

Advances and challenges in DFT-based energy materials design

Jun Kang(康俊)[†], Xie Zhang(张燮)[‡], and Su-Huai Wei(魏苏淮)[§]

Citation: Chin. Phys. B, 2022, 31 (**10**): 107105. DOI: 10.1088/1674-1056/ac89d7
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

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

Interfacial defect engineering and photocatalysis properties of hBN/MX_2 (M = Mo, W, and X = S, Se heterostructures

Zhi-Hai Sun(孙志海), Jia-Xi Liu(刘佳溪), Ying Zhang(张颖), Zi-Yuan Li(李子源), Le-Yu Peng(彭乐宇), Peng-Ru Huang(黄鹏儒), Yong-Jin Zou(邹勇进), Fen Xu(徐芬), and Li-Xian Sun(孙立贤)
Chin. Phys. B, 2022, 31 (**6**): 067101. DOI: 10.1088/1674-1056/ac43b2

Topological properties of Sb(111) surface: A first-principles study

Shuangxi Wang(王双喜) and Ping Zhang(张平)
Chin. Phys. B, 2022, 31 (**4**): 047105. DOI: 10.1088/1674-1056/ac46bc

Tunable electronic properties of GaS-SnS_2 heterostructure by strain and electric field

Da-Hua Ren(任达华), Qiang Li(李强), Kai Qian(钱楷), and Xing-Yi Tan(谭兴毅)
Chin. Phys. B, 2022, 31 (**4**): 047102. DOI: 10.1088/1674-1056/ac3a62

A density-wave-like transition in the polycrystalline V_3Sb_2 sample with bilayer kagome lattice

Ningning Wang(王宁宁), Yuhao Gu(顾雨豪), M. A. McGuire, Jiaqiang Yan, Lifen Shi(石利粉), Qi Cui(崔琦), Keyu Chen(陈科宇), Yuxin Wang(王郁欣), Hua Zhang(张华), Huaixin Yang(杨槐馨), Xiaoli Dong(董晓莉), Kun Jiang(蒋坤), Jiangping Hu(胡江平), Bosen Wang(王铂森), Jianping Sun(孙建平), and Jinguang Cheng(程金光)
Chin. Phys. B, 2022, 31 (**1**): 017106. DOI: 10.1088/1674-1056/ac4227

First-principles study on improvement of two-dimensional hole gas concentration and confinement in AlN/GaN superlattices

Huihui He(何慧卉) and Shenyuan Yang(杨身园)
Chin. Phys. B, 2022, 31 (**1**): 017104. DOI: 10.1088/1674-1056/ac00a0

TOPICAL REVIEW — Celebrating 30 Years of *Chinese Physics B*

Advances and challenges in DFT-based energy materials design

Jun Kang(康俊)[†], Xie Zhang(张燮)[‡], and Su-Huai Wei(魏苏淮)[§]

Beijing Computational Science Research Center, Beijing 100193, *China*

(Received 20 June 2022; revised manuscript received 12 August 2022; accepted manuscript online 16 August 2022)

The growing worldwide energy needs call for developing novel materials for energy applications. *Ab initio* density functional theory (DFT) calculations allow the understanding and prediction of material properties at the atomic scale, thus, play an important role in energy materials design. Due to the fast progress of computer power and development of calculation methodologies, DFT-based calculations have greatly improved their predictive power, and are now leading to a paradigm shift towards theory-driven materials design. The aim of this perspective is to introduce the advances in DFT calculations which accelerate energy materials design. We first present state-of-the-art DFT methods for accurate simulation of various key properties of energy materials. Then we show examples of how these advances lead to the discovery of new energy materials for photovoltaic, photocatalytic, thermoelectric, and battery applications. The challenges and future research directions in computational design of energy materials are highlighted at the end.

Keywords: density functional theory, materials design, energy applications

PACS: 71.15.Mb, 71.15.Nc, 71.20.−b **DOI:** 10.1088/1674-1056/ac89d7

1. Introduction

To sustain the growing worldwide energy needs, there are strong demands for developing novel materials for energy applications, such as photovoltaics, photocatalysts, thermoelectrics, and battery materials. Such developments require explorations beyond known materials and deep into the chemical space. Since first formulated in 1964–1965,[1,2] density functional theory (DFT) calculations have had an increasing important impact on materials design.[3,4] Material-specific studies during the 1970s and 1980s demonstrated the feasibility to predict material properties with *ab initio* DFT calculations at the atomic scale.[5–8] In the late 1980s, the developments in exchange–correlation functionals greatly improved the accuracy of DFT, and many DFT calculation packages appeared, such as FLAPW[9,10] and CASTEP.[11] Since then, DFT became widely accepted and was applied to a broad range of materials from molecules to solids.

Historically, materials design is usually carried out in a "trial-and-error" manner, inspired by empirical intuitions and rules obtained from known materials. In this paradigm, computations provide retrospective insights into experiments. Experimental groups synthesize and characterize a material and computations help to understand the experimental observations. Theorists take a given material to study its properties and then inform experimentalists if they find a suitable material with good properties for potential applications. With rapid developments over the past six decades, the continuously improving predictive power of DFT is now leading to a paradigm change in collaborations between computation and experiment in modern materials science, towards "materials-by-design".[12,13] In this new paradigm, the desired functionality is specified first, and DFT calculations are then employed to construct structure-property relationship and to predict completely new and synthesizable materials with the target properties, followed by an experimental validation. Through iterative design and data-exchange cycles between computation and experiment, the accuracy of the prediction model improves, greatly accelerating materials design.

The change of the role of DFT calculations in materials design is, of course, facilitated by the growing computing power, which follows the Moore's law; but more importantly, it should be attributed to the advances in DFT methodologies that greatly increase the accuracy and enhance the capability of DFT simulations through removal of long-standing approximations and/or extending the simulation scale. For energy materials design, the key physical quantities of concern include stability, band gap, density of states, optical absorption, carrier mobility, thermal conductivity, defect concentration, and ionization, excited-state dynamics, carrier relaxation, and recombination, *etc*. The construction of reliable prediction models relies on accurate DFT calculations of these key quantities, without which the "materials-by-design" approach is impossible.

The objectives of this perspective are to introduce the methodological advances in the predictive capabilities of DFT for energy materials design, to show examples of how these advances accelerate the discovery of new energy materials,

[†]Corresponding author. E-mail: jkang@csrc.ac.cn
[‡]Corresponding author. E-mail: xiezhang@csrc.ac.cn
[§]Corresponding author. E-mail: suhuaiwei@csrc.ac.cn

© 2022 Chinese Physical Society and IOP Publishing Ltd

and to discuss challenges and future research directions in computational design of energy materials. We will present state-of-the-art DFT methods for accurate simulation of various key properties of energy materials, including total energy-related properties, band structure, optical and excited-state properties, transport properties, as well as defect and alloy properties. Then we will provide examples to show how the advances of the DFT-based computational techniques have helped the discovery of novel photovoltaic, photocatalytic, thermoelectric, and battery materials. Finally, we conclude with an outlook on the remaining challenges in DFT methods for energy materials design.

2. DFT methods for accurate property prediction

The main idea of DFT can be summarized as follows:[1] (i) every observable (including energy) of a quantum mechanical system can be written as a unique functional of electron density; and (ii) the ground-state electron density minimizes the total energy functional. The most popular implementation of DFT is the Kohn–Sham (KS) scheme.[2] It assumes that the ground-state electron density of any interacting system can be reproduced by a non-interacting system, and all of the complexities of the electron–electron interaction effects are attributed to the exchange–correlation (xc) energy E_{xc}, resulting in the well-known single-particle KS equation. In principle, the KS-DFT is exact. However, the exact form of E_{xc} is unknown yet, and it has to be calculated with approximations. The predictive power of a KS DFT calculation largely depends on the approximations adopted. Development of reliable xc functionals is one of the major issues of DFT research. So far, there have been a huge number of approximated functionals proposed. The levels of sophistication and complexity for different approximations are often represented by the rungs of a ladder, namely, Jacob's Ladder of DFT[14] as schematically shown in Fig. 1. The three lower rungs are local or semilocal functionals. The local density approximation (LDA)[2,15] employs only the local value of electron density (ρ), while the generalized gradient approximation (GGA)[16] adds the gradient of ρ, and meta-GGA[17,18] further adds the kinetic energy density. The hybrid functionals[19,20] on the fourth rung include a portion of exact exchange energy. On the top rung of the ladder, namely, the random phase approximation (RPA) functionals, information of the unoccupied KS orbitals is also taken into consideration.

In the remaining part of this section, we discuss how the advances in methodology developments eventually improve the predictive power of DFT calculations for the key properties of energy materials, such as total energy related properties, band structure, optical and excited-state properties, transport properties, and defect and alloy properties.

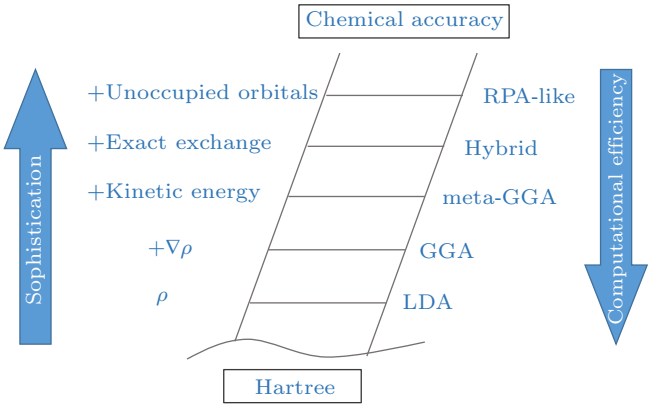


Fig. 1. The Jacob's ladder of DFT.

2.1. Total energy related properties

The ground-state total energy is a key quantity that can be directly computed from DFT calculations. The accuracy of the calculated total energy is, to a large extent, a reflection of the predictive power of DFT methods. Extensive tests have been carried out on the performance of various xc functionals in predicting the total energy and related properties such as lattice constant, cohesive energy, and bulk moduli across a broad range of materials.[21–23] A general trend is that the error of the calculated total energy significantly reduces over time, with 3.7 times reduction from 1980s to 2010s,[23] reflecting the advances in the xc functionals (it is also noted recently that, in addition to energy, accurate prediction of electron density should also play an important role in xc functional development[24]). Although the functionals on higher rungs of Jacob's Ladder are in principle more accurate, this is not guaranteed, and in practice the performance can be case-dependent.

In Fig. 2 the errors of different xc functionals in the predicted lattice constants, bulk moduli, and cohesive energies for 64 strongly bound solids were plotted (the data were taken from Ref. [21]). It is seen that even the rung 1 LDA can already give reasonable lattice constants, typically with an underestimation of 1%–2%. However, LDA tends to over-bind atoms, resulting in severely overestimated cohesive energy by $\sim$ 20%. The GGA functionals such as the one formulated by Perdew, Burke, and Ernzerhof (PBE)[16] reduce the over-binding of LDA, mostly by more accurate prediction of total energy of isolated atoms and molecules, but with a slight over-correction. In general, to compare the relative structural stability of a compound, GGA prefers the structure with larger volume, whereas LDA favors the structure with smaller volume. The GGA-predicted lattice constants are typically $\sim$ 1%–2% larger than experimental values, and the cohesive energy is underestimated by $\sim$ 5%. GGA is also found to be more realistic for magnetic solids compared to LDA.[25] Overall, the total energy related properties predicted at the GGA level are

fairly good, although higher-rung functionals may have further improvements. In addition, there could be error cancellation when comparing energy differences between different materials. Thus, for strongly bound solids, considering the balance between accuracy and efficiency, the GGA-level functional is often a reasonable starting point. For instance, it has been shown that the phase diagrams for the Cs–Pb–Cl system calculated with GGA and the hybrid functional are quite similar.[26]

Nevertheless, meta-GGA or hybrid functionals may be necessary for specific cases. For example, no GGA can be excellent for both solids and molecules at the same time,[22] and meta-GGA like SCAN[18] should be more suitable for systems involving both finite and infinite systems. The hybrid functionals partially remove the self-interaction error in LDA/GGA and tend to localize electrons, leading to better descriptions of materials with highly localized and strongly correlated d or f electrons. For example, compared to LDA/GGA, hybrid functionals significantly improve the predicted magnetic ground state and magnetic coupling constants of transition-metal monoxides.[27,28] They can also correctly describe the formation of intrinsic polaronic defects in strongly correlated systems like TiO_2.[29] Moreover, hybrid functionals are superior in predicting the band gap, as will be discussed later.

Despite their successes, a major shortcoming of the commonly used functionals at rungs 1–4 is that the long-range van der Waals (vdW) dispersion interactions are formally not included. As a result, for weakly bound systems where dispersion interactions are significant, none of these non-dispersion corrected functionals is able to give qualitatively correct total energies in most cases.[22] There have been many approaches proposed to include the long-range dispersion interaction. One way is adding a posteriori correction, similar to an atomic force field, on top of standard functionals. Examples are the DFT-D[30] and DFT-TS[31] methods. An alternative approach is treating the long-range interaction as explicit nonlocal correlation functionals of the density, such as the vdW-DF[32,33] and rVV10[34] methods. It is also possible to fully account for the dispersion interaction with high-level methods like quantum Monte Carlo (QMC)[35] and RPA to the correlation energy,[36,37] but the applications are currently limited by their high computational costs. With these approaches, the dispersion interactions in weakly bound materials can be described quite well. For instance, the typical error in binding energies of layered materials predicted by the SCAN + rVV10 method is $\sim 2\%$.[38] The accuracy of various posteriori vdW corrections and nonlocal vdW functionals was systematically examined in Ref. [39], with test sets including rare gases, layered solids, and molecular solids. It is concluded that the rev-vdW-DF2 functional leads overall to the best performance.

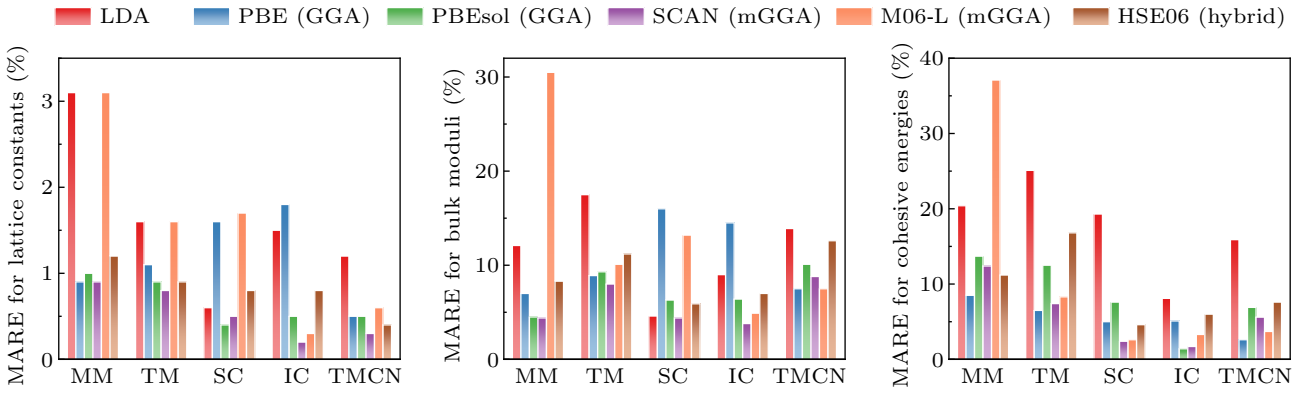


Fig. 2. The mean absolute errors (MARE) in calculated lattice constants, bulk moduli, and cohesive energies using different xc functionals for 64 solids including main-group metals (MM), transition metals (TM), semiconductors (SC), ionic crystals (IC), and transition metal carbides and nitrides (TMCN). Plotted with the data taken from Ref. [21].

When accurate total energy is available, the force and stress can be computed from its derivatives according to the Hellmann–Feynman theorem,[40,41] with a proper treatment of the Pulay force/stress in case the basis is atomic-position dependent. With the total energy, force, and stress, a vast number of materials properties could be evaluated. Structural optimizations can be performed by minimizing the energy or force/stress. The force constant matrix can be obtained by the finite-displacement method[42] or perturbation theory,[43] and used to compute phonon spectrum and vibrational properties. By a careful search of the potential energy surface with methods like nudged elastic band[44] or dimer,[45] one can obtain the minimum energy path and energy barrier for a chemical reaction. Taking all these total energy related properties into account, it is possible to evaluate the stability and efficiency of materials for energy applications.[46,47]

2.2. Band structure

2.2.1. Band gap

In practice, the lower-rung semilocal xc functionals like LDA and GGA are often adopted for a broad range of materials, due to the balance of accuracy and computational efficiency. In many cases they yield reasonable band structures (thus density of states). However, the fundamental band

gaps predicted by these functionals, which are usually determined by the eigen-energy difference between the highest-occupied and the lowest-unoccupied KS orbitals, suffer from severe underestimation (see Fig. 3). The typical error is 30%–100%, and sometimes a metallic band structure is predicted for a semiconductor. Correspondingly, the predicted conduction-band minimum (CBM) and valence-band maximum (VBM) positions are incorrect. This band-gap problem of the semilocal functionals is attributed to the lack of derivative continuity[48] and the delocalization error[49] in the xc functionals, which are different for different atomic orbitals. Therefore, the band-gap problem is most severe if the VBM and CBM have very different atomic orbital characters.

One of the solutions to the DFT band-gap problem is hybrid functionals [such as the one developed by Heyd, Scuseria, and Ernzerhof (HSE)[20]], which mix the LDA/GGA functional with exact (and possibly screened) Hatree–Fock (HF) exchange energy. The localization error of the HF exchange largely compensates the delocalization error of LDA/GGA,[49] therefore resulting in much improved band gap. The typical value of the mixing ratio of the HF exchange in the HSE hybrid functional is 0.25, and it can be further tuned to better reproduce experimental band gaps, *e.g.*, using smaller mixing ratio for more covalent materials and larger ratio for more ionic materials.

A more rigorous method is the many-body perturbation theory under the *GW* approximation,[50,51] which goes beyond the single-particle DFT framework. In this method the xc functional in DFT is effectively replaced by a nonlocal, energy-dependent self-energy Σ, which is given by integrating over the product of the electronic Green's function G and the screened interaction W. In principle, the *GW*-calculation should be performed self-consistently, and the final result is independent of the starting point, which is often provided by a DFT calculation. However, due to numerical cost and algorithmic difficulties, practical calculations are often carried out with the partially self-consistent GW_0 scheme where only the G is evaluated iteratively, or with the one-shot, non-self-consistent G_0W_0 scheme. In these schemes, the results could be dependent on the starting point. In general, the G_0W_0 method works quite well. It has been shown that hybrid-functional calculations could be a good starting point,[52] resulting in a typical error of 2.3% for band gap. While different levels of self-consistency may sometimes lead to better results than G_0W_0, systematic improvements over a broad range of materials are not guaranteed. For example, for extended solids, fully self-consistent *GW* tends to overestimate the band gap due to an underestimation of the dielectric screening, thus the vertex correction in W is important.[53–55]

In Fig. 3 we plotted the calculated band gaps for the SC40 solid test set[20] using PBE, HSE, and G_0W_0, as well as the experimental values. The data are collected from various literatures.[56–61] It can be clearly seen that the performance of PBE is poor. The slope of the corresponding linear fitting between theoretical and experimental results is only 0.68, far away from the ideal value of 1. HSE and G_0W_0 show significant improvements, with the slopes close to 1, and G_0W_0 performs slightly better.

While providing quite accurate band structures, a major disadvantage of the hybrid functional and *GW* methods is their high computational costs. Compared to an LDA/GGA calculation, the required computation time is one to two and two to three orders of magnitude longer for hybrid-functional and *GW* calculations, respectively. Many methods have been developed to correct the band-gap problem at a moderate computational cost, such as the DFT + U method,[62] meta-GGA-type TB-mBJ potentials,[63] the generalized Delta self-consistent-field method,[64] modification of pseudopotentials,[65] the DFT-1/2 method,[66] and the Wannier–Koopman method.[67] These methods have shown great successes in predicting the band gaps, but may also suffer from certain disadvantages like the lack of self-consistency, the absence of a total-energy formalism, and the ad hoc nature. Careful tests are always recommended when they are applied to a particular material of interest.

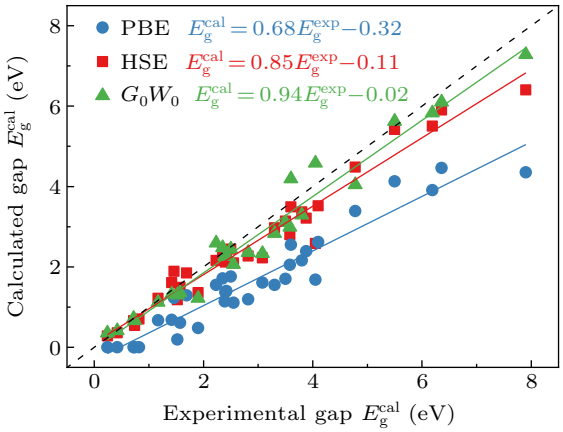


Fig. 3. Calculated and experimental band gaps for the SC40 solid test set.[20] The data were taken from Refs. [56–61] . Solid lines are linear fits. The dashed line corresponds to $E_g^{cal} = E_g^{exp}$.

2.2.2. Deformation potential and band alignment

Once correct band gap and eigenvalues of band edges are available, it is possible to compute other important band-related properties such as deformation potentials and band alignment between materials. The deformation potentials reflect the strength of electron–phonon coupling and determine the eigen-energy shift due to volume change. The band alignment between materials affects the carrier distribution and transport across the interface, as well as the dopability of semiconductors. A critical issue in computing these properties is finding a common energy reference, since the DFT-calculated eigenvalues for different material systems cannot be compared directly due to the periodic boundary conditions used in these calculations.

For low-dimensional materials, the vacuum level is a natural choice for the energy reference,[68] but it is absent in bulk-material calculations. For bulk materials, a widely adopted method is carrying out an explicit calculation for a heterostructure supercell constructed by the two materials of interest.[69] The atomic core-level or the averaged electrostatic potential at the two sides of the heterostructure can serve as the common energy reference to determine the relative position of the band edges of the two materials. Because of the periodic boundary condition, lattice-match is required at the heterostructure interface when constructing the supercell. If the volume deformation effect is not considered, then this method is only applicable to materials with negligible lattice mismatches. A correction taking into account the absolute deformation potential was later proposed,[70] resulting in improved band offset for materials with small lattice mismatches. Based on the transitivity rule, a three-step approach[71] and its modification[72] were developed to deal with materials with large lattice mismatches and low symmetries. For materials with dissimilar structures, the direct construction of a heterostructure supercell is difficult. In this case, one can construct an intermediate structure that well interfaces with both materials, and a searching algorithm for such structures was reported in Ref. [73]. With these advanced methods, the band offset between materials can be accurately predicted, in good agreement with experiments.

2.3. Optical and excited state properties

2.3.1. Optical properties

The optical property of a material is basically the response of its electron to the light-induced, time-dependent electromagnetic perturbation. The linear optical properties, such as absorption spectrum, reflectivity, and loss function, can all be deduced from the frequency-dependent complex dielectric function.[74] Under the RPA of polarization, which assumes independent particle transitions, one can directly use the DFT-calculated wavefunctions and eigenvalues to compute the matrix elements for optical transitions and then evaluate the dielectric function.[75,76] With LDA/GGA, the static macroscopic dielectric constant is often quite reasonable, but the absorption spectrum is not satisfactory. This can be attributed to the poor description of electron–electron interaction by LDA/GGA and the lack of electron–hole interaction in RPA. The inaccuracy in the electron–electron interaction results in incorrect band gap and eigenvalues as discussed above. This can be greatly improved through the hybrid-functional or *GW* methods. The lack of electron–hole interaction can lead to large errors for materials with significant exciton effects, especially for low-dimensional materials where the Coulomb screening is weak.[77] This effect can be included from full solution of the Bethe–Salpeter equation (BSE)[78] built on top of the *GW* approximation.[51,79,80] For example, the $GW+\mathrm{BSE}$ approach well reproduces the first peak of the imaginary part of the dielectric function of bulk BN, which is not well captured with $GW+\mathrm{RPA}$, as shown in Fig. 4(a).[81]

Another class of methods which can be used to calculate the optical properties is based on the time-dependent density functional theory (TDDFT),[82,83] which extends DFT to the time domain. For the time-dependent case, Runge and Gross proved that the one-to-one correspondence between external potential and charge density remains hold, establishing the foundation of TDDFT.[84] According to the linear-response formalism of TDDFT (LR-TDDFT), the inverse dielectric function for a periodic solid can be obtained from the full response function,[51,85,86] with the latter expressed by the KS response function and a proper xc kernel. The xc kernel is defined as the functional derivative of the time-dependent xc potential with respect to the time-dependent density. A correct long-range behavior of the xc kernel is critical for periodic systems.[87,88] Alternatively, within LR-TDDFT the optical spectrum can be computed by solving the so-called Casida equation,[82] which treats electric excitations as the eigenmodes of the system.[89] This approach usually works quite well for finite systems like molecules. Beyond the linear response regime, real-time TDDFT (RT-TDDFT) directly propagates the time-dependent KS equation.[90] With RT-TDDFT, the optical spectrum can be computed with the Fourier transform of the time-dependent dipole moment (for isolated systems) or current (for periodic systems) under an external field.[91,92]

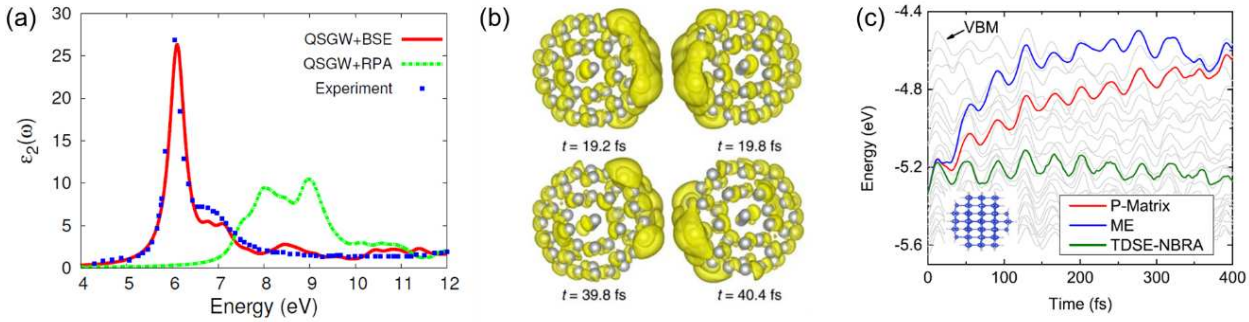


Fig. 4. (a) Imaginary part of the macroscopic dielectric function for bulk hexagonal BN calculated with $GW+\mathrm{RPA}$ and $GW+\mathrm{BSE}$. Reproduced with permission from Ref. [81]. Copyright 2018 American Physical Society. (b) RT-TDDFT study of plasmon in an Ag_{55} nanocluster. Reproduced from Ref. [93] in accordance with the Creative Commons Attribution (CC BY) license. Copyright 2015 Springer Nature. (c) NAMD simulation of hot carrier cooling in an Si quantum dot. Reproduced with permission from Ref. [94]. Copyright 2020 American Physical Society.

2.3.2. Excited-state properties

RT-TDDFT provides time-domain evolutions of electronic wavefunctions together with ionic motions, thus it is a powerful tool to study the dynamics of various photo-excited processes [Fig. 4(b)], for example, ultrafast dynamics of photo-excited carriers,[95] optically excited structural transitions,[96] photochemical reaction,[97] light-induced demagnetization,[98] and plasmonic processes.[93] Strategies have also been developed to include critical features like decoherence, detailed balance and trajectory branching properties into the RT-TDDFT Ehrenfest dynamics.[99] For accurate evolution of wavefunctions, RT-TDDFT usually requires a quite small simulation time step at the attosecond scale, which limits the simulation time scale. A great acceleration can be achieved with a recently developed Hamiltonian interpolation technique, which increases the time step to 0.1 fs–0.5 fs.[100]

To further extend the simulation size and time scales of excited carrier dynamics, nonadiabatic molecular dynamics (NAMD) methods are widely adopted, with either the Ehrenfest dynamics[101] or the surface hopping formalism.[102] For condensed matter systems, NAMD is often combined with the neglect of the back-reaction approximation (also called classic path approximation), in which the effect of the wavefunction evolution on the nuclear movement is explicitly ignored, and the nuclear trajectory is generated from the conventional ground-state Born–Oppenheimer molecular dynamics.[103] This leads to considerable computational saving and makes the simulation of systems with several hundred atoms feasible. Many DFT-based NAMD methods with detailed balance and decoherence effect have been proposed, such as DISH[103] and *P*-Matrix,[94] with successful applications in studying hot carrier dynamics like relaxation[104,105] and transfer,[106,107] as shown in Fig. 4(c). NAMD simulation with electron–hole interactions for exciton dynamics has also been realized recently.[108]

2.4. Transport properties

The performance of energy materials strongly depends on their carrier and phonon transport properties. For example, the challenge in improving the figure of merit of thermoelectric materials is how to increase the carrier conductivity and reduce the thermal conductivity at the same time.[109] The design of efficient energy materials thus calls for accurate prediction of transport properties.

2.4.1. Carrier transport

For perfect crystals, the diffusive carrier transport is limited by phonon scattering. In case diffusive transport is dominant, the phonon-limited mobility can be calculated by solving the Boltzmann transport equation (BTE) for the electronic distribution function.[110] Due to its complex integro-differential nature, direct solution of BTE is difficult, thus various approximations are often used to simplify the BTE. One approximation is using a constant relaxation time, either taken as an empirical parameter,[111] or estimated through the effective mass approximation combined with the deformation potential theory,[112,113] but the results are merely qualitative rather than quantitative due to the neglect of detailed scattering events.

A more accurate approach is computing the scattering rate and state-dependent relaxation time from *ab initio* electron–phonon coupling (EPC) matrix elements.[110,114] This is quite challenging, especially for semiconductors, since a very dense $\boldsymbol{q}$ grid in the reciprocal space is required to obtain converged scattering rates. For instance, $\sim 10^6$ $\boldsymbol{q}$-points in the first Brillouin zone are needed for Si.[115] The difficulty is overcome by two technical advances. One is the density functional perturbation theory (DFPT), which allows efficient calculations of electron–phonon coupling matrix elements at arbitrary $\boldsymbol{q}$ vectors using a primitive unit cell.[43] The other one is the Wannier interpolation method, with which one can interpolate the matrix elements from more coarse grids.[116]

The approach that evaluates the mobility with EPC from full DFT calculations was first reported for bulk silicon in 2009,[117] and has been successfully extended to other materials in recent years.[118,119] A direct solution of the BTE beyond the relaxation time approximation has also been demonstrated with iterative algorithms.[115,120] Besides phonon scattering, another factor limiting the carrier mobility in realistic materials is the Coulomb scattering due to ionized impurities. This effect can also be included in the BTE, with the impurity scattering potential either described by analytical models,[120] or computed self-consistently with DFT.[117] Considering both phonon and impurity scatterings, the mobility and conductivity can be well described by DFT calculations, as shown in Fig. 5(a).

When the size of materials goes down to the nanoscale, the dominant transport mechanism is not carrier diffusion but quantum tunneling. Nanoscale devices are often modeled by a central scattering region connected to left and right electrodes, and the carrier transport is described as a one-dimensional coherent scattering process for the incoming electrons from the electrodes. The problem can be solved by either combining DFT calculations with the non-equilibrium Green's function method using localized basis sets,[121–123] or by direct calculation of the scattering states using planewave basis sets under auxiliary periodic boundary conditions.[124,125] The latter method, together with a linear scaling algorithm,[126] has been applied to large-scale systems with 4000 atoms.[127]

Carrier transport can also occur between highly localized states, such as the cases of many organic semiconductors and disordered/defective structures. This process can be described by a hopping model using the Marcus theory.[128–130] The required electronic coupling strength and reorganization energy can be obtained from DFT calculations. For disordered/defective structures where a very large supercell is needed, *ab initio* multiscale simulations in combination with the Marcus theory have been demonstrated, as illustrated in Fig. 5(b).[131,132]

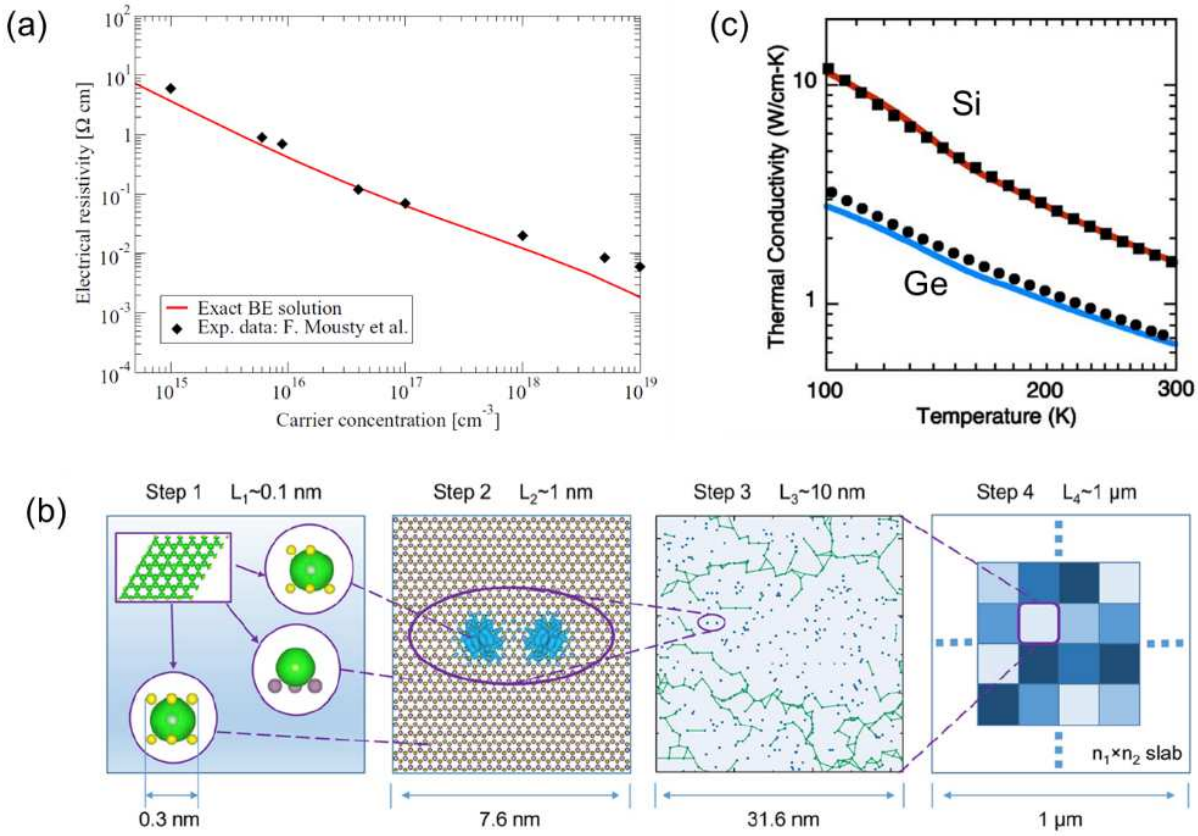


Fig. 5. (a) The calculated electrical resistivity of Si at 300 K as a function of carrier concentration from exact solution of BTE. Reproduced with permission from Ref. [120]. Copyright 2016 American Physical Society. (b) A multiscale model used to study the hopping transport in defective 2D material combining DFT calculations with the Marcus theory. Reproduced with permission from Ref. [132]. Copyright 2020 American Chemical Society. (c) Lattice thermal conductivities of Si and Ge calculated by solving phonon Boltzmann equation. Blue and red lines are calculated values, and squares and dots are experimentally measured values. Reproduced with permission from Ref. [133]. Copyright 2007 AIP Publishing.

2.4.2. Thermal conductivity

Thermal conductivity is an important fundamental property of materials. For semiconductors, lattice thermal conductivity dominates, and now it can be theoretically predicted by advanced DFT methods. For perfect crystals at low temperatures, thermal conductivity can be accurately calculated by solving the BTE for phonon transport with phonon spectrum and anharmonic interactions from DFT, as shown in Fig. 5(c).[133–135] The size effect in this method is minor thus one can use a small unit cell. However, it is not suitable for cases where anharmonicity is rather complex, such as amorphous materials and liquid which are highly disordered, or high-temperature conditions under which higher-order anharmonicity becomes important.

Another class of methods is based on *ab initio* molecular dynamics (MD), in either a nonequilibrium or an equilibrium manner.[136] In principle *ab initio* MD simulation contains full anharmonic interactions and can be applied to arbitrary temperatures. Disorder can also be included with reasonable construction of supercells. In the nonequilibrium MD approach one simulates the heat flux directly using a sustained temperature gradient, and then obtains the thermal conductivity by applying Fourier's law.[137,138] However, it suffers from huge size effects since a rather large supercell is needed to get a reasonably small temperature gradient. Therefore, extrapolation to infinite supercell length is usually required.[139]

In contrast, the equilibrium MD-based Green–Kubo (GK) method has a weak system-size dependence.[140,141] It relates the thermal conductivity with the heat current autocorrelation function (HC-ACF) through the dissipation-fluctuation theorem. Due to the ill-definition of heat current in DFT calculations of periodic systems, for a long time, the applications of the GK method were limited to force-field Hamiltonians. This ill-definition problem has been overcome recently by several works. One approach is evaluating atom-decomposed DFT energy and then using a two-step integration method in MD.[142] Other solutions include converting the position operator into the well-defined commutator between the position operator and the Hamiltonian,[143] and quantum-mechanically defining a stress tensor for each atom.[144]

2.5. Defect and alloy properties

2.5.1. Defect properties

Defect control is one of the most important issues for energy materials. In the past few decades, the basic formalisms of point-defect modeling through DFT calculations have been established with great successes. The formation energy $\Delta H_{\rm f}$

of a defect characterizes how easily it can be formed. Under thermodynamic equilibrium conditions, the defect concentration $n_{\rm D}$ at temperature T follows the Boltzmann relation, $n_{\rm D} = N\exp(-\Delta H_{\rm f}/k_{\rm B}T)$, where N is the available lattice sites in the host per volume. As illustrated in Fig. 6, the formation energy $\Delta H_{\rm f}(\alpha,q)$ for a particular defect α with charge state q is calculated by[145,146]

$$\Delta H_{\rm f}(\alpha,q) = E(\alpha,q) - E({\rm host}) + \sum_i n_i\mu_i + q(E_{\rm VBM} + E_{\rm F}), \quad (1)$$

where $E(\alpha,q)$ and $E({\rm host})$ are the total energies of the defective and pristine materials in a supercell, respectively. $E_{\rm VBM}$ is the host VBM energy, and $E_{\rm F}$ is the Fermi energy referenced to VBM. n_i is the number of atom i exchanged with the reservoirs to create the defect in the host. The chemical potential μ_i involves the effects of growth environments. Thermodynamic equilibrium growth requires that the host material is stable, and there is no elemental precipitates or secondary phases, thus μ_i is confined in a finite range by these conditions. By considering surfaces and edges also as "defects", the formula of defect formation energy can be further extended to calculate surface/edge energy.[147] It should be noted that when determining the available range of the chemical potentials, it is critical to consider all possible competing phases and dissociation paths, otherwise the stability of the host material would be overestimated,[148,149] and the trends of the defect formation energy might be qualitatively wrong. The defect charge transition energy level is defined as the $E_{\rm F}$ at which the formation energies of a defect in two charge states q and q' are equal, namely, $\varepsilon(q/q') = [E(\alpha,q) - E(\alpha,q') + (q-q')E_{\rm VBM}]/(q'-q)$. With $\Delta H_{\rm f}$ and ε, the equilibrium Fermi level and carrier density can be self-consistently evaluated by considering charge neutrality condition.[150] Computational frameworks to automate the calculations of these defect properties have been developed recently.[151–153]

A major impact of defect is that it can capture carriers and cause nonradiative carrier recombination. The carrier capture is a multi-phonon assisted electronic transition process.[154–156] The capture rate can be computed from DFT calculations,[157] and used to identify detrimental defects.[158,159] The key issue here is the calculation of electron–phonon coupling matrix for a large supercell containing defects. A direct calculation using the finite displacement method or DFPT is expensive. Efficient calculations can be realized by a variational approach[160] or using a one-dimensional effective phonon mode.[161] There have also been applications of NAMD to the calculations of carrier capture rates, but a proper rescaling of the carrier and/or defect densities to realistic values is critical.[162]

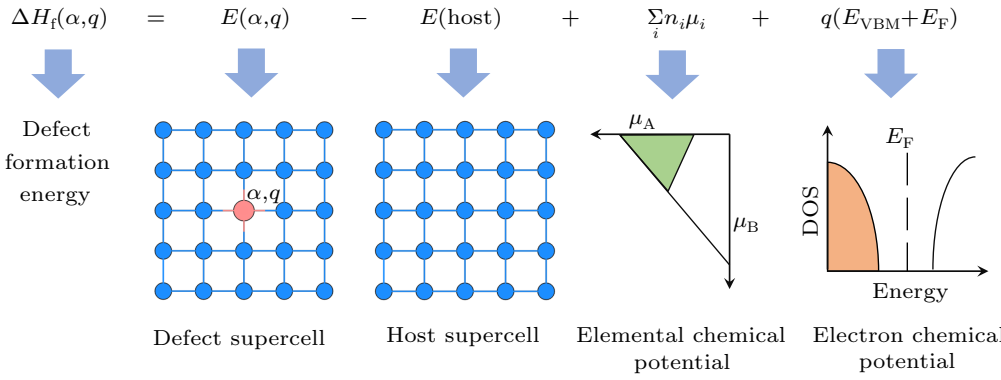


Fig. 6. A schematic illustration of defect formation energy calculation.

To achieve high accuracy in practical defect simulations, it is important to adopt a proper level of DFT methods [bare-DFT, hybrid, *GW*, spin–orbit coupling (SOC), *etc.*] that can correctly describe the structural and electronic properties of the material of interest. For example, to accurately calculate the structure and transition levels of certain intrinsic defects in halide perovskites, one has to use high-level xc functionals together with the SOC effect to obtain proper band-edge positions, otherwise there could be qualitative errors.[26,163–166]

Another important issue is how to appropriately take into account the finite-size effect for charged defects. This effect originates from the unphysical electrostatic interaction among the defect, its periodic images, and the compensating background jellium charge. The resulting error can be huge for systems with weak dielectric screening, and sometimes may even lead to divergence in the calculated formation energy (such as the case of 2D materials.[167]) For bulk materials, widely used corrections include the Makov–Payne scheme (MP),[168] the Lany–Zunger scheme (LZ),[169] and the Freysoldt–Neugebauer–Van de Walle scheme (FNV).[170] Their performances have been systematically investigated.[171] The MP correction works well for defects with point-like charge density, but leads to overcorrections for defects with more extended charge density. The LZ approach undercorrects the

formation energy for point-like defect states, but improves the results for defects with less localization. The FNV scheme leads to overall meaningful formation energies, and it is also extended to low-dimensional systems.[172] However, all these three methods adopt a macroscopic dielectric constant/profile, which is adjustable and does not give information of microscope screening effects. There have been efforts to avoid the use of such parameters by using a rigorous defect screening model,[173] or self-consistent potential correction.[174] Especially, by replacing the virtual jellium state with realistic host band-edge states, a physical, straightforward, and dimension-independent universal model to calculate the formation energy of charged defects was recently demonstrated.[175]

2.5.2. Alloy properties

Most of the calculation methods described above are for materials with periodic boundaries. For the simulation of random or disordered alloys without a periodic boundary, a key issue is how to construct a proper supercell to describe the chemical disorder so that standard DFT calculation methods can be applied. The simple approaches are the virtual crystal approximation (VCA),[176] which describes the alloys using single type of hybrid atoms that are composition-averaged mixture of atoms in the parent compounds, and coherent-potential approximation (CPA),[177] where all atoms with the same chemical types are assumed to be equivalent and each is embedded in a structureless averaged uniform medium. The VCA or CPA is computationally efficient since one can use a parent primitive cell for the calculation, but at the price of ignoring any possible short-range order and local atomic distortion, which could be critical in determining alloys properties.[178] These effects can be captured by the special quasi-random structures (SQS).[179,180] The SQS are periodic finite-size supercells, in which the atomic correlations between neighboring sites best match those of the fully random structures, and have been successfully applied to predict various alloy properties, such as elastic constants,[164] band bowing,[181] and doping.[182] Similar approaches can also be used for partially disordered alloys with short-range order.[183]

3. DFT-guided design of energy materials

The above DFT-based approaches, with increasing accuracy and predictive power, allow efficient design and exploration of novel energy materials. Very often this process is combined with high-throughput screening towards target properties and/or predictive models constructed via machine learning. In this section we first briefly discuss the typical routes for computational materials design and then present representative examples of energy materials design guided by advanced DFT calculations.

3.1. Routes for computational materials design

There are three typical routes for computational materials design: (i) knowledge-based materials design, (ii) high-throughput materials design, and (iii) genetic algorithm-based materials design. We illustrate the three routes in more detail in the following.

3.1.1. Knowledge-based materials design

This is a conventional approach to materials design. In order to design novel materials for a specific application, we focus on one or few of the prototypical materials that have been identified and demonstrated for the application in earlier experiments. Then, we thoroughly compute and analyze their basic properties with the aim to establish the correlations between structure, composition, and properties. Using the identified correlations, we could then design new structures and/or new compositions that yield certain properties required by the application of interest.

3.1.2. High-throughput materials design

With the rapid developments in DFT calculations and computer power, the basic materials properties become more easily accessible, leading to the construction of numerous materials databases, *e.g.*, the Materials Project.[184] Very often, it is already clear that for a specific application what the requirements for the basic properties are. Hence, we could build up a series of selection/filtering criteria, and then use them to screen a large materials database, which would likely lead to a small set of materials candidates for the target application. For these newly identified materials, we can perform more refined calculations and narrow down to a few of the most promising candidates, which can be transferred to experimentalists for verification and demonstration.

3.1.3. Genetic algorithm-based materials design

High-throughput materials screening relies largely on the availability of materials databases. Even if one could compute such a database of their own, one needs input structural and compositional information from established structural databases, *e.g.*, the Inorganic Crystal Structure Database (ICSD). However, when these databases are not available, there is also an alternative approach that one can use for the design of novel materials, which is based on the genetic algorithm (GA). Inspired by Darwin's idea of evolution, we may think of the design of novel materials for a target application as finding the ideal solutions to a difficult optimization problem (that is defined by an optimizable objective function).[185]

Starting from a population of materials candidates (if unknown, they could also be random structures or compositions), selection, crossover, and mutation operations can be applied to generate a new generation of materials. Keep repeating this procedure, the objective function gets optimized, which yields promising material candidates after many generations.

Although we discussed the three routes separately, we note that the three routes are not mutually exclusive. In fact, they can be well combined to accelerate materials design. For instance, the construction of selection criteria relies on knowledge extracted from extensively studied known materials. Very often, even before the high-throughput screening, it might be already clear that some material compositions can be simply excluded based on common sense or established knowledge. In this case, it is unnecessary to perform first-principles calculations, which might save lots of time and effort. The identified materials from high-throughput screening may also require detailed computations and investigations, which may serve as a feedback loop that allows to optimize the screening process. For the GA-based route, the initial population may also come from the candidates identified from the first two routes.

Another issue worth noting is that it could be beneficial to introduce certain degree of flexibility to the criteria of material screening in high-throughput calculations. For example, a material is typically considered to be unstable if the formation energy is calculated to be positive. However, there could be inaccuracy in DFT calculations, leading to an overall shift of the computed formation energy. In this case the calculated trend in the formation energy is more informative, rather than the absolute value. In addition, for materials with positive but small formation energies, experimental synthesis remains possible.[186,187] Thus, a reasonable choice of the screening criteria based on common sense or established knowledge of systematic calculation errors is important for efficient materials design.

3.2. Examples for DFT-guided design of energy materials

In this section, we present selected examples for DFT-guided design of various types of energy materials, including photovoltaic, photocatalytic, thermoelectric, and battery applications.

3.2.1. Photovoltaic materials

Solar absorber plays a vital role for the efficiency of a photovoltaic device. The design of stable and efficient solar absorbing materials is key in realizing high-performance solar cells. We first present an example to illustrate how we can make use of established knowledge to design novel efficient solar absorbers.

Yu *et al.*[188] tackled this challenge by analyzing a key metric — the spectroscopic limited maximum efficiency (SLME)[189] for various I–III–VI materials. One interesting finding was that comparing to the case of Tl occupying the III sites and Cu occupying the I sites in I–III–VI compounds, the SLME increases when Tl changes from high-valence Tl^{3+} to low-valence Tl^{+} by occupying some of the I sites. Careful investigations revealed that the origin of this enhancement is because for the materials with low-valence Tl, there is a considerable contribution of Tl s orbital hybridization in the VBM, and the CBM is characterized by Tl p orbitals. As the major character of the VBM is still Cu d states, both d $\rightarrow$ p and s $\rightarrow$ p transitions are possible, which enhances solar absorption.

Tl is rare and toxic, but the knowledge established here inspired to design Cu–V–VI materials, in which group-V elements could have both $+5$ and $+3$ oxidation states, mimicking the effect of Tl. Hence, Yu *et al.* investigated the SLMEs for different Cu–V–VI compounds, through which they found seven compounds that have comparable or higher SLMEs than $CuInSe_2$ [Fig. 7(a)].

To further validate the theoretical predictions, they compared the theoretically computed band gaps [Fig. 7(b)] and absorption spectra [Fig. 7(c)] against experimental results. The band gaps exhibited only an average of 0.15-eV deviation from the experimental gaps. The absorption profiles were also similar between theory and experiment, which further demonstrated that Sb in the low-valence state leads to a strong absorption.

The above example nicely illustrates how to employ knowledge derived from careful investigations of known materials to design novel high-absorption compounds. However, in this example we have not taken into account another two requirements for a high-performance solar absorber: stability and tolerance to defects. We use another two examples to demonstrate these two aspects.

Lead-halide perovskites are known for their highly efficient solar absorption. However, they suffer from severe instabilities and toxicity due to the presence of Pb. To overcome these issues, Zhao *et al.*[190] proposed to search for stable and non-toxic halide perovskites through cation transmutation. As shown in Fig. 8(a), one can double the formula of a cubic perovskite $APbX_3$ and then replace the two Pb^{2+} cations with one metallic cation in the $1+$ oxidation state and another in the $3+$ oxidation state, forming the so-called double perovskite. Such atomic transmutation can greatly enhance the Coulomb energy of the system, thus making the materials more stable.

To have optimal absorption efficiencies, Zhao *et al.* focused on materials with band gaps in the range of 0.8 eV–2.0 eV [Fig. 8(b)]. For good thermodynamic stabilities, they excluded materials with negative decomposition enthalpies. In addition, they took into account electron and hole effective masses, both of which should be smaller than $1.0m_0$. The small carrier effective masses are beneficial for ambipolar carrier conduction and carrier extraction. Exciton binding energies were limited to be lower than 100 meV in order to avoid the formation of excitons. Using these criteria, they identified 14 candidate materials. Excluding three compounds that contain toxic element Tl, 11 optimal compounds were found.

The absorption spectra of the 11 compounds were calculated and depicted in Fig. 8(c). Despite the differences in the absorption onset, all of these materials exhibit strong absorption profiles that compare well with that of the prototypical Pb-based hybrid perovskite $CH_3NH_3PbI_3$ (or $MAPbI_3$). Among the 11 compounds, there are two materials with ideal direct band gaps: $Cs_2InSbCl_6$ (1.02 eV) and $Cs_2InBiCl_6$ (0.91 eV), for which they further computed the SLMEs as shown in Fig. 8(d). Clearly, the SLMEs of these two materials are similar to that of $MAPbI_3$, demonstrating their promise in being highly efficient solar absorbers.

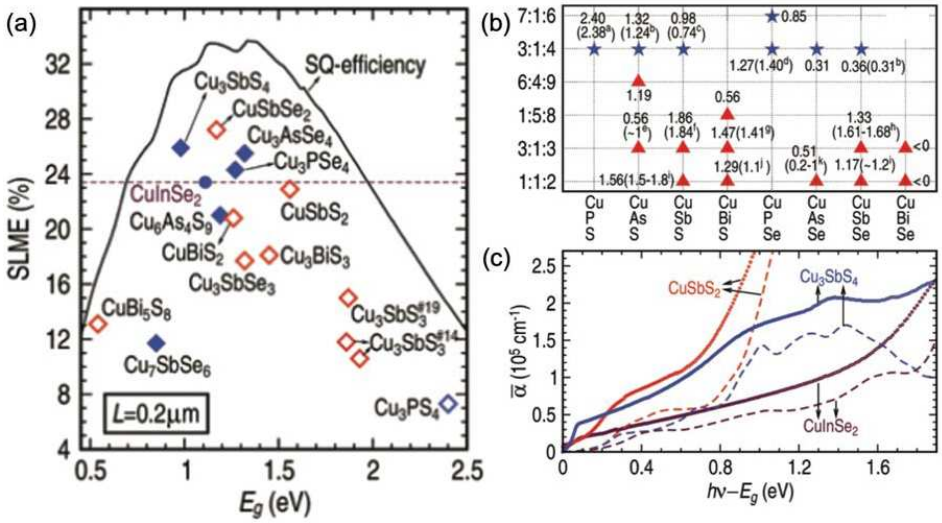


Fig. 7. (a) Theoretically computed SLMEs for Cu–V–VI compounds. (b) Accurate band gaps for Cu–V^{3+}–VI (red triangles) and Cu–V^{5+}–VI (blue stars) compounds with V = P, As, Sb, Bi, and VI = S, Se, computed using G_0W_0, which are compared with available experimental data in parenthesis. (c) Theoretically computed (dashed lines) and experimental (open circles) absorption spectra for $CuSbS_2$, Cu_3SbS_4, and $CuInSe_2$. Reproduced with permission from Ref. [188]. Copyright 2013 Wiley.

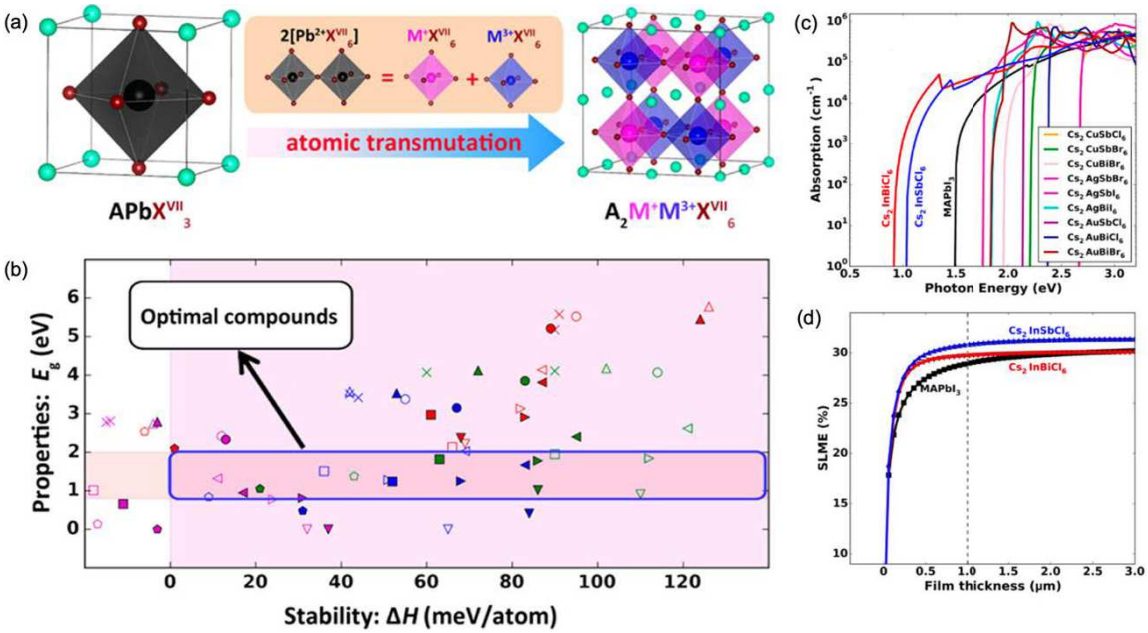
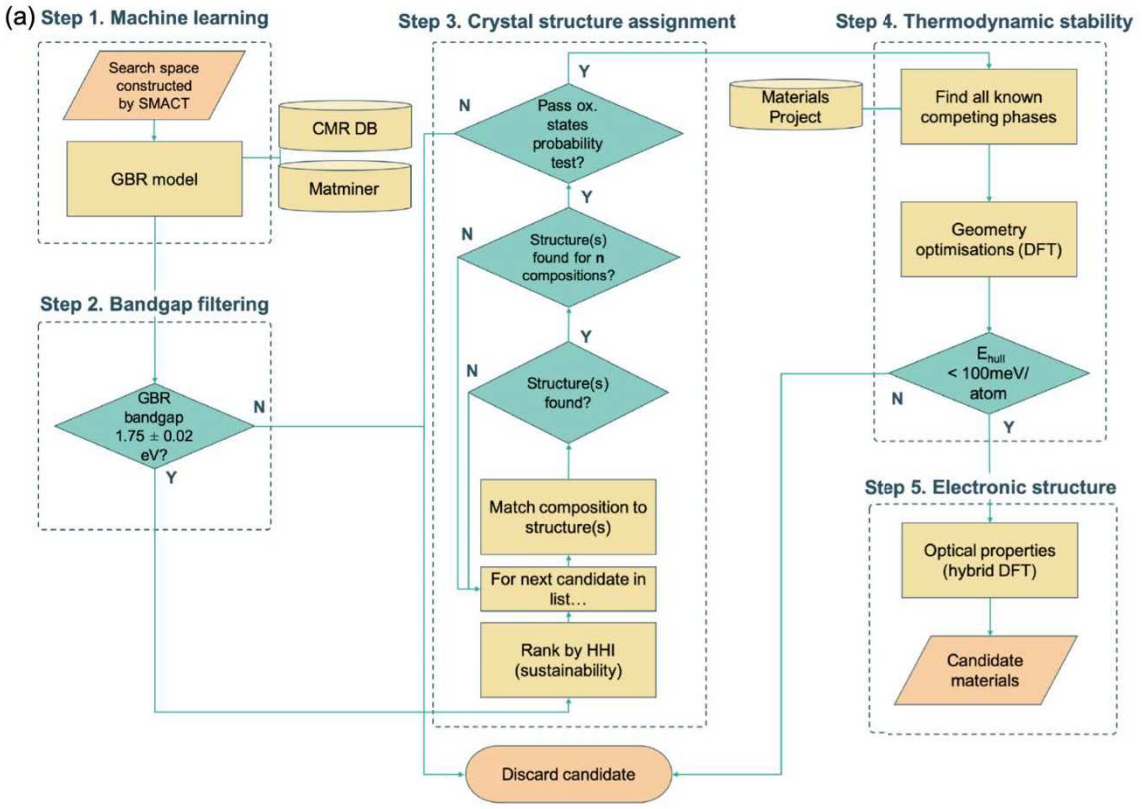


Fig. 8. (a) Candidate double-perovskite materials obtained through atomic transmutation. (b) Identification of optimal compounds based on band gap and stability. (c) Computed absorption spectra of the optimal compounds obtained from the screening. (d) Simulated SLMEs of $Cs_2InSbCl_6$ and $Cs_2InBiCl_6$ in comparison with $CH_3NH_3PbI_3$. Reproduced with permission from Ref. [190]. Copyright 2017 American Chemical Society.



(b) Identified materials

Formula	Space group	E_{hull} (meV/atom)	Bandgap (eV)
Li_2MnSiO_5	P4/nmm	86	2.24
$MnCdGe_2O_6$	$P2_1/c$	99	2.47
$MnCdGe_2O_6$	C2/c	99	1.76
$MnAg(SeO_3)_2$	$Pna2_1$	36	2.31

(c) $MnAg(SeO_3)_2$

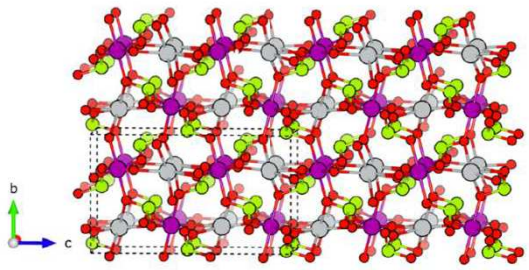


Fig. 9. (a) Workflow for the materials search. (b) Information for the candidate materials identified from the high-throughput screening. (c) Crystal structure of the most stable compound $MnAg(SeO_3)_2$. Reproduced with permission from Ref. [183]. Copyright 2019 American Chemical Society.

Maximally absorbing sunlight is necessary, but does not guarantee excellent solar conversion efficiency, for which we also need to consider the loss mechanisms during energy conversion. For solar cells, one major energy loss is due to the presence of deep-level defects that induce carrier recombination, leading to energy dissipation in the form of phonons, *i.e.*, heat.[191] By compiling various experimental data from the literature, Dahliah *et al.*[192] showed that for the same material composition the experimental solar conversion efficiency can be quite different from the theoretical ones, which, to a large extent, is due to the ignored impacts of defects in the material.

Hence, to design solar absorbers that will more likely to enable high power conversion efficiencies, Dahliah *et al.*[192] further included defect-assisted nonradiative recombination in their high-throughput screening process. They employed the configuration coordinate diagram to study the charge-state transitions of point defects, which allows to analyze and compute the electron and hole capture barriers and coefficients.[161] The capture coefficients can be further employed to estimate the theoretical efficiency with the impacts of defects considered.

Using this improved screening scheme for more than 7000 Cu-based materials in the Materials Project database, Dahliah *et al.* identified handful defect-tolerant solar absorbers. Most notably, $K_3Cu_3P_2$ and Na_2CuP exhibit not only high theoretical efficiencies, but they are also earth abundant as indicated by their low Herfindahl–Hirschman index (HHI),[193] making them very strong candidates for potential high-performance solar absorbers.

High-throughput screening relies critically on the availability of accurate materials properties computed using DFT. DFT calculations are overall expensive. For commonly known materials, one can make use of publicly available databases such as the Materials Project. However, for many new mate-

rials, one would have to compute the required materials properties by themselves. This can be very expensive if one aims for a large number of structures and chemical compositions. To address this challenge, one can combine machine learning with high-throughput screening.

As shown in Fig. 9(a), Davies *et al.*[194] built a statistical model to predict the band gap from chemical compositions based on supervised machine learning. By making use of the machine-learning model, Davies *et al.* could efficiently screen more than one million quaternary oxides, through which they have successfully identified four materials (two of them have the same chemical composition, but are different polymorphs) as shown in Fig. 9(b). These four materials are thermodynamically stable or metastable, and have direct band gaps in the visible range of the solar spectrum. Figure 9(c) shows the crystal structure of one of the candidate materials $MnAg(SeO_2)_2$. One can see that the chemical composition and the corresponding crystal structure are relatively complicated; without efficient machine-learning algorithm to accelerate the high-throughput workflow, it is very challenging to identify promising yet complex materials like the ones found here.

3.2.2. Photocatalytic materials

In addition to solar cells which covert solar energy to electrical energy, photo-induced water splitting is another way to utilize the sunlight by converting solar energy to chemical energy. However, to achieve high conversion efficiency, photoelectrocatalysts for the oxygen evolution reaction (OER) in water splitting are critical.[195] Through experimental efforts of multiple decades, only 16 metal-oxide photoanodes were found, which have band gaps in the desired range of 1.2 eV–2.8 eV. Apparently, this is a challenging and costly task. By combining theory and experiment in a high-throughput fashion, Yan *et al.*[196] developed a very efficient computational approach to the design of novel photocatalysts, which successfully yielded 12 new photocatalysts for the OER in water splitting.

They approached this challenge by first carefully analyzing known photoanode materials such as monoclinic $BiVO_4$[197] and β-$Mn_2V_2O_7$[198] (even though β-$Mn_2V_2O_7$ is not photoactive for the OER, it has a suitable band gap of 1.8 eV). One feature in common between the two materials is the existence of a VO_4 structural motif, in which hybridization between V 3d orbitals and O 2p orbitals leads to a desired band gap and a suitable VBM energy level (E_{VBM}). This inspires a hypothesis that the presence of VO_4 motif is a common feature in ternary vanadates that are potentially suitable for being photoanode materials in water splitting.

Motivated by this hypothesis, Yan *et al.*[196] designed a high-throughput screening scheme. Starting from more than 6.6×10^4 ternary vanadates with VO_4 motif in the Materials Project, they first used DFT calculations to search for the compounds with suitable band gaps, stabilities, and E_{VBM}, which resulted in 34 candidate materials. Further, they carried out experimental characterizations of these materials to verify their applicability to photocatalysts for the OER in water splitting. In the end, 15 materials passed through the screening, which also triggered lots of follow-up experimental studies.[199–205] For instance, the photoelectrocatalytic performance of β-$Cu_2V_2O_7$ was experimentally evaluated by Song *et al.*;[206] a solar-to-hydrogen efficiency of 16% was demonstrated.

They also analyzed the relation between the VBM position and the O 2p band character at the VBM for the newly identified materials. They found that both the band gap and E_{VBM} can be tuned in a broad range through changing the VO_4 band-edge character. This demonstrates flexible and tunable properties of these materials, making them very promising materials for photoelectrocatalysts in water splitting.

This example here again nicely shows how to combine knowledge extracted from known materials and high-throughput screening to achieve successful materials design. In this case, the authors identified the structural motif by their physical intuition and careful analyses of the existing materials. However, this process is non-trivial and very often rather challenging. However, one can actually more systematically search for "descriptors" that could well characterize the photocatalytic (or other applications related) performance.

For instance, employing symbolic regression Weng *et al.*[207] systematically derived a simple descriptor for the OER performance of various perovskite oxides. As shown in Fig. 10(a), they first synthesized a number of oxide perovskite catalysts and characterized their OER activities. Based on the obtained activity data and selected materials features, they performed symbolic regression for the datasets using different descriptors; the descriptors were generated by combining materials parameters in different mathematical forms, as implemented in the genetic programming code gplearn.[208]

Figure 10(b) shows the mean absolute errors (MAE) of the descriptors at different levels of complexities. One can see that descriptor C (μ/t) can best balance between simplicity and accuracy, where μ and t are the octahedral and tolerance factors, respectively. Using this new descriptor, one can nicely capture the chemical trend of the potentials of the oxide perovskite catalysts (V_{RHE}) using a linear correlation with the descriptor, even for materials beyond the ones synthesized by the authors (*i.e.*, the ones used for the symbolic regression) [see Fig. 10(c)]. This demonstrates the predictive capability of the descriptor and the reliability of its linear correlation with V_{RHE}.

With the newly derived descriptor, Weng *et al.* further screened a large compositional space for oxide perovskite catalysts as shown in Fig. 10(a). By doing so, they identified 13 potential candidates for enhanced OER activ-

ity, and managed to successfully synthesize and characterize five of them. Impressively, four out of the five new materials (including $Cs_{0.4}La_{0.6}Mn_{0.25}Co_{0.75}O_3$, $Cs_{0.3}La_{0.7}NiO_3$, $SrNi_{0.75}Co_{0.25}O_3$, and $Sr_{0.25}Ba_{0.75}NiO_3$) exhibited the highest intrinsic activities among oxide perovskite catalysts. This example again nicely shows the powerfulness of machine-learning techniques in accelerating the DFT-guided materials design.

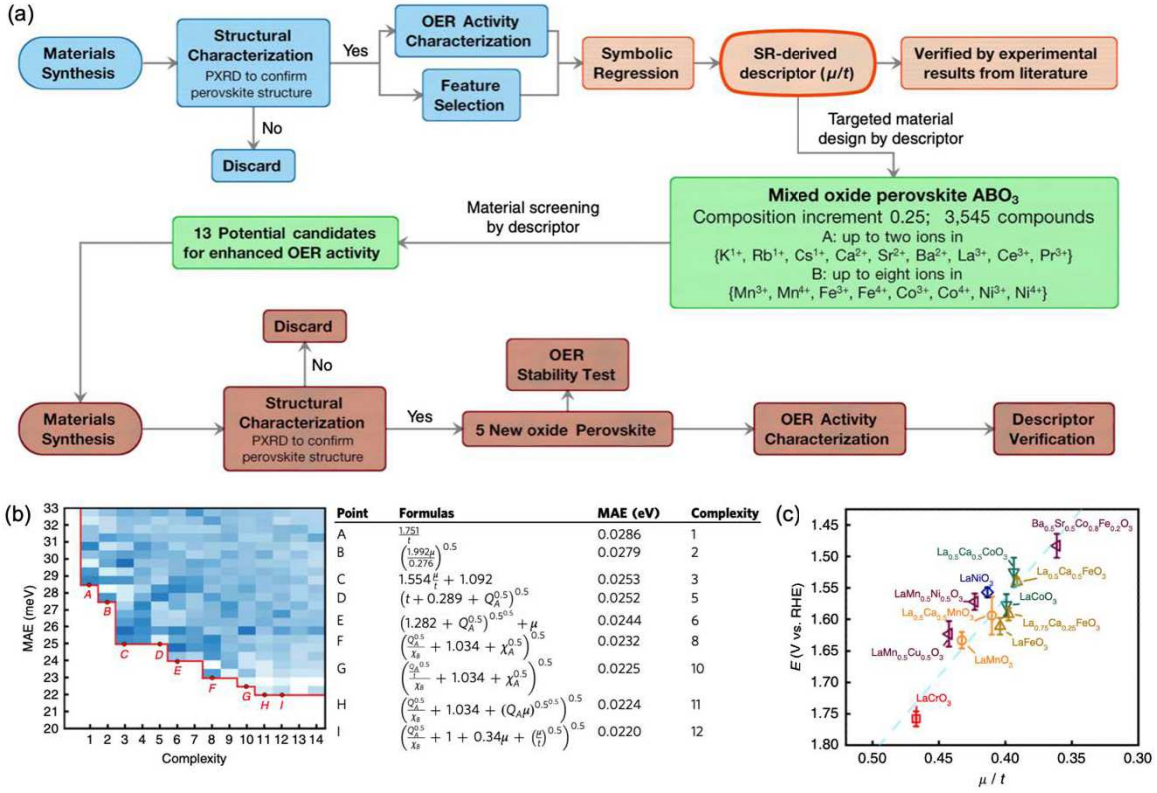


Fig. 10. (a) Workflow for the identification of OER descriptors. (b) The mean absolute errors (MAE) versus complexity for 8640 mathematical formulas, and the well-performing ones at different levels of complexities. (c) Performance of the descriptor μ/t in determining the potential relative to reversible hydrogen electrode (RHE). Reproduced from Ref. [207] in accordance with the Creative Commons Attribution (CC BY) license. Copyright 2020 Springer Nature.

3.2.3. Thermoelectric materials

Thermoelectric materials can make use of temperature gradient and convert it into electricity, and thus they are also promising energy materials for applications such as power generation and refrigeration.

The efficiency of a thermoelectric material is determined by a number of materials properties, including electrical conductivity (σ), thermal conductivity (κ), and Seebeck coefficient (S). The efficiency is also temperature (T)-dependent. Hence, a figure of merit zT is defined to characterize the maximum energy conversion efficiency of a thermoelectric material, namely,[209,210]

$$zT = \frac{\sigma S^2 T}{\kappa}. \tag{2}$$

To design efficient thermoelectric materials, there are multiple target properties required, and there often exists a trade-off between some properties. For instance, high electrical conductivity and low thermal conductivity cannot be easily achieved simultaneously. These complications make the computational design of novel efficient thermoelectric materials difficult.

Recently, Zhang *et al.*[211] systematically implemented a package for inverse design of materials by multi-objective differential evolution (IM^{2}ODE), which is based on the GA algorithm and supports both single-objective and multi-objective optimizations [see Figs. 11(a) and 11(b)]. Each candidate solution to the objective function is considered as a vector in the multidimensional space. As discussed above, three types of operations on the vectors (mutation, crossover, and selection) are involved during the evolution process.

By using the GA algorithm as implemented in the IM^{2}ODE package, Yang *et al.*[212,213] successfully identified two new thermoelectric materials: SiS and SiSe monolayers with the Pmma symmetry [see Fig. 11(c)]. These two materials are comprised of earth-abundant and non-toxic elements, and exhibit attractive thermoelectric performance because of intrinsically low thermal conductivities and high power factors (σS^2). The low thermal conductivity is due to low electron–phonon scattering, and the high power factor is primarily a result of the high density of states in the vicinity of the band edges [see Fig. 11(d)]. Combining these advantages, the final zT values are greater than 1 over a wide range of temperatures, and may even exceed 2.

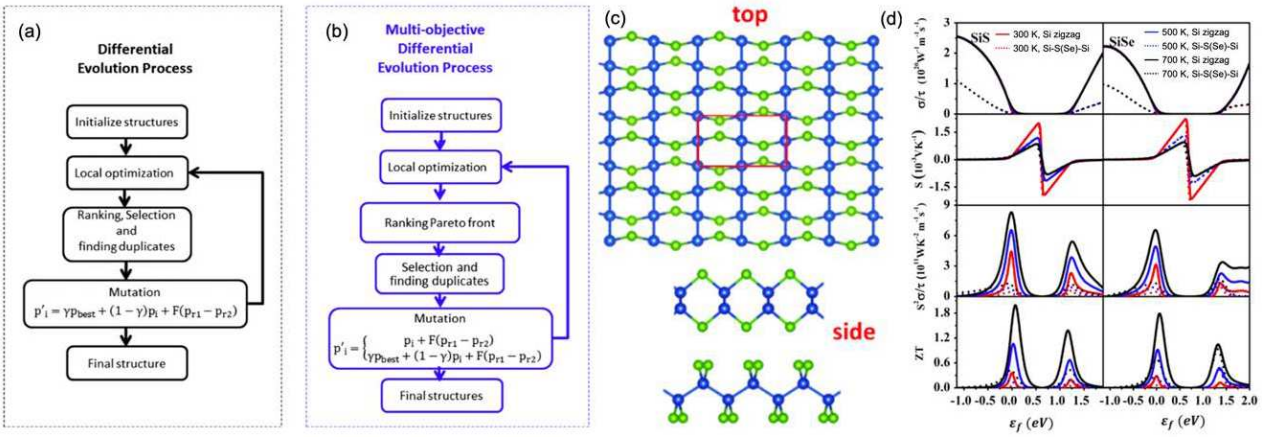


Fig. 11. Two differential evolution processes in the IM^{2}ODE package: (a) single-objective optimization and (b) multi-objective optimization. Reproduced with permission from Ref. [211]. Copyright 2015 Elsevier. (c) Crystal structure of Pmma SiS (or SiSe). (d) Thermoelectric performance of *Pmma* SiS and SiSe at different temperatures. Reproduced with permission from Ref. [213]. Copyright 2017 American Chemical Society.

3.2.4. Battery-cathode materials

After energy conversion have been successfully realized, we still need materials for energy storage, for which batteries are a key technology. Currently, Li-ion batteries have been successfully commercialized and used in various electronic devices and electrical vehicles. However, the abundance of Li on earth is relatively limited. In particular, with the rapid development of electrical vehicles in recent years, the Li demand and consumption have dramatically increased. The search for alternative battery cathode materials is extremely important and useful.

Zhang *et al.*[214] developed an effective scheme to screen cathode materials for Na-based batteries. Among the known Na-based cathode materials, Na-bearing layered materials are of particular interest, since they allow fast and reversible intercalation of Na ions without pronounced lattice distortions. As shown in Fig. 12(a), Zhang *et al.* formulated a simple but effective algorithm to extract Na-containing layered materials from the Materials Project database. By doing this, they obtained more than 150 candidate materials, which cover a variety of space groups (38 in total) [Fig. 12(b)].

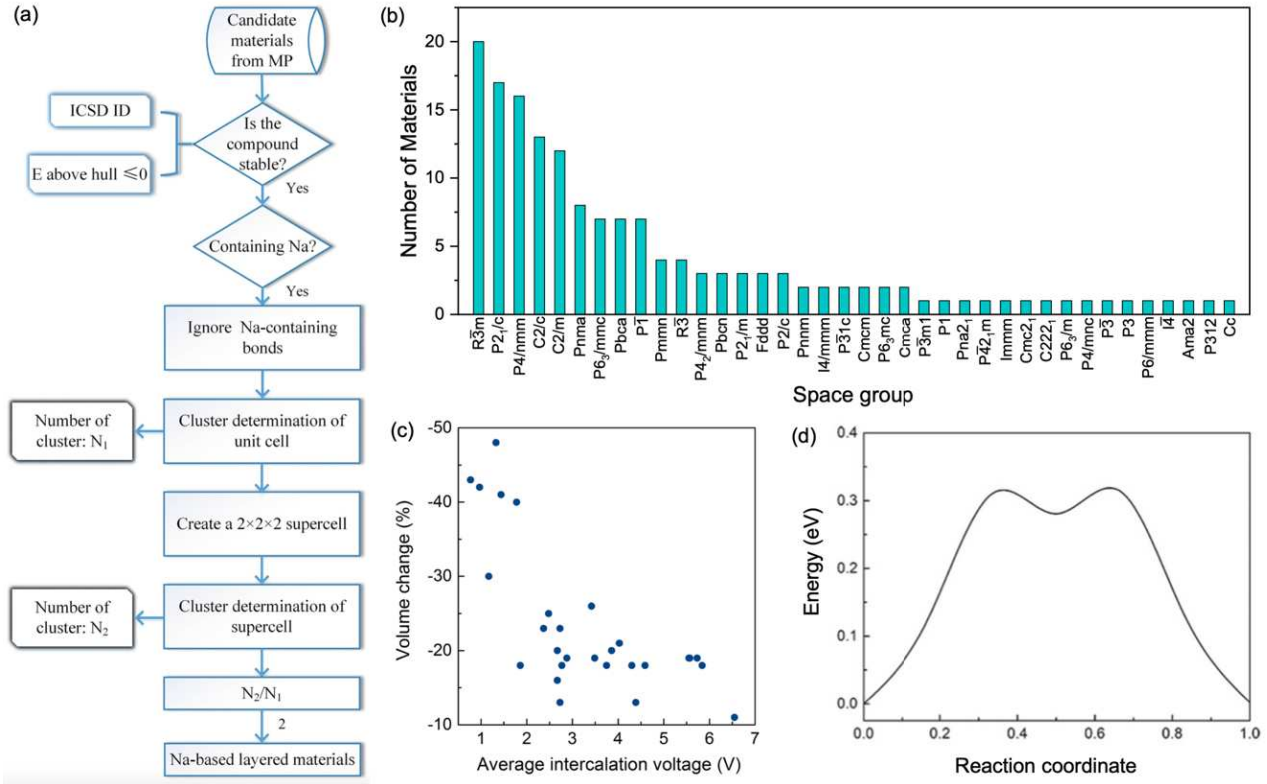


Fig. 12. (a) Workflow for the high-throughput search of Na-battery cathode materials. (b) Distribution of the materials space groups. (c) Correlation between the volume change due to Na de-intercalation and the average intercalation voltage. (d) Na interlayer migration barrier in $NaCoO_2$. Reproduced from Ref. [214] in accordance with the Creative Commons Attribution (CC BY) license. Copyright 2018 Springer Nature.

In addition to basic thermodynamic consideration, there are more factors that need to be taken into account, such as volume change due to Na de-intercalation, average intercalation voltage, and Na ion migration barriers. As shown in Figs. 12(c) and 12(d), Zhang *et al.* further filtered the candidate materials based on the three additional properties. In the end, they identified a number of promising cathode materials for Na–ion batteries, including $Na(CuO)_2$, $NaTiF_4$, $Na_2Zr(CuS_2)_2$, $Na_3Co_2SbO_6$, and $Na_2Cu(CO_3)_2$. These materials have suitable voltages, high capacities, low volume changes during Na intercalation/de-intercalation, as well as small Na diffusion barriers. This finding could then serve as a guide for experimentalists to synthesize these materials of great promise, boosting the development of Na-ion battery technology.

4. Summary and outlook

In summary, we show that the rapid developments in computational methodologies and computer power greatly enhance the predictive power of DFT calculations for various key properties of energy materials, and result in the discovery of many novel materials for potential energy applications. There is no doubt that the robustness and accuracy of DFT prediction will continuously improve in the future, *e.g.*, with the development of better DFT functionals that are accurate for both molecules and solids and better *GW* approaches that have less dependence on the DFT calculated input parameters. We expect a growing number of successful applications of DFT calculations to practical developments of energy materials will emerge. Nevertheless, there remain challenges regarding how to make DFT simulations closer to real materials and conditions, in terms of length- and time-scales, material complexities, and temperature. Here we discuss a few topics of potential interest for future studies.

(i) How to extend the length- and time-scales of DFT simulations towards realistic experimental conditions? Due to the $O(N^3)$ scaling, the system size of conventional DFT calculations is usually limited to $< 10^3$ atoms, and the time-scale is typically on the order of picoseconds. However, energy conversion involves many physical processes with different length- and time-scales, which can go beyond the capacity of conventional DFT calculations. To extend the length- and time-scales, it is desirable to develop linear-scaling DFT methods.[215] A few examples of such methods are the charge patching method[216] and the linear scaling 3D fragment method,[126] which can deal with structures containing over 10^4 atoms at moderate computational costs.[217,218] Another promising approach is combining machine learning techniques with DFT inputs. MD simulation of 100 million atoms with machine learning force fields has been demonstrated, at a speed of one nanosecond per day.[219] In low-dimensional materials with inhomogeneous deformation such as torsion and bending, translational symmetry is missing, thus an extremely large supercell is required for direct DFT calculations. In this case, the generalized Bloch method could be a good choice, which considers the helical symmetry and rotational symmetry to deal with torsion and bending, allowing the simulations of deformed structures with a relatively small repeating cell.[220,221]

(ii) How to take into account different types of disorder or partial order/disorder in complex materials? The flexibility in materials design and emerging new properties of structurally and compositionally complex materials facilitate their applications in energy conversion. The increased complexity often results in structural and chemical disorder. It is important to develop algorithms to efficiently generate accurate models to describe the disorder, *e.g.*, the SQS for multicomponent systems, and amorphous structures. To investigate partial short-range or long-range disorder, the concept of SQS can be extended to special quasi-disordered structure (SQDS), whose correlation functions best match those of the target structure with partial order/disorder. There are also conceptual challenges in defect simulations of disordered or partially disordered materials, such as the definition of defects in this case,[222] and how to conduct the statistical averages over site-dependent defect properties.

(iii) How to properly include temperature effects during property prediction? Standard DFT ground-state calculations only yield properties at 0 K. Temperature is a key factor affecting the performance of energy materials. Some materials that are unstable at 0 K can be stabilized by vibrational or configurational entropies.[223] Entropy also plays important roles in chemical reactions. In addition, lattice anharmonicity could be greatly enhanced by temperature.[224,225] Developments of methods considering these temperature effects are highly desirable.[226,227]

Acknowledgment

Project supported by the National Natural Science Foundation of China (Grant Nos. 12088101, 11991060, 12074029, 52172136, and U1930402).

References

[1] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
[2] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
[3] Jones R O 2015 *Rev. Mod. Phys.* **87** 897
[4] Louie S G 2021 *Nat. Mater.* **20** 8
[5] Gunnarsson O and Lundqvist B I 1976 *Phys. Rev. B* **13** 4274
[6] Zunger A and Cohen M L 1978 *Phys. Rev. B* **18** 5449
[7] Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471
[8] Connolly J W D and Williams A R 1983 *Phys. Rev. B* **27** 5169
[9] Wimmer E, Krakauer H, Weinert M and Freeman A J 1981 *Phys. Rev. B* **24** 864

[10] Wei S H and Krakauer H 1985 *Phys. Rev. Lett.* **55** 1200
[11] Payne M C, Teter M P, Allan D C, Arias T A and Joannopoulos J D 1992 *Rev. Mod. Phys.* **64** 1045
[12] Zunger A 2018 *Nat. Rev. Chem.* **2** 0121
[13] Sanchez-Lengeling B and Aspuru-Guzik A 2018 *Science* **361** 360
[14] Perdew J P and Schmidt K 2001 *AIP Conf. Proc.* **577** 1
[15] Ceperley D M and Alder B J 1980 *Phys. Rev. Lett.* **45** 566
[16] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
[17] Zhao Y and Truhlar D G 2006 *J. Chem. Phys.* **125** 194101
[18] Sun J, Ruzsinszky A and Perdew J P 2015 *Phys. Rev. Lett.* **115** 036402
[19] Stephens P J, Devlin F J, Chabalowski C F and Frisch M J 1994 *J. Phys. Chem.* **98** 11623
[20] Heyd J, Scuseria G E and Ernzerhof M 2003 *J. Chem. Phys.* **118** 8207
[21] Zhang G X, Reilly A M, Tkatchenko A and Scheffler M 2018 *New J. Phys.* **20** 063020
[22] Tran F, Stelzl J and Blaha P 2016 *J. Chem. Phys.* **144** 204120
[23] Peverati R and Truhlar D G 2014 *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **372** 20120476
[24] Medvedev M G, Bushmarinov I S, Sun J, Perdew J P and Lyssenko K A 2017 *Science* **355** 49
[25] Bagno P, Jepsen O and Gunnarsson O 1989 *Phys. Rev. B* **40** 1997
[26] Hu S, Xia B, Yan Y and Xiao Z 2020 *Phys. Rev. Mater.* **4** 115201
[27] Rivero P, Moreira I de P R, Scuseria G E and Illas F 2009 *Phys. Rev. B* **79** 245129
[28] Tran F, Blaha P, Schwarz K and Novák P 2006 *Phys. Rev. B* **74** 155108
[29] Elmaslmane A R, Watkins M B and McKenna K P 2018 *J. Chem. Theory Comput.* **14** 3740
[30] Grimme S, Antony J, Ehrlich S and Krieg H 2010 *J. Chem. Phys.* **132** 154104
[31] Tkatchenko A and Scheffler M 2009 *Phys. Rev. Lett.* **102** 073005
[32] Dion M, Rydberg H, Schröder E, Langreth D C and Lundqvist B I 2004 *Phys. Rev. Lett.* **92** 246401
[33] Berland K, Cooper V R, Lee K, Schröder E, Thonhauser T, Hyldgaard P and Lundqvist B I 2015 *Rep. Prog. Phys.* **78** 066501
[34] Sabatini R, Gorni T and de Gironcoli S 2013 *Phys. Rev. B* **87** 041108
[35] Foulkes W M C, Mitas L, Needs R J and Rajagopal G 2001 *Rev. Mod. Phys.* **73** 33
[36] Schimka L, Harl J, Stroppa A, Grüneis A, Marsman M, Mittendorfer F and Kresse G 2010 *Nat. Mater.* **9** 741
[37] Lebègue S, Harl J, Gould T, Ángyán J G, Kresse G and Dobson J F 2010 *Phys. Rev. Lett.* **105** 196401
[38] Peng H, Yang Z H, Perdew J P and Sun J 2016 *Phys. Rev. X* **6** 041005
[39] Tran F, Kalantari L, Traoré B, Rocquefelte X and Blaha P 2019 *Phys. Rev. Mater.* **3** 063602
[40] Hellmann H 1937 *Einführung in die Quantenchemie* (Leipzig: Franz Deuticke)
[41] Feynman R P 1939 *Phys. Rev.* **56** 340
[42] Alfè D 2009 *Comput. Phys. Commun.* **180** 2622
[43] Baroni S, de Gironcoli S, Dal Corso A and Giannozzi P 2001 *Rev. Mod. Phys.* **73** 515
[44] Henkelman G, Uberuaga B P and Jónsson H 2000 *J. Chem. Phys.* **113** 9901
[45] Henkelman G and Jónsson H 1999 *J. Chem. Phys.* **111** 7010
[46] Kulkarni A, Siahrostami S, Patel A and Nørskov J K 2018 *Chem. Rev.* **118** 2302
[47] Seh Z W, Kibsgaard J, Dickens C F, Chorkendorff I, Nørskov J K and Jaramillo T F 2017 *Science* **355** eaad4998
[48] Perdew J P and Levy M 1983 *Phys. Rev. Lett.* **51** 1884
[49] Mori-Sánchez P, Cohen A J and Yang W 2008 *Phys. Rev. Lett.* **100** 146401
[50] Hedin L 1965 *Phys. Rev.* **139** A796
[51] Onida G, Reining L and Rubio A 2002 *Rev. Mod. Phys.* **74** 601
[52] Fuchs F, Furthmüller J, Bechstedt F, Shishkin M and Kresse G 2007 *Phys. Rev. B* **76** 115109
[53] Cao H, Yu Z, Lu P and Wang L W 2017 *Phys. Rev. B* **95** 035139
[54] Chen W and Pasquarello A 2015 *Phys. Rev. B* **92** 041115
[55] Grumet M, Liu P, Kaltak M, Klimeš J and Kresse G 2018 *Phys. Rev. B* **98** 155143
[56] Lucero M J, Henderson T M and Scuseria G E 2012 *J. Phys.: Condens. Matter* **24** 145504
[57] Das P, Mohammadi M and Bazhirov T 2018 arXiv: 1807.05623
[58] Shishkin M and Kresse G 2007 *Phys. Rev. B* **75** 235102
[59] Chen W and Pasquarello A 2012 *Phys. Rev. B* **86** 035134
[60] Rinke P, Scheffler M, Qteish A, Winkelnkemper M, Bimberg D and Neugebauer J 2006 *Appl. Phys. Lett.* **89** 161919
[61] Kim Y S, Hummer K and Kresse G 2009 *Phys. Rev. B* **80** 035203
[62] Anisimov V I, Aryasetiawan F and Lichtenstein A I 1997 *J. Phys.: Condens. Matter* **9** 767
[63] Tran F and Blaha P 2009 *Phys. Rev. Lett.* **102** 226401
[64] Chan M K Y and Ceder G 2010 *Phys. Rev. Lett.* **105** 196403
[65] Wang L W 2001 *Appl. Phys. Lett.* **78** 1565
[66] Ferreira L G, Marques M and Teles L K 2008 *Phys. Rev. B* **78** 125116
[67] Ma J and Wang L W 2016 *Sci. Rep.* **6** 24924
[68] Kang J, Tongay S, Zhou J, Li J and Wu J 2013 *Appl. Phys. Lett.* **102** 012111
[69] Wei S H and Zunger A 1998 *Appl. Phys. Lett.* **72** 2011
[70] Li Y H, Walsh A, Chen S, Yin W J, Yang J H, Li J, Da Silva J L F, Gong X G and Wei S H 2009 *Appl. Phys. Lett.* **94** 212109
[71] Lang L, Zhang Y Y, Xu P, Chen S, Xiang H J and Gong X G 2015 *Phys. Rev. B* **92** 075102
[72] Hu Y, Qiu C, Shen T, Yang K and Deng H 2021 *J. Semicond.* **42** 112102
[73] Gu H J, Zhang Y Y, Chen S Y, Xiang H J and Gong X G 2018 *Phys. Rev. B* **97** 235308
[74] Ambrosch-Draxl C and Sofo J O 2006 *Comput. Phys. Commun.* **175** 1
[75] Del Sole R and Girlanda R 1993 *Phys. Rev. B* **48** 11789
[76] Gajdoš M, Hummer K, Kresse G, Furthmüller J and Bechstedt F 2006 *Phys. Rev. B* **73** 045112
[77] Ramasubramaniam A 2012 *Phys. Rev. B* **86** 115409
[78] Salpeter E E and Bethe H A 1951 *Phys. Rev.* **84** 1232
[79] Rohlfing M and Louie S G 1998 *Phys. Rev. Lett.* **81** 2312
[80] Ping Y, Rocca D and Galli G 2013 *Chem. Soc. Rev.* **42** 2437
[81] Cunningham B, Grüning M, Azarhoosh P, Pashov D and van Schilfgaarde M 2018 *Phys. Rev. Mater.* **2** 034603
[82] Casida M E and Huix-Rotllant M 2012 *Annu. Rev. Phys. Chem.* **63** 287
[83] Maitra N T 2016 *J. Chem. Phys.* **144** 220901
[84] Runge E and Gross E K U 1984 *Phys. Rev. Lett.* **52** 997
[85] Sharma S, Dewhurst J K, Sanna A and Gross E K U 2011 *Phys. Rev. Lett.* **107** 186401
[86] Rigamonti S, Botti S, Veniard V, Draxl C, Reining L and Sottile F 2015 *Phys. Rev. Lett.* **114** 146402
[87] Botti S, Sottile F, Vast N, Olevano V, Reining L, Weissker H C, Rubio A, Onida G, Del Sole R and Godby R W 2004 *Phys. Rev. B* **69** 155112
[88] Refaely-Abramson S, Jain M, Sharifzadeh S, Neaton J B and Kronik L 2015 *Phys. Rev. B* **92** 081204
[89] Yang Z and Ullrich C A 2013 *Phys. Rev. B* **87** 195204
[90] Liu W H, Wang Z, Chen Z H, Luo J W, Li S S and Wang L W 2022 *WIREs Comput. Mol. Sci.* **12** e1577
[91] Lopata K and Govind N 2011 *J. Chem. Theory Comput.* **7** 1344
[92] Kolesov G, Grånäs O, Hoyt R, Vinichenko D and Kaxiras E 2016 *J. Chem. Theory Comput.* **12** 466
[93] Ma J, Wang Z and Wang L W 2015 *Nat. Commun.* **6** 10107
[94] Kang J and Wang L W 2019 *Phys. Rev. B* **99** 224303
[95] Zhang J, Hong H, Zhang J, Fu H, You P, Lischner J, Liu K, Kaxiras E and Meng S 2018 *Nano Lett.* **18** 6057
[96] Bang J, Sun Y Y, Liu X Q, Gao F and Zhang S B 2016 *Phys. Rev. Lett.* **117** 126402
[97] Miyamoto Y, Zhang H, Cheng X and Rubio A 2019 *Phys. Rev. B* **99** 165424
[98] Chen Z, Luo J W and Wang L W 2019 *Proc. Natl. Acad. Sci.* **116** 19258
[99] Wang L W 2020 *J. Phys. Chem. A* **124** 9075
[100] Wang Z, Li S S and Wang L W 2015 *Phys. Rev. Lett.* **114** 063004
[101] Meyer H and Miller W H 1980 *J. Chem. Phys.* **72** 2272
[102] Tully J C 1990 *J. Chem. Phys.* **93** 1061
[103] Akimov A V and Prezhdo O V 2013 *J. Chem. Theory Comput.* **9** 4959
[104] Gao Q and Kang J 2022 *Phys. Chem. Chem. Phys.* **24** 9891
[105] Madjet M E, Berdiyorov G R, El-Mellouhi F, Alharbi F H, Akimov A V and Kais S 2017 *J. Phys. Chem. Lett.* **8** 4439
[106] Akimov A V, Asahi R, Jinnouchi R and Prezhdo O V 2015 *J. Am. Chem. Soc.* **137** 11517
[107] Zheng F and Wang L W 2019 *J. Phys. Chem. Lett.* **10** 6174
[108] Jiang X, Zheng Q, Lan Z, Saidi W A, Ren X and Zhao J 2021 *Sci. Adv.* **7** eabf3759

[109] Yang K, Xiao J, Ren Z, Wei Z, Luo J W, Wei S H and Deng H X 2021 *J. Phys. Chem. Lett.* **12** 7832
[110] Giustino F 2017 *Rev. Mod. Phys.* **89** 015003
[111] Madsen G K H, Carrete J and Verstraete M J 2018 *Comput. Phys. Commun.* **231** 140
[112] Beleznay F B, Bogár F and Ladik J 2003 *J. Chem. Phys.* **119** 5690
[113] Xi J, Long M, Tang L, Wang D and Shuai Z 2012 *Nanoscale* **4** 4348
[114] Zhou J J and Bernardi M 2016 *Phys. Rev. B* **94** 201201
[115] Li W 2015 *Phys. Rev. B* **92** 075405
[116] Noffsinger J, Giustino F, Malone B D, Park C H, Louie S G and Cohen M L 2010 *Comput. Phys. Commun.* **181** 2140
[117] Restrepo O D, Varga K and Pantelides S T 2009 *Appl. Phys. Lett.* **94** 212103
[118] Gunst T, Markussen T, Stokbro K and Brandbyge M 2016 *Phys. Rev. B* **93** 035414
[119] Park C H, Bonini N, Sohier T, Samsonidze G, Kozinsky B, Calandra M, Mauri F and Marzari N 2014 *Nano Lett.* **14** 1113
[120] Fiorentini M and Bonini N 2016 *Phys. Rev. B* **94** 085204
[121] Taylor J, Guo H and Wang J 2001 *Phys. Rev. B* **63** 245407
[122] Brandbyge M, Mozos J L, Ordejón P, Taylor J and Stokbro K 2002 *Phys. Rev. B* **65** 165401
[123] Smidstrup S, Markussen T, Vancraeyveld P, Wellendorff J, Schneider J, Gunst T, Verstichel B, Stradi D, Khomyakov P A, Vej-Hansen U G, Lee M E, Chill S T, Rasmussen F, Penazzi G, Corsetti F, Ojanperä A, Jensen K, Palsgaard M L N, Martinez U, Blom A, Brandbyge M and Stokbro K 2019 *J. Phys.: Condens. Matter* **32** 015901
[124] Wang L W 2005 *Phys. Rev. B* **72** 045417
[125] Garcia-Lekue A and Wang L W 2006 *Phys. Rev. B* **74** 245404
[126] Wang L W, Zhao Z and Meza J 2008 *Phys. Rev. B* **77** 165113
[127] Ye M, Jiang X, Li S S and Wang L W 2021 *Comput. Phys. Commun.* **260** 107737
[128] Marcus R A 1993 *Rev. Mod. Phys.* **65** 599
[129] Wang L, Nan G, Yang X, Peng Q, Li Q and Shuai Z 2010 *Chem. Soc. Rev.* **39** 423
[130] Chu I H, Radulaski M, Vukmirovic N, Cheng H P and Wang L W 2011 *J. Phys. Chem. C* **115** 21409
[131] Vukmirović N and Wang L W 2009 *Nano Lett.* **9** 3996
[132] Kang J 2020 *J. Phys. Chem. Lett.* **11** 3615
[133] Broido D A, Malorny M, Birner G, Mingo N and Stewart D A 2007 *Appl. Phys. Lett.* **91** 231922
[134] Garg J, Bonini N, Kozinsky B and Marzari N 2011 *Phys. Rev. Lett.* **106** 045901
[135] Li W, Carrete J, Katcho A N and Mingo N 2014 *Comput. Phys. Commun.* **185** 1747
[136] Schelling P K, Phillpot S R and Keblinski P 2002 *Phys. Rev. B* **65** 144306
[137] Müller-Plathe F 1997 *J. Chem. Phys.* **106** 6082
[138] Lee Y and Hwang G S 2012 *Phys. Rev. B* **86** 075202
[139] Stackhouse S, Stixrude L and Karki B B 2010 *Phys. Rev. Lett.* **104** 208501
[140] McGaughey A J H and Larkin J M 2014 *Annu. Rev. Heat Transfer* **17** 1
[141] Fan Z, Pereira L F C, Wang H Q, Zheng J C, Donadio D and Harju A 2015 *Phys. Rev. B* **92** 094301
[142] Kang J and Wang L W 2017 *Phys. Rev. B* **96** 020302
[143] Marcolongo A, Umari P and Baroni S 2016 *Nat. Phys.* **12** 80
[144] Carbogno C, Ramprasad R and Scheffler M 2017 *Phys. Rev. Lett.* **118** 175901
[145] Wei S H 2004 *Comput. Mater. Sci.* **30** 337
[146] Freysoldt C, Grabowski B, Hickel T, Neugebauer J, Kresse G, Janotti A and Van de Walle C G 2014 *Rev. Mod. Phys.* **86** 253
[147] Sin C K, Zhang J, Tse K and Zhu J 2020 *J. Semicond.* **41** 061101
[148] Xiao Z, Du K Z, Meng W, Wang J, Mitzi D B and Yan Y J 2017 *J. Am. Chem. Soc.* **139** 6054
[149] Xiao Z, Du K Z, Meng W, Mitzi D B and Yan Y 2017 *Angew. Chem.* **129** 12275
[150] Kim S, Hood S N, Park J S, Whalley L D and Walsh A 2020 *J. Phys. Energy* **2** 036001
[151] Huang M, Zheng Z, Dai Z, Guo X, Wang S, Jiang L, Wei J and Chen S 2022 *J. Semicond.* **43** 042101
[152] Goyal A, Gorai P, Peng H, Lany S and Stevanović V 2017 *Comput. Mater. Sci.* **130** 1
[153] Broberg D, Medasani B, Zimmermann N E R, Yu G, Canning A, Haranczyk M, Asta M and Hautier G 2018 *Comput. Phys. Commun.* **226** 165
[154] Henry C H and Lang D V 1977 *Phys. Rev. B* **15** 989
[155] Huang K 1981 *Sci. Sin.* **24** 27
[156] Huang K 2019 *J. Semicond.* **40** 090102
[157] Wang L 2019 *J. Semicond.* **40** 091101
[158] Zhang X, Shen J X, Turiansky M E and Van de Walle C G 2021 *Nat. Mater.* **20** 971
[159] Zhang X and Wei S H 2022 *Phys. Rev. Lett.* **128** 136401
[160] Shi L and Wang L W 2012 *Phys. Rev. Lett.* **109** 245501
[161] Alkauskas A, Yan Q and Van de Walle C G 2014 *Phys. Rev. B* **90** 075202
[162] Wang S, Huang M, Wu Y N, Chu W, Zhao J, Walsh A, Gong X G, Wei S H and Chen S 2022 *Nat. Comput. Sci.* **2** 486
[163] Meggiolaro D and De Angelis F 2018 *ACS Energy Lett.* **3** 2206
[164] Kang J 2020 *Phys. Rev. Mater.* **4** 085405
[165] Du X, Liu Y, Pan W, Pang J, Zhu J, Zhao S, Chen C, Yu Y, Xiao Z, Niu G and Tang J 2022 *Adv. Mater.* **34** 2110252
[166] Hu S, Xia B, Lin Y P, Katase T, Fujioka J, Kamiya T, Hosono H, Du K Z and Xiao Z 2020 *Adv. Funct. Mater.* **30** 1909906
[167] Wang D, Han D, Li X B, Xie S Y, Chen N K, Tian W Q, West D, Sun H B and Zhang S B 2015 *Phys. Rev. Lett.* **114** 196801
[168] Makov G and Payne M C 1995 *Phys. Rev. B* **51** 4014
[169] Lany S and Zunger A 2008 *Phys. Rev. B* **78** 235104
[170] Freysoldt C, Neugebauer J and Van de Walle C G 2009 *Phys. Rev. Lett.* **102** 016402
[171] Komsa H P, Rantala T T and Pasquarello A 2012 *Phys. Rev. B* **86** 045112
[172] Komsa H P, Berseneva N, Krasheninnikov A V and Nieminen R M 2014 *Phys. Rev. X* **4** 031044
[173] Suo Z J, Luo J W, Li S S and Wang L W 2020 *Phys. Rev. B* **102** 174110
[174] Chagas da Silva M, Lorke M, Aradi B, Farzalipour Tabriz M, Frauenheim T, Rubio A, Rocca D and Deák P 2021 *Phys. Rev. Lett.* **126** 076401
[175] Xiao J, Yang K, Guo D, Shen T, Deng H X, Li S S, Luo J W and Wei S H 2020 *Phys. Rev. B* **101** 165306
[176] Nordheim L 1931 *Ann. Phys.* **401** 607
[177] Soven P 1967 *Phys. Rev.* **156** 809
[178] Bellaiche L, Wei S H and Zunger A 1997 *Appl. Phys. Lett.* **70** 3558
[179] Wei S H, Ferreira L G, Bernard J E and Zunger A 1990 *Phys. Rev. B* **42** 9622
[180] Zunger A, Wei S H, Ferreira L G and Bernard J E 1990 *Phys. Rev. Lett.* **65** 353
[181] Gao Q, Sahin H, Kang J and Wei S H 2021 *Phys. Rev. B* **104** 064204
[182] Yang J and Wei S H 2019 *Chin. Phys. B* **28** 086106
[183] Yang J, Zhang P and Wei S H 2018 *J. Phys. Chem. Lett.* **9** 31
[184] Jain A, Ong S P, Hautier G, Chen W, Richards W D, Dacek S, Cholia S, Gunter D, Skinner D, Ceder G and Persson K A 2013 *APL Mater.* **1** 011002
[185] Chakraborti N 2004 *Int. Mater. Rev.* **49** 246
[186] Ji G, Han C, Hu S, Fu P, Chen X, Guo J, Tang J and Xiao Z 2021 *J. Am. Chem. Soc.* **143** 10275
[187] Han C, Ji G, Hu S, Mi R, Fu P, Du K Z and Xiao Z 2021 *Chem. Mater.* **33** 7106
[188] Yu L, Kokenyesi R S, Keszler D A and Zunger A 2013 *Adv. Energy Mater.* **3** 43
[189] Yu L and Zunger A 2012 *Phys. Rev. Lett.* **108** 068701
[190] Zhao X G, Yang J H, Fu Y, Yang D, Xu Q, Yu L, Wei S H and Zhang L 2017 *J. Am. Chem. Soc.* **139** 2630
[191] Zhang X, Shen J X and Van de Walle C G 2020 *Adv. Energy Mater.* **10** 1902830
[192] Dahliah D, Brunin G, George J, Ha V A, Rignanese G M and Hautier G 2021 *Energy Environ. Sci.* **14** 5057
[193] Gaultois M W, Sparks T D, Borg C K H, Seshadri R, Bonificio W D and Clarke D R 2013 *Chem. Mater.* **25** 2911
[194] Davies D W, Butler K T and Walsh A 2019 *Chem. Mater.* **31** 7221
[195] Woodhouse M and Parkinson B A 2008 *Chem. Mater.* **20** 2495
[196] Yan Q, Yu J, Suram S K, Zhou L, Shinde A, Newhouse P F, Chen W, Li G, Persson K A, Gregoire J M and Neaton J B 2017 *Proc. Natl. Acad. Sci. USA* **114** 3040

[197] Walsh A, Yan Y, Huda M N, Al-Jassim M M and Wei S H 2009 *Chem. Mater.* **21** 547

[198] Yan Q, Li G, Newhouse P F, Yu J, Persson K A, Gregoire J M and Neaton J B 2015 *Adv. Energy Mater.* **5** 1401840

[199] Gadiyar C, Strach M, Schouwink P, Loiudice A and Buonsanti R 2018 *Chem. Sci.* **9** 5658

[200] Hossain M K, Sarker H P, Sotelo P, Dang U, Rodríguez-Gutiérrez I, Blawat J, Vali A, Xie W, Oskam G, Huda M N, Macaluso R T and Rajeshwar K 2020 *Chem. Mater.* **32** 6247

[201] Kumari S, Helt L, Junqueira J R C, Kostka A, Zhang S, Sarker S, Mehta A, Scheu C, Schuhmann W and Ludwig A 2020 *Int. J. Hydrog. Energy* **45** 12037

[202] Botella P, López-Moreno S, Errandonea D, Manjón F J, Sans J A, Vie D and Vomiero A 2020 *J. Phys.: Condens. Matter* **32** 385403

[203] Scarongella M, Gadiyar C, Strach M, Rimoldi L, Loiudice A and Buonsanti R 2018 *J. Mater. Chem. C* **6** 12062

[204] Kumari S, Junqueira J R C, Sarker S, Mehta A, Schuhmann W and Ludwig A 2020 *J. Chem. Phys.* **153** 014707

[205] Ninova S, Strach M, Buonsanti R and Aschauer U 2020 *J. Chem. Phys.* **153** 084704

[206] Song A, Chemseddine A, Ahmet I Y, Bogdanoff P, Friedrich D, Abdi F F, Berglund S P and van de Krol R 2020 *Chem. Mater.* **32** 2408

[207] Weng B, Song Z, Zhu R, Yan Q, Sun Q, Grice C G, Yan Y and Yin W J 2020 *Nat. Commun.* **11** 3513

[208] Stephens T gplearn

[209] Li C, He W, Wang D and Zhao L D 2021 *Chin. Phys. B* **30** 067101

[210] Liu Q, Li G, Zhu H and Zhao H 2022 *Chin. Phys. B* **31** 047204

[211] Zhang Y Y, Gao W, Chen S, Xiang H and Gong X G 2015 *Comput. Mater. Sci.* **98** 51

[212] Yang J H, Zhang Y, Yin W J, Gong X G, Yakobson B I and Wei S H 2016 *Nano Lett.* **16** 1110

[213] Yang J H, Yuan Q, Deng H, Wei S H and Yakobson B I 2017 *J. Phys. Chem. C* **121** 123

[214] Zhang X, Zhang Z, Yao S, Chen A, Zhao X and Zhou Z 2018 *Npj Comput. Mater.* **4** 1

[215] Wang L W 2014 *Natl. Sci. Rev.* **1** 604

[216] Wang L W 2002 *Phys. Rev. Lett.* **88** 256402

[217] Ma J and Wang L W 2015 *Nano Lett.* **15** 248

[218] Hou X, Kang J, Qin H, Chen X, Ma J, Zhou J, Chen L, Wang L, Wang L W and Peng X 2019 *Nat. Commun.* **10** 1750

[219] Lu D, Wang H, Chen M, Lin L, Car R, E W, Jia W and Zhang L 2021 *Comput. Phys. Commun.* **259** 107624

[220] Zhang D and Wei S 2021 *Chin. Sci. Bull.* **66** 674

[221] Zhang D B and Wei S H 2017 *Npj Comput. Mater.* **3** 1

[222] Medvedeva J E, Zhuravlev I A, Burris C, Buchholz D B, Grayson M and Chang R P H 2020 *J. Appl. Phys.* **127** 175701

[223] Qian X and Yang R 2018 *Phys. Rev. B* **98** 224108

[224] Glensk A, Grabowski B, Hickel T and Neugebauer J 2014 *Phys. Rev. X* **4** 011018

[225] Glensk A, Grabowski B, Hickel T and Neugebauer J 2015 *Phys. Rev. Lett.* **114** 195901

[226] Hickel T, Grabowski B, Körmann F and Neugebauer J 2011 *J. Phys.: Condens. Matter* **24** 053202

[227] Qiao S, Wu Y N, Yan X, Monserrat B, Wei S H and Huang B 2022 *Phys. Rev. B* **105** 115201