respect to the orbital energy \bar{E}_{2p}^O for both interfaces. (b) Two-variable dependence plots of the work functions of both orbital energy \bar{E}_{2p}^O and the orbital radius r_s^A for AO-terminated interfaces. (c, d) Partial dependence of the work function with respect to the radius of the p (d) orbital of the A (B) element for the AO (BO_2) termination. The colored regions represent groups of elements in the periodic table.

choice of A cation across the periodic table, we observe a clear separation of the work function between surfaces that are terminated with alkali/alkaline-earth metals and the post-transition metals/metalloids. The low work function of the alkali/alkaline terminated perovskites make them potential candidates for designing thermionic converters.

We can now discuss the trend between the size of the d orbital radii of the B atoms (r_d^B) and the work function of BO_2 surfaces with similar arguments. Figure 7(d) shows that the work functions also decrease with r_d^B . The increasing r_d^B radii reflect a decrease in electronegativity, thus causing a diminution in the work function. Compared to the PDP of the AO surface, we do not observe significant separation in the dependence of the work functions as a function of BO_2 surface interactions. This is likely due to the fact that the B cations are mainly composed of the transition metals and metalloids, with no alkaline/alkali metals included. These partial dependence analyses reveal that the work function is determined by both bulk electronic properties and surface electronegativity. By controlling compositions and structures, these two effects can be leveraged simultaneously to design materials

with desired work functions.

In closing, we underscore the practical importance of our statistical observations. A low work function is a crucial requirement for designing electron emitters and thermionic energy converters, and we find here that perovskites with alkali or alkaline-earth metals at the A site are promising candidates for these applications, as shown in Fig. 7. We conclude that although the center of the oxygen $2p$ band is a sensitive descriptor of the work function for a number of perovskites, AO-terminated surfaces with shallow Fermi energy are much better described by the orbital radii of the A-site elements. This analysis demonstrates the possibility of optimizing surface structure and chemistry to effectively reduce the work function for *e.g.* thermionic energy conversion.

4 Conclusions

We have examined the work functions of cubic perovskites by statistical means. We have constructed a database of perovskites and have employed a random forest regression to predict their work functions, achieving predictive accuracy with only a few features included. Two central features that primarily control the per-

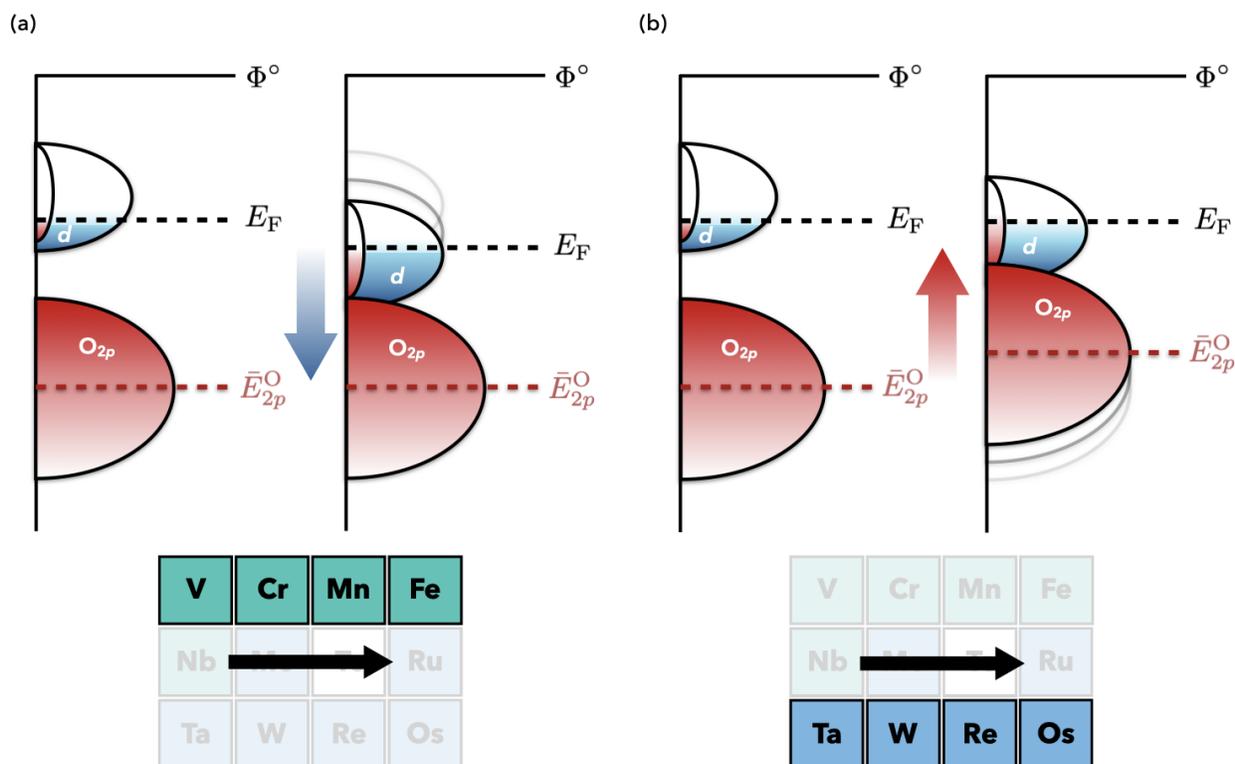


Fig. 8 Schematized densities of states for perovskites that contain (a) 3d transition metals and (b) 5d transition metals. Φ° stands for the energy level of vacuum. The red and blue regions correspond to the p and d bands of oxygen and of the B-site transition metal, respectively. The black and red dash lines represent the Fermi level and the O 2p band center.

ovskite work functions have been identified: the oxygen 2p band center and the valence orbital radii of the surface-terminating cations. The oxygen 2p band center is found to be crucial to the determination of the BO_2 -termination work functions, while r_s^A predominantly influences the AO-termination work functions. We have explained how those electronic descriptors affect the work functions of perovskites using partial dependence analysis, and have found that the general trends are related to the stability of oxygen energy levels and atomic electronegativities. These correlations may benefit the search for metal oxides with desired surface electronic properties. For instance, optimizing the compositions of the perovskites to achieve deep oxygen 2p band centers while simultaneously terminating the interface with alkali or alkaline-earth elements may yield optimally low work functions, which is essential for thermionics. Conversely, the perovskites that have shallow oxygen 2p band centers, coupled with p -block metal or metalloid terminations, may be of interest for designing hole collectors.

Conflicts of interest

The authors declare no competing conflicts of interest.

Acknowledgements

Y. X. and I. D. acknowledge primary financial support from the National Science Foundation under grant DMREF-1729338. W. C. acknowledges support from the National Science Founda-

tion under grant DMR-2011839 through the MRSEC Center for Nanoscale Science of the Pennsylvania State University. The computational work was performed using the Roar supercomputer of the Penn State Institute for Computational and Data Sciences.

Data availability

The work-function and descriptor databases, and the code repository are available online at <https://github.com/yyx5048/Cubic-Perovskite-WorkFunction>.

Notes and references

- 1 S. Yamamoto, *Reports on Progress in Physics*, 2005, **69**, 181–232.
- 2 I. E. Castelli, F. Hüser, M. Pandey, H. Li, K. S. Thygesen, B. Seger, A. Jain, K. A. Persson, G. Ceder and K. W. Jacobsen, *Adv. Energy Mater.*, 2015, **5**, 1400915.
- 3 A. K. Singh, J. H. Montoya, J. M. Gregoire and K. A. Persson, *Nat. Commun.*
- 4 Y. Xiong and I. Dabo, *Phys. Rev. Materials*, 2019, **3**, 065801.
- 5 S. Trasatti, *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry*, 1972, **39**, 163 – 184.
- 6 A. R. Zeradjanin, A. Vimalanandan, G. Polymeros, A. A. Topalov, K. J. J. Mayrhofer and M. Rohwerder, *Phys. Chem. Chem. Phys.*, 2017, **19**, 17019–17027.
- 7 Y. Song, D. Johnson, R. Peng, D. K. Hensley, P. V. Bonnesen, L. Liang, J. Huang, F. Yang, F. Zhang, R. Qiao, A. P. Baddorf,

- T. J. Tschaplinski, N. L. Engle, M. C. Hatzell, Z. Wu, D. A. Cullen, H. M. Meyer, B. G. Sumpter and A. J. Rondinone, *Sci. Adv.*, 2018, **4**, year.
- 8 J. Hwang, R. R. Rao, L. Giordano, Y. Katayama, Y. Yu and Y. Shao-Horn, *Sci.*, 2017, **358**, 751–756.
- 9 T. Liu, C. Xi, C. Dong, C. Cheng, J. Qin, S. Hu, H. Liu and X.-W. Du, *J. Phys. Chem. C*, 2019, **123**, 28319–28326.
- 10 R. E. Schaak and T. E. Mallouk, *Chemistry of Materials*, 2002, **14**, 1455–1471.
- 11 L. Zhang, Y. Zhou, L. Guo, W. Zhao, A. Barnes, H.-T. Zhang, C. Eaton, Y. Zheng, M. Brahlek, H. F. Haneef, N. J. Podraza, M. H. W. Chan, V. Gopalan, K. M. Rabe and R. Engel-Herbert, *Nat. Mater.*, 2016, **15**, 204–210.
- 12 R. C. Haislmaier, E. D. Grimley, M. D. Biegalski, J. M. LeBeau, S. Trolrier-McKinstry, V. Gopalan and R. Engel-Herbert, *Adv. Funct. Mater.*, 2016, **26**, 7271–7279.
- 13 R. Tran, X.-G. Li, J. H. Montoya, D. Winston, K. A. Persson and S. P. Ong, *Surf. Sci.*, 2019, **687**, 48 – 55.
- 14 R. Jacobs, J. Booske and D. Morgan, *Adv. Funct. Mater.*, 2016, **26**, 5471–5482.
- 15 C. Oses, C. Toher and S. Curtarolo, *Nat. Rev. Mater.*, 2020, **5**, 295–309.
- 16 J. G. N. Z. J. N. M. Q. T. H. K. V. Sicong Jiang, Tao Hu and J. Luo, *Scr. Mater.*, 2018, **142**, 116 – 120.
- 17 N. A. Benedek and C. J. Fennie, *J. Phys. Chem. C*, 2013, **117**, 13339–13349.
- 18 I. E. Castelli and K. W. Jacobsen, *Model. Simul. Mater. Sci. Eng.*, 2014, **22**, 055007.
- 19 A. A. Emery and C. Wolverton, *Sci. Data*, 2017, **4**, 170153.
- 20 G. Pizzi, A. Cepellotti, R. Sabatini, N. Marzari and B. Kozinsky, *Comput. Mater. Sci.*, 2016, **111**, 218–230.
- 21 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.
- 22 P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, *J. Condens. Matter Phys.*, 2009, **21**, 395502.
- 23 A. M. Rappe, K. M. Rabe, E. Kaxiras and J. Joannopoulos, *Phys. Rev. B*, 1990, **41**, 1227.
- 24 H. J. Monkhorst and J. D. Pack, *Phys. Rev. B*, 1976, **13**, 5188.
- 25 N. Marzari, D. Vanderbilt, A. De Vita and M. Payne, *Phys. Rev. Lett.*, 1999, **82**, 3296.
- 26 O. Andreussi and N. Marzari, *Phys. Rev. B*, 2014, **90**, 245101.
- 27 I. Dabo, B. Kozinsky, N. E. Singh-Miller and N. Marzari, *Phys. Rev. B*, 2008, **77**, 115139.
- 28 Y. Li and I. Dabo, *Phys. Rev. B*, 2011, **84**, 155127.
- 29 N. E. Singh-Miller and N. Marzari, *Phys. Rev. B*, 2009, **80**, 235407.
- 30 A. Stroppa and G. Kresse, *New J. Phys.*, 2008, **10**, 063020.
- 31 Z. Guo, F. Ambrosio, W. Chen, P. Gono and A. Pasquarello, *Chem. Mater.*, 2018, **30**, 94–111.
- 32 T. Ma, R. Jacobs, J. Booske and D. Morgan, *APL Materials*, 2020, **8**, 071110.
- 33 S. A. Chambers and P. V. Sushko, *Phys. Rev. Materials*, 2019, **3**, 125803.
- 34 A. V. Krukau, O. A. Vydrov, A. F. Izmaylov and G. E. Scuseria, *J. Chem. Phys.*, 2006, **125**, 224106.
- 35 A. Liaw and M. Wiener, *R news*, 2002, **2**, 18–22.
- 36 S. R. Safavian and D. Landgrebe, *IEEE Trans. Syst. Man Cybern.*, 1991, **21**, 660–674.
- 37 S. Weisberg, *Applied linear regression*, John Wiley & Sons, 2005, vol. 528.
- 38 L. Breiman, *Machine Learning*, 2001, **45**, 5–32.
- 39 J. H. Friedman, *Ann. Stat.*, 2001, **29**, 1189–1232.
- 40 J. Schmidt, J. Shi, P. Borlido, L. Chen, S. Botti and M. A. L. Marques, *Chem. Mater.*, 2017, **29**, 5090–5103.
- 41 A. van Roekeghem, J. Carrete, C. Oses, S. Curtarolo and N. Mingo, *Phys. Rev. X*, 2016, **6**, 041061.
- 42 G. Pilania, A. Mannodi-Kanakithodi, B. P. Uberuaga, R. Ramprasad, J. E. Gubernatis and T. Lookman, *Sci. Rep.*, 2016, **6**, 19375.
- 43 S. Lu, Q. Zhou, Y. Ouyang, Y. Guo, Q. Li and J. Wang, *Nat. Commun.*, 2018, **9**, 3405.
- 44 V. Stanev, C. Oses, A. G. Kusne, E. Rodriguez, J. Paglione, S. Curtarolo and I. Takeuchi, *Npj Comput. Mater.*, 2018, **4**, 29.
- 45 J. H. Montoya, A. D. Doyle, J. K. Nørskov and A. Vojvodic, *Phys. Chem. Chem. Phys.*, 2018, **20**, 3813–3818.
- 46 Y.-L. Lee, J. Kleis, J. Rossmeisl, Y. Shao-Horn and D. Morgan, *Energy Environ. Sci.*, 2011, **4**, 3966–3970.
- 47 A. Zunger, *Phys. Rev. B*, 1980, **22**, 5839–5872.
- 48 D. Pettifor, *Solid State Communications*, 1984, **51**, 31 – 34.
- 49 L. M. Ghiringhelli, J. Vybiral, S. V. Levchenko, C. Draxl and M. Scheffler, *Phys. Rev. Lett.*, 2015, **114**, 105503.
- 50 L. Ward, A. Agrawal, A. Choudhary and C. Wolverton, *npj Computational Materials*, 2016, **2**, 16028.
- 51 F. Ren, L. Ward, T. Williams, K. J. Laws, C. Wolverton, J. Hattrick-Simpers and A. Mehta, *Sci. Adv.*, 2018, **4**, year.
- 52 W. Li, R. Jacobs and D. Morgan, *Comput. Mater. Sci.*, 2018, **150**, 454–463.
- 53 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.
- 54 Z. Zhong and P. Hansmann, *Phys. Rev. B*, 2016, **93**, 235116.
- 55 S. Halas and T. Durakiewicz, *Vacuum*, 2010, **85**, 486 – 488.
- 56 T. T. Mayeshiba and D. D. Morgan, *Solid State Ion.*, 2016, **296**, 71 – 77.
- 57 B. Han, M. Risch, Y.-L. Lee, C. Ling, H. Jia and Y. Shao-Horn, *Phys. Chem. Chem. Phys.*, 2015, **17**, 22576–22580.
- 58 R. Jacobs, J. Hwang, Y. Shao-Horn and D. Morgan, *Chem. Mater.*, 2019, **31**, 785–797.