

STRUCTURAL SCIENCE CRYSTAL ENGINEERING MATERIALS

ISSN 2052-5206

Received 13 July 2018 Accepted 6 November 2018

Edited by A. J. Blake, University of Nottingham, England

Keywords: bond-valence site energy; diffusion in solids; high-throughput materials screening; solid electrolytes; energy-storage systems.



© 2019 International Union of Crystallography

research paper

SoftBV – a software tool for screening the materials genome of inorganic fast ion conductors

Haomin Chen, Lee Loong Wong and Stefan Adams*

Department of Materials Science and Engineering, National University of Singapore, 9 Engineering Drive 1, Singapore 117575, Singapore. *Correspondence e-mail: mseasn@nus.edu.sg

The identification of materials for advanced energy-storage systems is still mostly based on experimental trial and error. Increasingly, computational tools are sought to accelerate materials discovery by computational predictions. Here are introduced a set of computationally inexpensive software tools that exploit the bond-valence-based empirical force field previously developed by the authors to enable high-throughput computational screening of experimental or simulated crystal-structure models of battery materials predicting a variety of properties of technological relevance, including a structure plausibility check, surface energies, an inventory of equilibrium and interstitial sites, the topology of ion-migration paths in between those sites, the respective migration barriers and the site-specific attempt frequencies. All of these can be predicted from CIF files of structure models at a minute fraction of the computational cost of density functional theory (DFT) simulations, and with the added advantage that all the relevant pathway segments are analysed instead of arbitrarily predetermined paths. The capabilities and limitations of the approach are evaluated for a wide range of ion-conducting solids. An integrated simple kinetic Monte Carlo simulation provides rough (but less reliable) predictions of the absolute conductivity at a given temperature. The automated adaptation of the force field to the composition and charge distribution in the simulated material allows for a high transferability of the force field within a wide range of Lewis acid-Lewis base-type ionic inorganic compounds as necessary for high-throughput screening. While the transferability and precision will not reach the same levels as in DFT simulations, the fact that the computational cost is several orders of magnitude lower allows the application of the approach not only to pre-screen databases of simple structure prototypes but also to structure models of complex disordered or amorphous phases, and provides a path to expand the analysis to charge transfer across interfaces that would be difficult to cover by ab initio methods.

1. Introduction

The current trial-and-error based evolution of battery technologies on a decade timescale is obviously progressing too slowly to exploit opportunities in the 'materials genome' in trying to keep up with the rapidly increasing demand for highperformance energy-storage systems. As an example, the emerging development of low-cost Na-ion batteries is unnecessarily slowed down by many research groups aiming to copy what worked well for Li-ion batteries, thus misunderstanding the fundamental differences between the structural chemistries of the two elements. Artificial intelligence systems should be able to come up with a more rational basis for predicting promising candidate systems.

In this context, ab initio methods, and especially density functional theory (DFT) methods, are widely explored as they generally offer a higher accuracy and greater transferability than empirical potentials. DFT methods are intrinsically less sensitive to the types of bonding (except for van der Waals interactions), while the applicability of an empirical force field typically remains restricted to a limited set of compounds and properties. A few large-scale projects have started to build databases by automating DFT calculations on a set of experimental or theoretical structures (Jain et al., 2013; Calderon et al., 2015; Saal et al., 2013) that can provide useful insight into the static properties of a wide range of compounds. However, though computer hardware advances continuously, high-fidelity ab initio calculation methods are still too slow or too limited in size to cover the genome of complex battery materials yielding dependable predictions of mass transportrelated properties. Force-field methods are therefore still actively used whenever some degree of coarse graining is desired, such as in the cases of screening a large number of compounds, creating prohibitively large structures (e.g. of amorphous or otherwise disordered phases), quickly estimating the properties of a new structure, efficiently exploring configurational space and as an initial assessment to prescreen candidates for more in-depth DFT analyses, concentrating the computational effort of DFT studies on promising candidates. In order to be suitable for screening a wide range of compounds, the empirical force field needs to contain some element of adaptability to the compounds to be analysed.

Herein we briefly review an adaptable two-body Morsetype force field developed in our group based on the bondvalence approach (Adams & Rao, 2009; Adams, 2014; Chen & Adams, 2017) and we present a new software tool utilizing this force field. Finally, we illustrate its capabilities and limitations in predicting a variety of properties from a crystal structure (or a local structure model of an amorphous material, interface etc.). This method is mostly suitable for studying ion transport in solids, but can be used for other purposes such as molecular dynamics and structure verification. The software is written in ANSI C for performance and portability, and is designed in the form of a library, where modules of new functions can be extended without much difficulty. The compiled binaries for Windows and Linux, with a full list of instructions, can be found at http://www.dmse.nus.edu.sg/asn/ software.html or requested from the authors.

2. Computational method

The augmentation of the classical bond-valence approach by systematically factoring in bond softness (Adams, 2001; Adams & Swenson, 2005), and a link between the bondvalence scale and the energy scaled, have been derived in our earlier work (Adams & Rao, 2009, 2011; Adams, 2014) and will only briefly be summarized here. Moreover, both our group and several others have derived comprehensive sets of bond-valence parameters (Adams, 2014; Gagné & Hawthorne, 2015; Brown, 2016; Chen & Adams, 2017) that can be converted by this approach into Morse-type pair potentials for computationally predicting the properties of a wide range of inorganic crystal structures within the fundamental limitations of any pair potential. In this section some basics of this force field are reviewed.

2.1. Bond-valence method

In the classical bond-valence approach (Brown, 2002), an inorganic solid is described as a network of atoms of known integer oxidation states, where pairs of adjacent atoms with opposite signs in their oxidation states are connected via a bond. A detailed discussion of what constitutes an 'adjacent' atom in this context is given by Chen & Adams (2017). Each bond is characterized by a real number (called the bond valence) depending on the nature of the interacting atoms and their distance. A key postulate of the bond-valence method is the 'bond-valence sum rule' requiring that the sum of all bond valences connected to an atom approximately equals the absolute value of its oxidation state

$$V(A) = \sum_{X} s_{A-X},\tag{1}$$

where V(A) denotes the oxidation state of a cation A and s_{A-X} denotes the bond valence of the bond connecting the cation A and anion X. The value of the bond valence may naturally be expressed as

$$s_{A-X} = \exp\left(\frac{R_0 - R}{b}\right),\tag{2}$$

in accordance with the exponential drop in electron density with distance from the nucleus (Brown & Altermatt, 1985; Filsø *et al.*, 2013). Here, R_0 and b are empirical parameters that depend on the nature of A and X, while R refers to the distance between A and X. Both R_0 and b may in principle be fitted from reliably known crystal structures. If, however, as is common, only interactions in the first coordination shell are considered, the short range of experimentally observed bond distances limits the reliability of independent determinations of R_0 and b (or completely prevents it, if the reference structures contain the cation only in a single high-symmetry coordination). Therefore, b was initially proposed to be a constant of 0.37 Å regardless of A and X (Brown & Altermatt, 1985). More recently, it was demonstrated that, by incorporating information from higher coordination shells, b can in most cases still be refined independently or derived systematically and the thereby bond-softness-adapted parameter yields a closer agreement with the experimental structures (Adams, 2001; Krivovichev & Brown, 2001; Locock & Burns, 2004; Gagné & Hawthorne, 2015). It can also be shown that bond valences defined using the bond-softness-sensitive parameters are closely related to the valence electron density at the bond critical point in the Bader theory (Filsø et al., 2013; Adams, 2014).

The choices of this formalization and the parameter refinement approach obviously constitute a compromise between accuracy and simplicity, as for any individual compound it will be possible to come up with a unique set of parameters that describe the bonding situation more precisely. As an example, the softness parameter for compounds containing several types of anion is adjusted by a global equalization scheme [for details see Adams (2014)] rather than by a more local coordination polyhedral-based approach, mostly for the sake of simplicity. Still, this transferable formulation has been found to allow for a dependable quick assessment of the plausibility of crystal structure models and, as shown below, can be turned into a versatile tool for the approximate assessment of a wide range of properties of a compound from its structure model.

The actual structure of interest will always exhibit bondvalence sums that deviate to some extent from the expected values (*i.e.* from the oxidation states). The root-mean-square bond-valence sum mismatch averaged over all atoms in a crystal structure is termed the global instability index (GII) (Salinas-Sanchez *et al.*, 1992),

GII =
$$\left[\frac{\sum_{i=1}^{N} \left(\sum_{j} s_{ij} - V_{i}\right)^{2}}{N}\right]^{1/2}$$
, (3)

and can serve as a simple estimate of the plausibility of a crystal structure refinement or of the structural stability of a compound.

2.2. The softBV force field

The conventional bond-valence approach conveniently avoids the issue of calculating charges carried by ions, which can be ambiguous even with ab initio methods. As a consequence, it is not straightforward to link the mismatch of the bond-valence sum of an ion or the GII to an experimentally observable quantity or to quantify the energy required to break a bond of a certain bond valence in conventional energy units, e.g. electronvolts (eV). Similarly, the expected values for many observable quantities cannot be defined in terms of bond-valence units. Moreover, the bond-valence concept as described so far neglects Coulomb repulsions among cations or anions. In order to take into account the effect of Coulomb repulsions in analysing a crystal structure, the bond-valence term (which incorporates Born repulsion as well as Coulomb and van der Waals attraction terms for the interaction between Lewis acids and Lewis bases) and Coulomb repulsions (among Lewis acids and Lewis bases) have to be put onto a comparable scale. Hence, it appears reasonable to convert the bondvalence contribution into an absolute energy scale, which gives rise to the softBV force field. Lewis acids and bases will be identified with cations and anions in the following discussion.

The *softBV* force field consists of two terms, a Morse-type term and a Coulomb term. The Morse-type term that, as explained above, describes both the attractive interactions and a short-range Born repulsion is applied to cation–anion pairs only, while a separate Coulomb repulsion term covers cation–cation and anion–anion interactions. This implicitly postulates that Born repulsions among cations or anions can be neglected, as the Coulomb repulsions will normally prevent the occurrence of cation–cation or anion–anion distances where significant Born repulsions would occur. The three parameters characterizing the Morse potential,

$$E_{\text{bond}} = D_0 \left[\exp\left(\frac{R_{\min} - R}{b}\right) - 1 \right]^2 - D_0, \qquad (4)$$

namely, the bond breaking energy D_0 , the equilibrium bond distance R_{\min} , and the bond-softness parameter *b* characterizing the curvature of the potential around R_{\min} and hence the compressibility of the bond, are derived from the bond-valence parameters R_0 and *b* and the (average) coordination number $N_{\rm C}$. The expected value of the equilibrium bond distance R_{\min} is derived according to

$$R_{\min} = \left(0.9185 + 0.2285 \left| \sigma_A - \sigma_X \right| \right) R_0 - b \ln\left(\frac{V}{N_{\rm C}}\right), \quad (5)$$

where the numerical factors in the first term represent an empirical approach to factor in the effect of bond ionicity and atomic polarizability on the realized bond lengths, while the second term can be thought of as a bond-order term. σ_A and σ_X represent the absolute softnesses of the cation and anion, respectively, which are derived according to equation (6) from the absolute value of hardness of the ion (Parr & Pearson, 1983):

$$\sigma_n = \frac{1}{h_n} = \frac{2}{I_n + I_{n-1}},$$
(6)

where I_n is the ionization energy of the ion in oxidation state *n*. Softness data are taken from Adams (2000) or calculated from ionization energy data taken from the NIST database (NIST, 2018).

The bond-breaking energy D_0 is given by

$$D_0 = \frac{b^2}{2}k = \frac{b^2}{2}14.4 \frac{\text{eV}}{\text{\AA}} \frac{c(V_A V_X)^{1/c}}{R_{\min}(n_A n_X)^{1/2}},$$
(7)

where k, the force constant of the Morse potential at R_{\min} , depends on the oxidation states of the cation and anion, V_A and V_X , as well as on the principal quantum numbers of the cation and anion, n_A and n_X . The numerical factor in equation (7) originates from the vacuum permittivity, and the constant c = 1 if the cation A is an s- or p-block element or c = 2 if it belongs to a d- or f-block element.

While both Coulomb attraction and Born repulsion are included in the bond valence and hence in the Morse potential term, Coulomb repulsions among cations or anions are not needed for the originally intended use of the bond-valence method as a simple plausibility check of a crystal structure. For an application of the approach as an atomistic force field, the inclusion of Coulomb repulsions is obviously indispensable. In order to maintain the local nature intrinsic to the bond-valence approach, and thus its high computational efficiency, screened charges are used here. The repulsive term for a cation–cation (or anion–anion) distance R is thus, by analogy with the real-space term of the Ewald sum, given by

$$E_{\text{Coul}} = \frac{q_1 q_2}{R} \operatorname{erfc} \left[\frac{R}{f(r_1 + r_2)} \right].$$
(8)

Here, q is the effective charge carried by the ion, r is the covalent radius of the atom, erfc(x) denotes the complementary error function, and f is a tuneable screening factor that determines the extent to which the Coulomb interaction is screened so that the Coulomb attraction treated as a part of

the bond-valence term and the separately treated Coulomb repulsions remain balanced. Similar approaches have been particularly widely used in simulating inorganic glasses (see *e.g.* Zirl & Garofalini, 1990). The effective charges q on each ion are simply estimated according to equations (9*a*) and (9*b*), ensuring both that the effective charge factors in the screening effect of core electrons and that the compound is overall electroneutral:

$$q_{X,i} = \frac{V_{X,i}}{\left(n_{X,i}\right)^{1/2}} \left[\frac{\sum_{j} \frac{V_{A,j} N_{A,j}}{\left(n_{A,j}\right)^{1/2}}}{\sum_{i} \frac{V_{X,i} N_{X,i}}{\left(n_{X,i}\right)^{1/2}}}\right]^{1/2},$$
(9a)

$$q_{A,j} = \frac{V_{A,j}}{\left(n_{A,j}\right)^{1/2}} \left[\frac{\sum_{i} \frac{V_{X,i} N_{X,i}}{\left(n_{X,i}\right)^{1/2}}}{\sum_{j} \frac{V_{A,j} N_{A,j}}{\left(n_{A,j}\right)^{1/2}}}\right]^{1/2}.$$
(9b)

The summation in equations (9*a*) and (9*b*) is performed over all anions X_i and all cations A_j , where the indices *i*, *j* refer to all types of anion and cation and in the compound, and $N_{X,i}$, $N_{A,j}$ are the total numbers of anions or cations of the respective type *i* or *j* in the unit cell. If equations (9*a*) and (9*b*) lead to charges exceeding the respective oxidation state of X_i or A_j , the resulting charges are scaled down to keep the absolute value of the charges less than or equal to the absolute values of the oxidation states.

Though both the Morse term and the Coulomb repulsion normally converge well within the pre-defined standard cutoff distance of 10 Å, both terms are energy shifted to prevent any step-like change in the potential at the cut-off value. This means that, for each interaction, the corresponding interaction for R = 10 Å is subtracted.

The screening term of equation (8) introduces a new structure-specific parameter f, named the screening factor. Initially, this screening factor was either derived from pressure minimizations through molecular dynamics simulations (which requires a set of preliminary dynamic calculations for each compound) or estimated using an empirical formula based on such pressure minimizations for a reference set of oxides (which turned out not to be reliable enough for extending the application to structures containing other anions). In the present version of the software we therefore chose to calculate it in a way that is both fast and automatic, as required for high-throughput applications, but which retains adjustment to the individual compound. This is achieved by minimizing the diagonal element of the stress tensor. This approach is found to be fast and leads to more reliable results than the previously employed empirical estimate of f, as long as the underlying crystal structure model is plausible (as indicated by a sufficiently low GII value) and approximately represents a model of the local structure. It should, however, be noted that the static nature of this approach entails minor variations in the screening factor for the same chemical composition, depending on the atomic arrangement. If this is deemed to be unsuitable, e.g. in comparing the energy landscapes in different structure models of the same compound, a

It should be noted that the above estimation of the repulsive energy is coarser than the assessment of the bonding energy, but the absolute magnitude of the screened repulsive energy is typically one order of magnitude smaller than that of the bonding energy, allowing for a somewhat simplified approach while maintaining acceptable accuracy of the total energy. It should be emphasized that the softBV formulation (and therein particularly the screening factor of the Coulomb charge estimation, the bond-softness equalization method and the influence of coordination numbers on R_{\min}) strictly speaking creates a unique force field for each compound. As a result, detailed comparisons of absolute energies resulting from this method may be restricted to compounds with closely similar immobile sub-structures. This condition should certainly be met if the only change in the structure is the relocation of a mobile ion, which makes this force field particularly suitable for the study of ion transport in solids.

3. Software architecture

In order to perform an automatic analysis of the properties of various materials from local structure models (such as experimental crystal structures, ab initio simulated structures, local structure models for glasses *etc.*) using *softBV* force-field methods, a software tool has been developed under the same name, *softBV*, within our group. It comprises the necessary routines for basic tasks and these are discussed in this section. While in this work we focus on the functionality of the command-line version of the software, a graphical user interface is being developed in parallel and will be described separately (Chew *et al.*, 2019).

3.1. Modular software design

As shown schematically in Fig. 1, the software is designed as a modular library, with a small number of core transparent



Dependency graph of the modules composing the command-line version of *softBV*.

structures, an in/out library handling structural files, and some core application programming interfaces defining operations on these structures. Actual functionalities are organized into their own modules built on top of the core libraries. The software is coded in ANSI C under the standard C99. The reason for choosing C99 is that the availability of boolean types therein makes the source code more easily readable, and the native extension to complex-number types simplifies routines employing Fourier analysis.

It is worth mentioning that the exploitation of symmetry information in generating the energy isosurface is crucial, as the time required to generate isosurfaces of constant site energy for mobile ions is typically shortened by one to two orders of magnitude, if symmetry information from the input structures is utilized.

In order to prevent unnecessary barriers to using our new software tool, we have built on commonly utilized input and output file formats: input and output structure files use the Crystallographic Information Format (CIF) (Hall et al., 2005) as a universally accepted and readily available format to document the structures of inorganic solids. It is particularly advantageous that the oxidation state of each atom is an intrinsic documented property in the CIF. If the oxidation state is not specified in the input CIF file, the program just assumes default oxidation states for the elements present, but does not conduct elaborate plausibility checks for the applicability of these default values in the specific compound. Thus, the user should ensure that the oxidation states of all ions are indicated. Non-integer oxidation states will be converted internally by the program to partial occupancies by a mixture of ions of that element with the nearest integer oxidation states.

Site energies for the mobile ion are stored as a discrete representation of a field in the *GAUSSIAN16* cube file format (Frisch *et al.*, 2016). A wide range of popular software tools are already available to visualize isosurfaces of constant bond-valence site energy superimposed on structure models from such cube-format grid files, including *e.g.* the commercial *Materials Studio* (Dassault Systèmes BIOVIA, San Diego, California, USA), or *VESTA* (Momma & Izumi, 2011) which is freely available for non-commercial applications. For further pieces of information logged by the software, the comma-separated values (csv) format is used so the information can be conveniently handled not only by the associated user interface (Chew *et al.*, 2019) but also by standard data-analysis software or be integrated into user-defined scripts.

3.2. Available functionality

3.2.1. Bond-valence sums. The most basic function of the software includes the checking of the bond-valence sums on each ion in a given structure, provided that the corresponding *softBV* bond-valence parameters exist in the program's builtin database. The latter is largely identical to the *softBV* parameters, including expected coordination numbers, recently published by Chen & Adams (2017). The bond-valence approach presumes that the bond-valence sum should match the oxidation state of the considered ion. In practice, there is always a minor deviation, which is expressed as the GII. The larger this GII value is for a compound, the less plausible is the structure model or the less structurally stable is the compound. A GII value larger than 0.2 is commonly assumed to indicate a plausibility issue for the structure model, unless there are special reasons (high-pressure structures, extremely high temperatures, chemically strained structures) explaining the deviating bond lengths (Brown, 2009). It may be noted in passing that heavier ions tend to have a higher impact on the GII, so a scaling by the principal quantum number n_i in a modified stability index MII as shown in equation (10) might lead to a more appropriate judgment of structural plausibility (Adams, 2014):

MII =
$$\left[\frac{\sum_{i=1}^{N} \left(\sum_{j} s_{ij} - V_{i}\right)^{2}}{n_{i} N}\right]^{1/2}$$
. (10)

Still, for the time being the more widely employed GII is implemented in the software.

3.2.2. Generation of bond-valence site energy landscape for a mobile species. The software, upon importing the crystal structure data of interest from a CIF file and indicating the mobile species to be assumed, divides the unit cell of the structure into a uniformly separated grid. It then calculates the interaction energy for a probe ion of the indicated mobile species placed at any of the grid points with the whole structure except for other ions of the same type. This interaction energy is termed the bond-valence site energy (BVSE). In order to ensure that the BVSE values at common special positions in a crystal structure (at fractional coordinates 0, 1/2, n/3, n/4, n/6, n/8, n/12) are explicitly included as grid points, the software adjusts the default value (0.1 Å) or user-specified value of the grid resolution slightly to enforce the number of grid points in each dimension to be always a multiple of 24. The resulting array of BVSE values at each grid point forms a discrete representation of the energy landscape that is then stored in the above-mentioned cube files and used in a variety of subsequent analyses.

We are aware that the popular approach in DFT analyses of ion-migration barriers is to predefine the start and end points of the ion-migration step and apply the nudged elastic band (NEB) method (Mills & Jónsson, 1994) for assessing the migration barrier between these points. Two issues should be addressed with the application of the NEB method:

(i) The determination of the initial and final structures is often not justified;

(ii) The chosen path along which the NEB method is performed may not be the critical path that controls the overall activation energy of ion migration.

The advantage of our comprehensive analysis of the entire energy landscape is that it automatically generates a complete picture of the conduction pathway. While with respect to the energy-barrier height values along a given pathway segment it is arguably not as accurate as the DFT NEB method, our approach reveals all the paths that exist for the given structure model and allows for an approximate assessment of the relative height of the barriers, which should be sufficient to prescreen databases of crystal structures and provides the necessary basis for further refinement of the relevant paths in shortlisted candidate structures by more precise methods.

As mentioned above, a screening factor needs to be provided as an additional parameter to calculate the repulsive energy. The originally chosen approach of running a separate set of constant volume constant temperature (NVT) molecular dynamics (MD) simulations with varying screening factors to identify the screening factor that leads to zero pressure yields reliable results, but turned out to be a toolengthy process for screening applications, whereas an empirical prediction of f values based on MD simulations for a set of reference structures turned out not to be reliable enough. Therefore, the screening factor f is, in the current version of *softBV*, estimated automatically by varying it iteratively to achieve a pressure close to zero in fast staticpressure relaxation calculations. As shown in Fig. 2, screening factors obtained in this way agree reasonably well with those obtained via MD for the same set of reference compounds with low GII.

In our software, the automatic calculation of the screening factor is triggered by default whenever a calculation of energy is required. Sometimes this might not be possible (*e.g.* due to a lack of bond-valence parameters for interactions within the immobile sub-structure), and in these cases a default value of f = 0.75 is enforced. An automatic calculation may also not be desirable, if different structure models of the same compound are to be compared. In these cases the user can supply a common fixed screening factor for all structures to be compared. It may be noted that applying the above static method of estimating f to incorrect structure models (with a high GII) will lead to a stronger bias on the f value than if f was refined from a series of MD simulations, as the structure model is not optimized at the same time. In such cases it may be advisable to recalculate f after a geometry optimization.

For crystal structures representing average structure models, including partially occupied alternative sites for the same ion that cannot be occupied in the same unit cell (*e.g.*



Figure 2

The correlation between screening factors determined from a series of molecular dynamics simulations and from static-pressure variation for a set of 32 Na compounds with GII < 0.10.

average structures of ferroelectric multidomain crystals), the fictitious interaction between the partially occupied sites would bias the calculation. Unusually low values of f that may result from such a bias are replaced by a default value of f = 0.75 and a warning is issued.

3.2.3. Geometry optimization and NEB. Since the determination of the screening factor requires an internal implementation of some geometry-evaluation routines, we extend it into a module capable of some simple geometry-optimization tasks. Firstly, a relaxation of the atomic positions using the steepest descent method with the softBV force field can be performed. While in principle the unit-cell size and shape might be relaxed, two-body force fields generally have little shear resistance. Therefore, at this point only the unit-cell lengths are relaxed, while the unit-cell angles are fixed. In this geometry optimization it is possible to fix certain ion positions. This is achieved by editing the input CIF file to add an additional refinement flag for each ion. Note that in that case the energy output will not include the interactions among the immobile ions. This setup also proves useful in calculating surface energies.

Our program also includes a simple implementation of the NEB routine (Mills & Jónsson, 1994) that accepts two input structures as initial and final images, and interpolates to create initial guesses of sampling images. This routine was included to verify the correctness of the scaling due to relaxation, as well as to facilitate a comparison of the BVSE findings on migration pathways with the NEB results from DFT simulations. Since the NEB method is based on the energy of specific atomic arrangements of the mobile species at the starting and target sites, while the energy-landscape analysis eliminates the influence of repulsions among the mobile ions, comparing the softBV NEB results with the BVSE energy-landscape results may also be used to assess the influence of repulsions among the mobile ions on the migration barriers. As the NEB approach requires manual manipulation of the local site occupancies as well as manual selection of the start and end points, it appears less suitable for the intended databasescreening applications.

3.2.4. Automated path analysis of the energy landscape. While cube files can be directly visualized to identify low energy-barrier migration pathways, it is convenient, especially for high-throughput screening, to design an algorithm that reduces the energy isosurface to pathways, so that equilibrium and interstitial sites and their connectivity via transition states can be clearly identified without the need for human intervention. The general idea behind this algorithm is first to identify voxels in the energy landscape that are local minima or saddle points, and then to connect them by tracking from saddle-point voxels to minima voxels.

As illustrated in Fig. 3, each voxel is analysed with respect to its surrounding 26 voxels [eight voxels in this twodimensional (2D) example]. These 26 environment voxels can be categorized into two groups according to whether their energy is (i) higher than or equal to, or (ii) lower than that of the central voxel. Each group is further broken into clusters in such a way that voxels from distinct clusters must not share any face or edge (ignoring common faces or edges with the central voxel). If the central voxel is surrounded by only one higher energy cluster, it is clearly a local minimum. If it is surrounded by two lower clusters and at least one higher cluster, it is a typical saddle point that connects two local minima. A general pathway point is a voxel surrounded by one higher cluster and one lower cluster. Other situations are possible, such as a saddle point that branches to three or more paths, each leading to some local minimum. However, according to our experience with many real structures tested, this rarely happens except in high-energy regions, which do not affect ion migration in practice.

From each saddle-point voxel with two (lower energy cluster) branches, we can travel to the voxel of lowest energy within each of the two lower energy clusters. Continuing this process until we finally reach a local minimum, all voxels traversed form a continuous path that connects the saddle point to two local minima. The section of a path connecting two local minima via a saddle point is termed a pathway segment. With all pathway segments identified, a topological graph is generated from the energy isosurface, where vertices represent local minima (sites that may be occupied) and edges represent saddle points that connect two local minima (hopping trajectory between sites). As an example, Fig. 4 shows the Li⁺ paths in Li₅La₃Ta₂O₁₂ (Thangadurai *et al.*, 2004; Cussen, 2006). The analysis demonstrates that, in this garnetrelated solid electrolyte, a low-energy migration path is formed based on alternating hops between tetrahedrally coordinated Li(1) and octahedrally coordinated Li(2), while a much higher energy barrier is required for direct Li(2)-Li(2)hops. Fig. 4(d) also shows the energy profile for an NEB analysis of possible hops between Li(2) atoms, and the lowest energy pathway is accordingly observed to be the pathway via the Li(1) site. Note that the overall migration barriers found by both approaches are nearly identical, while the detailed energy profiles differ, mostly because of the specific charge distribution that had to be assumed in the NEB calculation. Here, effectively, the partial occupancy of Li(1) was retained in the NEB calculations, so that the energy profile includes the displacement of the Li on the Li(1) site and therefore does not exactly reach the local minimum for a vacant Li(1) site that is observed in the energy-landscape analysis.



Figure 3

A schematic drawing of the clustering environment of a central saddlepoint voxel C in a 2D plane. (*a*) The energy values of each voxel. (*b*) The surrounding voxels classified into two lower energy clusters L1 and L2, and two higher energy clusters H1 and H2.

Besides information on which local minima are connected to which saddle points, and the corresponding bond-valence site energies and fractional coordinates, another crucial property to be obtained for saddle points during their identification is their index. The index of a saddle point or of the corresponding path segment is defined as the net number of times the path segment crosses the periodic boundary of the unit cell. This number is an integer and can be negative, which means the crossing happens in the reverse direction to the direction arbitrarily defined as positive. When a path crosses the boundary of the unit cell in the positive direction, the index is increased by one; when it crosses the same boundary in the reverse direction, the index is decreased by one. Since the boundaries for each unit cell correspond to the three principal directions, the path index is effectively an array of three integers, each evaluated independently by the protocol discussed above. The method is applicable to all types of path, regardless of their exact shape and of how many times they cross boundaries.

Each migration pathway network can be represented by a cycle in the conduction graph. Such a cycle always carries three indices, which are the sums of the respective indices of each of its constituent pathway segments. Obviously, any cycle with at least one of its three indices non-zero is equivalent to a one-dimensional (1D) percolating network that can be represented in a unit cell, and any cycle with all three indices zero is a non-percolating network that confines the ion locally. When two 1D networks meet at some point and their indices are not parallel, or more accurately their index arrays span the whole 2D linear space, a 2D pathway network is formed. This conclusion can be readily extended to three dimensions, defining the 3D network as a set of 1D networks that are connected and whose index arrays span the 3D linear space.

Analysing all cycles in the conduction graph would enable us to identify the shape of all percolating networks. However, finding all simple cycles in a graph is at least a 'nondeterministic polynomial-time hard' (NP-hard) problem (Leeuwen, 1998) whose complexity grows exponentially with the size of the graph. A preliminary test of a structure with a volume of about 1000 $Å^3$, a rather common size for unit cells of inorganic crystal structures, consumes more than 4 GB memory and fails to reach its end after two days on a typical desktop PC, using Johnson's algorithm for searching cycles (Johnson, 1975). Such performance is unacceptable for a screening tool and alternative methods are needed to analyse the percolating networks. One currently implemented method is still based on the composition of cycles. The idea for handling the complexity explosion is to select a small number of cycles for the analysis in the way mentioned above, instead of surveying all the cycles. To do so, a random spanning forest of the conduction graph is generated first, containing all vertices V in the graph and V - k edges, where k is the number of connected components in the graph. By adding any one of the remaining E - V + k edges to the forest, a unique cycle is formed. In total only E - V + k cycles are kept in the memory, which is a manageable number. Every combination of two such cycles is tested for the formation of a 2D network, and

research papers

every combination of three is tested for a 3D network. This is not an accurate way of evaluating the true percolation thresholds of the structure, because a large number of cycles are left out. In practice it is able to report the actual percolation thresholds for most but not all compounds, and at least it provides an upper limit for each percolation threshold. An alternative percolation analysis approach that addresses the limitations of the percolation analysis method adopted in *softBV* will be discussed elsewhere (Wong *et al.*, 2019) and implemented in the user-interface version *softBV-GUI* (Chew *et al.*, 2019).

3.2.5. Automatic analysis of attempt frequencies in the energy landscape. Apart from the height of the energy barriers along the conduction pathways and the site-occu-

pancy distribution, another factor affecting the absolute conductivity is the attempt frequency for hops of the mobile ion depending on the energetic environment of the local minimum. To gain some insight into the different attempt frequencies of the same ion at different sites, *softBV* includes the option of conducting a Fourier transform analysis of the generated energy landscape. With the assumption of harmonic oscillation at the bottom of each potential well, the attempt frequency can be related to the second derivative of force by

$$\nu = \frac{1}{2\pi} \left(\frac{k}{m}\right)^{1/2},\tag{11}$$





An example of a bond-valence site energy isosurface and conduction pathway graphs generated from the crystal structure of a garnet-type Li⁺ ionconducting solid. (*a*) The crystal structure of $Li_5La_3Ta_2O_{12}$ (LLT). Brown octahedra represent TaO_6 , blue atoms represent La^{3+} and green atoms represent Li⁺. (*b*) The isosurface of constant Li⁺ bond-valence site energy 0.4 eV above the site energy of the lowest energy Li⁺ site. Green spheres mark the local BVSE minima representing possible Li⁺ sites. (*c*) The topology of the Li⁺ migration pathways in LLT. Green spheres, corresponding to the BVSE minima, indicate the nodes of the pathway network and dotted lines indicate the pathways. (*d*) A reaction pathway diagram (blue triangles) showing that the 3D pathway network in LLT is based on hops between Li(1) and Li(2) sites with a barrier <0.4 eV, while additional direct hops between Li(2) sites require overcoming a 0.8 eV higher barrier. For the percolating Li(2)–Li(1)–Li(2) path, green triangles show the corresponding NEB migration-barrier analysis. The NEB analysis finds essentially the same overall migration barrier, though the path differs in its details as, in contrast with the energy-landscape analysis, it factors in Li–Li repulsions for the specific Li occupancy of the manually selected start and end models.

where m is the mass of the conducting ion and k is the spring constant or second derivative of the energy.

In order to work out the second derivative of the energy, it is necessary to differentiate the discrete representation of the energy field. While it is possible to rely on the finitedisplacement method, since the analytical formula for the energy is known, this would couple the generation and analysis of the energy field, which would complicate updating and maintenance of the software, as well as limit the ability to analyse energy fields generated by other tools.

Therefore, we chose to determine the frequency via Fourier analysis with the aid of the FFTW3 package (Frigo & Johnson, 2005). Following differentiation in Fourier space, the Hessian matrix can be constructed at each grid point. Combined with the first derivatives, the eigenvalues of the Hessian matrix suggest whether a point is a critical point, and if so which type of critical point it belongs to. In addition, each eigenvalue indicates a vibration frequency, associated with an eigenvector indicating the direction of the vibration. Therefore, the relevant vibration frequency will differ along different paths connecting to the same local minimum. It is hard to verify this, as experimental vibration frequencies are usually reported as a single average number or as a direction, but not as a sitedependent dispersion curve. Since the lowest vibration frequency usually corresponds to the lowest barrier (Meyer & Neldel, 1937), we consider the lowest vibration frequency to dominate over all other frequencies.

In order to validate whether our routine yields plausible predictions, vibration frequencies in a number of binary compounds whose bond-valence parameters are available in the *softBV* parameter set are now calculated and compared with a compilation of experimental results (Bilz & Kress, 1979) or ab initio computational results. Two sets of experimental vibration frequencies are used: the transverse optical mode ω_{to} is extracted from the lowest non-zero point along the γ direction in the dispersion curve, while the longitudinal optical mode ω_{lo} is the highest point. Different patterns of combining eigenvalues into frequencies fitted against experimental values are explored, and we find that ω_{to} is most reliably fitted to the minimal eigenvalues among the six eigenvalues from both cation and anion, and ω_{lo} is best fitted to the geometric average of the combined frequencies given by

$$\omega_{\text{low}} = 2 \min\left\{ \left[\frac{\min(k_{c,1}, k_{c,2}, k_{c,3})}{m_c} \right]^{1/2}, \left[\frac{\min(k_{a,1}, k_{a,2}, k_{a,3})}{m_a} \right]^{1/2} \right\},$$
(12)
$$\left[(k_{a,3})^{1/2} + (k_{a,3})^{1/2} + (k_{a,3})^{1/2} + (k_{a,3})^{1/2} + (k_{a,3})^{1/2} + (k_{a,3})^{1/2} \right]^{1/2} \right]^{1/2}$$

$$\omega_{\text{high}} = \left[\frac{(\kappa_{\text{c},1}) + (\kappa_{\text{c},2}) + (\kappa_{\text{c},3})}{(m_{\text{c}})^{1/2}} \frac{(\kappa_{\text{a},1}) + (\kappa_{\text{a},2}) + (\kappa_{\text{a},3})}{(m_{\text{a}})^{1/2}} \right] ,$$
(13)

where $k_{c,i}$ is the *i*th eigenvalue from cation vibration, $k_{a,i}$ is the *i*th eigenvalue from anion vibration, m_c is the mass of the cation and m_a is the mass of the anion. When fitting ω_{low} against ω_{to} , the calculated frequency is about half the literature value, which is compensated by the factor 2 in equation

(12). This might be due to the twofold minimization in this equation, which implies that only one frequency along one direction is used, thus systematically underestimating the real vibration frequency. When fitting ω_{high} against ω_{lo} , the slope of the linear correlation is almost one, since all three eigenvalues are factored in.

Since one might suspect that the vibration frequency is simply controlled by the mass of the ions in the system rather than by the energy landscape, in Fig. 6 we fit the vibration frequency to the literature data as purely mass dependent by assuming a constant eigenvalue for all compounds. The virtual fixed eigenvalue is adjusted so that the slope of the regression line is equal to unity in order to facilitate a comparison of the resulting scatter with the corresponding graphs in Fig. 5. A tendency can still be seen but the scatter in Fig. 6 is much wider, so clearly both the ion mass and the energy landscape play a significant role in determining the vibration frequency.

3.2.6. Towards estimation of absolute ionic conductivity. The most relevant functional property to be derived from the structure of an ion-conducting solid is, after all, its ionic conductivity. The software tool described here enables us to identify from a crystal structure:

(i) Equilibrium sites and interstitial sites for the mobile ion,

(ii) Their respective site energies and occupancies,

(iii) The topology of the relevant migration paths in the energy landscape,

(iv) The migration barriers along these paths.



Figure 5

(a) The calculated frequency $\omega_{\rm low}$ according to equation (12) versus the literature frequency transversal optical frequency $\omega_{\rm to}$. (b) The calculated frequency $\omega_{\rm high}$ according to equation (13) versus the literature frequency longitudinal optical frequency $\omega_{\rm lo}$.

research papers

This encourages a trial to estimate directly the absolute value of the ionic conductivity at a given temperature from the structure model. As our aim is to design a tool that is sufficiently fast to allow for high-throughput screening, we based this attempt on the rejection-free KMC algorithm (Bortz *et al.*, 1975; Gillespie, 1977) where the procedure is generally as follows.

(i) Construct a directed graph of conduction where each vertex represents a site (interstitial or equilibrium) and each edge represents a pathway joining two sites, whose weight is the migration barrier $(E_m + \Delta E)$. Herein, E_m is the constant contribution to the migration barrier from the bond-valence site energy landscape, *i.e.* the contribution from interactions of the mobile species with the immobile sub-structure, while ΔE refers to the difference between the repulsion that the selected mobile atom feels at the starting point and at the bottleneck towards the target site from interactions with all other ions of the same mobile species.

(ii) Generate a supergraph corresponding to a supercell where each lattice parameter is at least 25 Å.

(iii) Assign sites of lowest energies occupied and perform an equilibration phase.

(iv) Run the KMC and, at each step, adjust the migration barrier based on the repulsion energy from the instantaneous arrangement of mobile ions on nearby occupied sites.



Figure 6

(a) The calculated frequency ω_{low} versus the literature frequency transverse optical frequency ω_{to} . (b) The calculated frequency ω_{high} versus the literature frequency longitudinal optical frequency ω_{lo} . The globally fixed eigenvalues are chosen to ensure a slope of exactly one for both cases.

(v) For hops to vacant target sites, calculate the transition rate based on the Arrhenius equation,

$$p = \omega \exp\left(-\frac{E_{\rm m} + \Delta E}{k_{\rm B}T}\right),\tag{14}$$

where p refers to the transition rate from site i to site j, ω is the vibration frequency, $E_{\rm m} + \Delta E$ is the effective migration barrier, $k_{\rm B}$ is the Boltzmann constant and T is the temperature. To calculate the vibration frequency ω , each eigenvector is projected onto the direction of the path segment, and the dot product is used to scale the corresponding eigenvalue. The sum of the three scaled eigenvalues yields ω .

The conductivity σ is then calculated from the Nernst– Einstein equation,

$$\sigma = \frac{(ze)^2 cD}{k_{\rm B}T},\tag{15}$$

where σ refers to the ionic conductivity, z is the oxidation state of the conducting ion, e is the electronic charge, c is the concentration of the conducting ion and D is the diffusion constant,

$$D = \frac{\langle x^2 \rangle}{6t}.$$
 (16)

Here, t is the time and $\langle x^2 \rangle$ is the mean-square displacement averaged over all ions of the mobile species. It should be noted that the assumption of validity of the Nernst–Einstein relationship implies that subsequent hops are assumed to be uncorrelated, which in detail will not always be the case.

3.2.7. Surface-energy calculation. Considering the relative simplicity and low computational cost of our force field, it can be employed for certain tasks that require intensive resources when DFT methods are used. One such task is the determination of the surface index given a bulk structure. The ordinary procedure for calculating the surface energy of a particular index is:

(i) Perform a relaxation of the bulk structure and obtain the bulk energy per formula unit E_{bulk} ;

(ii) Construct a slab of surface area A along the given index of at least 20 Å, where the positions of the atoms in the lower 10 Å are fixed, while those in the upper 10 Å are allowed to relax;

(iii) Reduce the electric dipole moment of the slab as far as possible by moving some ions from the mobile surface to the immobile one or vice versa (Tasker, 1979), generating a number of reconstructed surface structures;

(iv) Perform a relaxation and calculate the total energy minus the energy among the immobile part, then divide by the number of formula units in the mobile part to give E_{slab} . The surface energy is then given by

$$E_{\rm surface} = \frac{E_{\rm slab} - E_{\rm bulk}}{A},\tag{17}$$

where all variables are normalized by the number of formula units.

This is somewhat different from the formula employed in DFT calculations, where the double-slab technique is

commonly used. There, the two surfaces of the slab are chosen to be symmetric and the whole slab is allowed to relax so that E_{surface} involves dividing by 2A (Sun & Ceder, 2013). This is because the energy calculated using DFT cannot be assigned to individual atoms and therefore much larger structures are required.

4. Application examples

4.1. Benchmarking of energy barriers predicted by softBV

The energy barriers estimated by the *softBV* method are conceptually migration barriers, though they may include the defect-formation step, if the BVSE for the saddle point of a Frenkel defect formation is lower than the BVSE of the highest energy saddle point along the lowest energy percolating path. The most readily available experimental comparison data are based on the temperature dependence of ionic conductivities determined by impedance spectroscopy, but the activation energies resulting from these data are sums of migration barriers and defect-formation energies. The available literature data on experimental activation energies are, however, severely biased towards compounds with relatively low defect-formation energy barriers and intentionally or unintentionally doped materials, where no defect formation is necessary. The most straightforward comparison for the softBV values is thus with migration barriers determined by NEB analysis in a series of DFT calculations where available.

As seen from Fig. 7, in general there is good agreement between the migration energies predicted by BVSE and DFT methods. Still, it may be cautioned that, for some compounds with an extremely high concentration of the mobile species (so that the actual site energy is dominated by the interactions among the ions of the mobile species), the *softBV* approach tends somewhat to underestimate the barriers. For the same reason, the approach is more reliable for estimating the mobility of cations than for estimating anion mobility in the common case where large anions control the overall packing of ions in a structure. At very high migration barriers > 1 eV, which are irrelevant for practical applications anyway, there seems also to be a tendency that the BVSE underestimates the barriers, as such high energy barriers can only be overcome by extensive structural relaxation of the host structure. It may also be noted that, among the Mg compounds in Fig. 7, one refers to an experimental data point by Anuar et al. (2014) reporting an extremely low migration barrier of 0.1 eV for Mg^{2+} in the NaSICON-type structure $Mg_{0.5}Zr_2(PO_4)_3$, which is not in line with either our BVSE model ($E_{\rm mig} \simeq 0.4 \, {\rm eV}$) for this monoclinic LT-NaSICON phase nor with the low magnitude of the conductivity observed in the same work.

For the case of silver ionic conductors, where abundant experimental literature data are available from vibrant research activity in the 1970s to 1990s, but hardly any DFT data, Fig. 8 shows a comparison of effective migration barriers from our BVSE models with experimental activation energies. For this purpose the migration barrier for 3D percolating paths was assumed to be the effective migration barrier, unless the migration barrier for 2D percolating paths was < 2/3 of the 3D barrier, and the barrier for a 1D percolating path was only assumed to be the effective migration barrier if the barriers for 2D and 3D percolating paths were at least 2.7 times higher. Unexpectedly, the quality of the correlation in Fig. 8 is comparable with that between the DFT- and BVSE-based migration barriers in Fig. 7. It should be cautioned that this probably just means that the available experimental Ag⁺ conductivity data almost exclusively refer to compounds where the (Frenkel) defect-formation energies are either small or are included because they occur along pathways within the observed migration paths, as explained above.





Comparison of migration barriers in (*a*) 57 Li⁺ ionic conductors and (*b*) 19 Mg²⁺ ionic conductors from our BVSE models with literature DFT NEB data. Dashed lines in each of the graphs indicate a 1:1 correlation and solid lines are polynomial fits to the data. The graph for Mg²⁺ is a redrawn and extended version of data previously published by Nishitani *et al.* (2018).

4.2. Correlation between energy barriers and absolute conductivities

In a simplistic model of an otherwise defect-free solid, where the ionic conductivity relies on the motion of one type of mobile ion of concentration $c_{\rm ion}$ via a single type of uncorrelated jump (of distance *a* with direction-independent attempt frequency v_0) on a regular percolating path of dimensionality *d* with a concentration $c_{\rm sites}$ of equivalent sites for the mobile species, the expected conductivity is given by equation (18) under the additional assumption that the exponential terms containing the defect-formation enthalpy $\Delta H_{\rm def}$ and entropy $\Delta S_{\rm def}$, and the migration enthalpy $\Delta H^{\neq} = E_{\rm mig}$ and entropy $\Delta S^{\neq}(d)$, are high enough that only a small fraction of the attempted hops are successful:

$$\sigma = \frac{c_{\rm ion} \left(1 - \frac{c_{\rm ion}}{c_{\rm sites}}\right) z^2 F^2 a^2 \nu_0}{2dRT} \exp\left[\frac{\Delta S^{\neq}(d) + \frac{\Delta S_{\rm def}}{2}}{R}\right] \\ \times \exp\left[-\frac{\Delta H^{\neq} + \frac{\Delta H_{\rm def}}{2}}{RT}\right]. \tag{18}$$

Here *R* is the universal gas constant. Equation (18) further implies that a site can only be occupied by one ion at a time and that hops to occupied sites are a priori unsuccessful, which is obviously not always the case as it would rule out the often crucial interstitialcy mechanism. Writing the activation entropy as $\Delta S^{\neq}(d)$ here emphasizes that its value will depend on the pathway dimensionality, so that in general the conductivity σ for a given migration barrier will increase with the dimensionality of the pathway network (*i.e.* with the number of potential target sites) despite the *d* in the denominator of the prefactor. The condition that only a small



Figure 8

Comparison of migration barriers in 112 Ag^+ ionic conductors from our BVSE models with literature data of experimental activation energies. The dashed black line indicates a 1:1 correlation and the solid coloured line is a linear fit to 109 of the data marked by diamond symbols, omitting three outliers (marked by circles): Ag_3YCl_6 , with a barrier that is considerably higher than the reported activation energy, and two high-temperature structure models of superionic HT-Ag₂S, for which the BVSE model predicts even lower migration barriers than observed experimentally.

fraction of the hops should be successful in this model factors in that the Arrhenius-type exponential dependence of the success rate on the barrier height is itself a large-barrier approximation, while a more general treatment may require a Fermi–Dirac type correlation between site energy and site occupancy, as visible *e.g.* from the apparent decrease in activation energy commonly observed in high-temperature data of superionic compounds (Adams, 2000).

In reality, the average jump distance, attempt frequency, concentration of mobile ions, activation entropy, defectformation enthalpy and entropy, and effective pathway dimensionality will – at least in a statistical sense – exhibit some degree of correlation with the effective migration barrier, the considered direction *etc.*, rendering the prediction of an absolute conductivity at a given temperature (*e.g.* room temperature) a complex problem. From a purely empirical point of view, one may try to circumvent this issue and instead indicate an expected range of conductivities based on existing data of other solid electrolytes with the same mobile ion. In Fig. 9 the literature data on the observed or extrapolated experimental room-temperature conductivities for 265 Na⁺ ionic conductors are plotted against the reported effective activation energy.

For the simplified model, a scatter around a linear variation with slope -1 would be expected, but the experimental data seem to cluster around two limiting activation-energy dependencies that can each be approximated by the functional form of equation (19) that could be used as a smooth 'activation function' in a neural-network property predictor:

$$\ln(\sigma_{\rm RT}) = a_1 - a_2 \frac{E_{\rm act}}{k_{\rm B}T} - a_3 \ln\left[1 + \exp\left(\frac{E_{\rm act}}{k_{\rm B}T} - a_4\right)\right].$$
 (19)

Here, a_1 represents the limiting ionic conductivity for migration barriers that are low compared with the thermal energy, $-a_2$ and $(-a_2 - a_3)$ represent the limiting slopes in the low and high migration-barrier domains, respectively, and a_4 represents the threshold activation-energy barrier between



Figure 9

The correlation between the experimentally observed activation energies of 265 Na⁺ ion-conducting solids and the observed (or extrapolated) room-temperature conductivity σ_{RT} . The two dashed lines are guides to the eye according to the four-parameter equation (19).



Figure 10

(a) A polyhedral representation of the structure of Li_6PO_5CI . (b) Isosurfaces of constant BVSE for Li^+ in the structure model. Purple regions have a BVSE energy less than 0.27 eV above the global minimum and grey regions have an energy 0.47 eV above global minimum. (c) A pathway topology graph for Li^+ .

the two domains. The equation can be understood as a generalization of the *softplus* activation function $y = \ln[1 + 1]$ exp(x) (Dugas *et al.*, 2000) or as a differentiable approximant to a 'leaky' rectifying linear unit $y = \min(a_1x, a_2x)$. In the example of the Na⁺ conductors in Fig. 9, a_4 for both fits corresponds to an activation energy of ca 0.4 eV and a_1 to a limiting room-temperature conductivity of about 30 mS cm^{-1} , and the limiting slopes are about 0.2-0.35 in the superionic regime or 0.8–1.2 for higher activation energies. It is tempting to identify the limiting slopes of this generalized activation function with extrinsic conductivities for which $E_{\rm act}$ is determined by the migration barriers alone and with intrinsic conductivities where significant defect-formation energies affect E_{act} , but in many cases the literature data cannot be unambiguously assigned to either case. Still, the deviation of the magnitude of the average slopes from unity demonstrates some degree of correlation between the pre-exponential factor and the effective activation energy. The magnitude of the slope is clearly below unity in the superionic low migration-barrier region, yet can be below or above unity for the high activation-energy domain.

4.3. KMC analysis of energy landscape and conduction pathways

From the empirical considerations in the previous section, it may appear obvious that the derivation of the absolute conductivity at a given temperature from the structure and BVSE energy landscape can, in the best case, only be approximate and to achieve that would have to account for a multitude of possible complex transport mechanisms. In order to illustrate how, in principle, the energy landscape can be constructed and analysed by the *softBV* software, argyroditetype Li₆PO₅Cl is used as an example. Lithium argyrodites Li_{7-x}PS_{6-x}X_x (X = Cl, Br) are known to be fast Li⁺ conductors with high ionic conductivities (Deiseroth *et al.*, 2008; Rayavarapu *et al.*, 2012) that can be used as solid electrolytes in all solid-state batteries (Chen & Adams, 2015), but they are structurally only metastable and are unstable against hydration (Chen *et al.*, 2015). A natural approach to enhancing the stability would be to move from sulfides to the corresponding oxides. Indeed, Kong *et al.* (2010) were able to synthesize and characterize $\text{Li}_6\text{PO}_5\text{Cl}$, which turned out to be anion-ordered and a poor conductor. In this section, the calculation of ionic conduction properties based on the experimental structure and *softBV* method is performed and the results are compared with the experimental data.

As the isosurface of constant $BVSE(Li^+)$ in Li_6PO_5Cl in Fig. 10 demonstrates, all experimentally observed Li positions



Figure 11

Conductivities calculated by KMC compared with experimental values. The solid blue curve and filled circles show all the simulated KMC conductivities. The solid red line marks effective KMC conductivities corresponding to long-range conduction (from high-temperature KMC simulations) and the dotted red line its extrapolation down to lower temperatures. The green curve and triangles show experimental conductivity data.

coincide with local minima of BVSE(Li⁺). The percolation analysis yields a barrier of $E_{\rm mig} = 0.50$ eV, which might be sufficiently low to classify Li₆PO₅Cl as a moderate conductor, but the global minimum does not correspond to any experimentally occupied Li site. It is suspected that, in the case of Li₆PO₅Cl, the activation energy is strongly affected by coulombic repulsion so it deviates from the percolation barrier, and the strong repulsion also forbids the global minimum from being occupied. A KMC simulation was conducted and the results are compared with experimental values in Fig. 11.

The KMC results suggest that Li₆PO₅Cl is indeed a poor conductor and the KMC simulations at room temperature correspond to localized Li⁺ movements. In order to obtain conductivity and activation energy at room temperature, the simulation temperature was raised until the average displacement of Li⁺ was larger than 4.0 Å. KMC gives an extrapolated result that $\sigma = 3.31 \times 10^{-12}$ S cm⁻¹ and $E_{act} = 0.95$ eV, while the experimental values are $\sigma = 1.41 \times 10^{-9}$ S cm⁻¹ and $E_{act} = 0.66$ eV. The equilibrated occupancies calculated from KMC are consistent with the experimental structure such that the global BVSE minima are not occupied due to Li–Li repulsion.

We now attempt to benchmark our results against our compilation of literature conductivity data for Li^+ and Ag^+ solid conductors. As a quick measure to check whether the literature data contain implausible data points in Fig. 12, the natural logarithm of conductivity at 300 K is plotted against activation energy (for compounds where the conductivity can be measured at high temperature only, the values shown are inevitably based on extrapolation). The results yield a limited scatter around a monotonic fitting curve according to equation (19), in line with the empirical finding that low experimental activation energies (which are not to be directly identified with migration barriers) are generally associated with high ionic conductivity. Closely similar trends are also found in the KMC results, which one might tentatively interpret as that the KMC algorithm yields physically plausible activation energies and



Figure 13

A comparison of simulated conductivities from the KMC method with literature values for Ag^+ conductors (red triangles) and Li^+ conductors (blue squares). The broken line marks a 1:1 correlation.

absolute conductivities. Still, for a given activation energy the conductivity may vary over five orders of magnitude.

In Fig. 13, KMC-simulated conductivities for a range of Li⁺ and Ag⁺ ion conductors are compared with experimental values from the literature. While there is a clustering of data points near the 1:1 line, the overall scattering is still wide and numerous outliers are present. It remains inconclusive whether the present implementation of the KMC approach has sufficient predictive power to identify fast ionic conductors. The limited agreement between the experimental literature data and the KMC model suggests, on one hand, that the characterization of experimental samples with respect to point-defect concentrations and interface effects is often unsatisfactory. On the other hand, the KMC model does not yet factor in all the relevant modes of correlated ion transport. Moreover, the chosen method of assigning the initial distribution of mobile ions based on the site energy may sometimes bias the observed concentration of mobile defects.



Figure 12

Plots of literature experimental (open symbols) or KMC simulated (filled symbols) ionic conductivity data at 300 K (partially extrapolated from higher temperatures) versus activation energy for (a) Ag^+ conductors and (b) Li^+ conductors. Solid lines represent fits according to equation (19)

4.4. Surface-energy calculation

In this section we investigate the surface energies of various surface indices in five compounds, LiCoO₂, NaCl, BaTiO₃, MgO and AlN, using the relaxation module within *softBV*, and compare the results with literature values given by DFT calculations (Kramer & Ceder, 2009; Nakayama *et al.*, 2013; Eglitis *et al.*, 2006; Trang *et al.*, 2010; Bruno *et al.*, 2008, 2009). As seen from Fig. 14, the results show a clear nearly linear correlation. For the studied compounds, the surface energies



Figure 14

A double-logarithmic plot of the correlation between the surface energies calculated by the *softBV* software compared with literature DFT values. The inset shows the same data on a linear scale. The broken lines in both graphs indicate the same power law regression fit to the data.



Figure 15

Reconstructions of the NaCl (011) surface (green denotes Cl and yellow Na). See text and Table 1 for discussion of each panel.

Structure	Surface energy (J m ⁻²)
(011)- (a)	0.339
(011)-(b)	0.338
(011)-(c)	0.526
(011)- (d)	0.444
(011)-(e)	0.448

thus agree reasonably well with the values found by DFT calculations.

A particular advantage of using force-field methods in surface-energy calculations is that the reduced computational cost allows for a quick examination of richer surface arrangements. As an example, Fig. 15 shows the (011) surface of NaCl with two formula units in the top atomic layer per unit cell. Confined in this layer, the surface can have five reconstructions, labelled (a) to (e). From experience, we tend to exclude the last two because their electrical dipole is not zero. The third reconstruction can also be argued as not being 'close-packed' on the surface. All these can, however, be explicitly qualified by running our low-cost calculations, eliminating the need to narrow down the range of inspected cases by error-prone 'prior experience'.

As seen from Table 1, we find that indeed the reconstructions (011)-(d) and (011)-(e) have high energies because of their violation of the zero-dipole rule. In reconstruction (011)-(c) all ions on the surface are under-bonded. Reconstructions (011)-(a) and (011)-(b) both have reasonably low surface energies and merit further examination with more precise computational techniques.

5. Concluding remarks

Based on the well established bond-valence approach, the presented set of softBV software tools yields a convenient method for high-throughput screening of both solid electrolytes and mixed conducting insertion electrode materials, as well as their interfaces. It integrates the established crystalchemical bond-valence-sum-based plausibility control of the structure model with an automated comprehensive analysis of the migration pathways in solids with respect to pathway topology, migration barriers along the path and a localized attempt-frequency analysis. For a typical size of structure model accessible to DFT studies, the presented software tools can yield a complete overview of all relevant ion-migration pathways with a 10^3 – 10^5 times smaller computational effort and with reasonable accuracy, so that more detailed and expensive calculations can be focused on the promising candidates. This greatly reduced computational effort can be utilized to apply the diffusion-pathway analysis to more complex realistic structure models of disordered solids and their interfaces as necessary to guide the discovery of new safe high-performance (solid-state) energy-storage systems.

Acknowledgements

Support by Wee Shin Chew in developing a user interface for softBV is acknowledged. The details of the user interface will be introduced in a separate article. The authors declare no competing financial interest. The software will be made available free of charge for academic use.

Funding information

The following funding is acknowledged: National Research Foundation, Prime Minister's Office, Singapore (CRP Award NRF-CRP 10-2012-6); National University of Singapore (Centre for Energy Research seed grant).

References

- Adams, S. (2000). Habilitation thesis, Universität Göttingen, Germany.
- Adams, S. (2001). Acta Cryst. B57, 278-287.
- Adams, S. (2014). Practical Considerations in Determining Bond Valence Parameters. Bond Valences, edited by I. D. Brown & K. R. Poeppelmeier, Structure and Bonding, Vol. 158, pp. 91–128. Berlin, Heidelberg: Springer.
- Adams, S. & Rao, R. P. (2009). Phys. Chem. Chem. Phys. 11, 3210–3216.
- Adams, S. & Rao, R. P. (2011). Phys. Status Solidi A, 208, 1746– 1753.
- Adams, S. & Swenson, J. (2005). J. Phys. Condens. Matter, 17, S87.
- Anuar, N., Adnan, S. & Mohamed, N. (2014). Ceram. Int. 40, 13719– 13727.
- Bilz, H. & Kress, W. (1979). *Phonon Dispersion Relations in Insulators. Springer Series in Solid-State Sciences*, Vol. 10. Berlin, Heidelberg: Springer-Verlag.
- Bortz, A. B., Kalos, M. H. & Lebowitz, J. L. (1975). J. Comput. Phys. 17, 10–18.
- Brown, I. D. (2002). *The Chemical Bond in Inorganic Chemistry The Bond Valence Model*. New York: Oxford University Press.
- Brown, I. D. (2009). Chem. Rev. 109, 6858-6919.
- Brown, I. D. (2016). *Bond Valence Parameters*. https://www.iucr.org/ resources/data/datasets/bond-valence-parameters.
- Brown, I. D. & Altermatt, D. (1985). Acta Cryst. B41, 244-247.
- Bruno, M., Aquilano, D., Pastero, L. & Prencipe, M. (2008). Cryst. Growth Des. 8, 2163–2170.
- Bruno, M., Aquilano, D. & Prencipe, M. (2009). Cryst. Growth Des. 9, 1912–1916.
- Calderon, C. E., Plata, J. J., Toher, C., Oses, C., Levy, O., Fornari, M., Natan, A., Mehl, M. J., Hart, G., Buongiorno Nardelli, M. & Curtarolo, S. (2015). *Comput. Mater. Sci. A*, **108**, 233–238.
- Chen, M. & Adams, S. (2015). J. Solid State Electrochem. 19, 697-702.
- Chen, H. & Adams, S. (2017). IUCrJ, 4, 614-625.
- Chen, H. M., Maohua, C. & Adams, S. (2015). *Phys. Chem. Chem. Phys.* **17**, 16494–16506.
- Chew, W. S., Wong, L. L., Chen, H. & Adams, S. (2019). In preparation.

- Cussen, C. (2006). Chem. Commun. pp. 412-413.
- Deiseroth, H.-J., Kong, S.-T., Eckert, H., Vannahme, J., Reiner, C., Zaiß, T. & Schlosser, M. (2008). *Angew. Chem. Int. Ed.* **47**, 755– 758.
- Dugas, C., Bengio, Y., Bélisle, F. C., Nadeau & Garcia, R. (2000). Adv. Neural Inf. Process. Syst. 13, 472–478.
- Eglitis, R. I., Borstel, G., Heifets, E., Piskunov, S. & Kotomin, E. (2006). J. Electroceram. 16, 289–292.
- Filsø, M. Ø., Turner, M. J., Gibbs, G. V., Adams, S., Spackman, M. A. & Iversen, B. B. (2013). *Chem. Eur. J.* **19**, 15535–15544.
- Frigo, M. & Johnson, S. G. (2005). Proc. IEEE, 93, 216-231.
- Frisch, M. J. et al. (2016). GAUSSIAN16. Revision B.01. Gaussian Inc., Wallingford, Connecticut, USA.
- Gagné, O. C. & Hawthorne, F. C. (2015). Acta Cryst. B71, 562-578.
- Gillespie, D. T. (1977). J. Phys. Chem. 81, 2340-2361.
- Hall, S. R., Allen, F. H. & Brown, I. D. (2005). Editors. *International Tables for Crystallography*, Vol. G, ch. 4.1, pp. 210–257. Chester: International Union of Crystallography.
- Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G. & Persson, K. A. (2013). APL Mater. 1, 011002.
- Johnson, D. B. (1975). SIAM J. Comput. 4, 77-84.
- Kong, S.-T., Deiseroth, H.-J., Maier, J., Nickel, V., Weichert, K. & Reiner, C. (2010). Z. Anorg. Allg. Chem. 636, 1920–1924.
- Kramer, D. & Ceder, G. (2009). Chem. Mater. 21, 3799-3809.
- Krivovichev, S. V. & Brown, I. D. (2001). Z. Kristallogr. 216, 245-247.
- Leeuwen, J. & v (1998). Handbook of Theoretical Computer Science. Vol. A, Algorithms and Complexity. Amsterdam: Elsevier.
- Locock, A. J. & Burns, P. C. (2004). Z. Kristallogr. 219, 259-266.
- Meyer, W. & Neldel, H. (1937). Z. Tech. Phys. 12, 588.
- Mills, G. & Jónsson, H. (1994). Phys. Rev. Lett. 72, 1124-1127.
- Momma, K. & Izumi, F. (2011). J. Appl. Cryst. 44, 1272-1276.
- Nakayama, M., Hotta, S., Nakamura, T. & Kasuga, T. (2013). J. Ceram. Soc. Jpn, **121**, 611–613.
- Nishitani, Y., Adams, S., Ichikawa, K. & Tsujita, T. (2018). Solid State Ionics, 315, 111–115.
- NIST (2018). NIST Atomic Spectra Database Ionization Energies Form. http://physics.nist.gov/PhysRefData/ASD/ionEnergy.html.
- Parr, R. G. & Pearson, R. G. (1983). J. Am. Chem. Soc. 105, 7512– 7516.
- Rayavarapu, P. R., Sharma, N., Peterson, V. & Adams, S. (2012). J. Solid State Electrochem. 16, 1807–1813.
- Saal, J. E., Kirklin, S., Aykol, M., Meredig, B. & Wolverton, C. (2013). JOM, 65, 1501–1509.
- Salinas-Sanchez, A., Garcia-Muñoz, J. L., Rodriguez-Carvajal, J., Saez-Puche, R. & Martinez, J. L. (1992). J. Solid State Chem. 100, 201–211.
- Sun, W. & Ceder, G. (2013). Surf. Sci. 617, 53-59.
- Tasker, P. W. (1979). J. Phys. C. Solid State Phys. 12, 4977-4984.
- Thangadurai, V., Adams, S. & Weppner, W. (2004). *Chem. Mater.* 16, 2998–3006.
- Trang, N. T., Chinh, N. V., Cong, B. T. & Linh, N. H. (2010). Proc. Natl Conf. Theor. Phys. 35, 209–220.
- Wong, L. L., Chen, H. & Adams, S. (2019). In preparation.
- Zirl, D. M. & Garofalini, S. H. (1990). J. Am. Ceram. Soc. 73, 2848– 2856.