Crystal structure optimisation using an auxiliary equation of state

Adam J. Jackson,[1] Jonathan M. Skelton,[1] Christopher H. Hendon,[1] Keith T. Butler,[1] and Aron Walsh[1,2,a)]
[1)] *Centre for Sustainable Chemical Technologies and Department of Chemistry, University of Bath, Claverton Down, Bath, BA2 7AY, United Kingdom*
[2)] *Global E^3 Institute and Department of Materials Science and Engineering, Yonsei University, Seoul 120-749, Korea*

Standard procedures for local crystal-structure optimisation involve numerous energy and force calculations. It is common to calculate an energy–volume curve, fitting an equation of state around the equilibrium cell volume. This is a computationally intensive process, in particular for low-symmetry crystal structures where each isochoric optimisation involves energy minimisation over many degrees of freedom. Such procedures can be prohibitive for non-local exchange-correlation functionals or other 'beyond' density functional theory electronic structure techniques, particularly where analytical gradients are not available. We present a simple approach for efficient optimisation of crystal structures based on a known equation of state. The equilibrium volume can be predicted from one single-point calculation, and refined with successive calculations if required. The approach is validated for PbS, PbTe, ZnS and ZnTe using nine density functionals, and applied to the quaternary semiconductor Cu_2ZnSnS_4 and the magnetic metal-organic framework HKUST-1.

I. INTRODUCTION

The standard operating procedure for computational investigations in solid-state chemistry is to begin with a crystal structure – obtained either from diffraction studies or through chemical analogy – and to optimise the lattice shape, volume and internal positions to minimise all forces. It is from this equilibrium crystal structure (athermal for the majority of electronic-structure approaches) that the full range of material response functions (e.g. elastic, dielectric, magnetic) are calculated.[1]

The optimisation of a crystal structure may require hundreds of self-consistent field iterations across a series of ionic configurations.[2] The most robust approach to optimisation is the calculation of an equation of state (EoS) for the material, relating the unit cell dimensions to energy and pressure.[3] This is based on a series of calculations at differing volumes, where ideally the shape and internal positions are optimised at each point. The simplest case is a cubic lattice with high internal symmetry, e.g. rocksalt, where the only degree of freedom is the volume, and computing the EoS reduces to a series of single-point calculations. For a triclinic cell, the lengths, angles and internal positions in principle all require optimisation. While it is sometimes possible to directly optimise the cell volume by *simultaneously* minimising the stress tensor of the unit cell, this approach can run into artifacts when using plane-wave basis sets (i.e. Pulay forces) unless an iterative procedure is employed.[4]

It has become commonplace to use an 'equilibrium' crystal geometry, determined using one exchange-correlation functional within density functional theory, for a 'single-shot' higher-level calculation performed to give a more accurate electronic structure. This methodology has been applied to the calculation of properties as diverse as workfunctions, electronic bandgaps, optical properties and defect formation energies.[5–12] The implicit assumption is that the qualitative behaviour is insensitive to small differences in the local structure. The approximation will fail where the electronic structure (chemical bonding) of a system is poorly described at the initial level of theory, e.g. the treatment of Mott insulators such as NiO within the local-density approximation (LDA).[13]

In this contribution we outline a simple procedure for the rapid volume optimisation (RVO) of crystal structures. It takes advantage of the similarity in the pressure-volume relationship for a given material between different theoretical approaches, here being exchange-correlation (XC) functionals within density functional theory. Where an EoS is known for one functional, the equilibrium volume for another functional can be predicted with reasonable accuracy using a single calculation, and further refined following an iterative procedure. The approach has particular utility for studies assessing material properties using a range of electronic-structure methods, and for studies employing methods with high computational cost (e.g. hybrid, meta-hybrid and double-hybrid treatments of electron exchange and correlation). We validate the approach for four Zn and Pb chalcogenides, and demonstrate its utility in describing the electronic and magnetic structure of one complex semiconductor (Cu_2ZnSnS_4) and one metal-organic framework (HKUST-1), respectively.

II. OUTLINE OF PROCEDURE

The goal of local crystal-structure optimisation is to minimise all degrees of freedom (cell size, shape and positions) with respect to the total energy of the system. It is convenient to employ an EoS based on an energy-volume (E-V) curve, where the remaining degrees of freedom (i.e. shape and positions) are minimised for each volume using standard numerical minimisation approaches (e.g. the conjugate-gradient method). Kohn-Sham density functional theory (DFT)[14] is one of the most widely

[a)] a.walsh@bath.ac.uk

used electronic structure techniques for modelling solid-state materials. Most DFT codes provide optimisation algorithms for this purpose, e.g. the `ISIF=4` setting in the Vienna Ab initio Simulations Package (VASP)[15], the `cell_dofree='volume'` setting in Quantum-Espresso[16] or the `CVOLOPT` setting in CRYSTAL.[17]

A superficial resemblance is clear between E-V curves obtained with different exchange-correlation functionals, with similar shapes but different minima (Figure 1a). The extent of the similarity becomes apparent when using pressure-volume (P-V) curves (Figure 1b), where "pressure" refers to the scalar hydrostatic pressure on the periodic system. As this pressure $P = -\frac{\mathrm{d}E}{\mathrm{d}V}$, the optimal geometries $\frac{\mathrm{d}E}{\mathrm{d}V} = 0$ are now those intersecting the $P = 0$ line. We note that while these still differ depending on the chosen XC functional, the P-V curves have similar curvature, with the same approximate slopes about their zero-crossing points. From these we make our key assumption: *the slope of one P-V curve may be used to estimate the crossing point of another.*

For certain beyond-DFT calculation methods, the stress tensor is not computed directly. However, where the energy is available the hydrostatic pressure may always be estimated with a finite difference:

$$P(V) \approx -\frac{E(V+\delta) - E(V)}{(V+\delta) - V} \tag{1}$$

The procedure, outlined in Figure 2, is:

1. Form a P-V curve using one description of the interatomic interactions (method A). This can be achieved by fitting an EoS to an energy-volume curve. If a system-specific set of classical potentials is available, this would be expected to perform very well as they are often fitted to the experimental lattice parameters and elastic properties. Within DFT, descriptions of electron exchange and correlation within the generalised-gradient approximation (GGA) are suitable,[18] given their low cost and the availability of analytical gradients for the rapid calculation of forces. Comparative studies suggest that PBEsol[19] gives especially good estimates for the lattice parameters and elastic properties of crystals.[20,21]

2. Calculate the slope about $P = 0$ for method A. This will form our linear approximation.

$$m = \frac{\mathrm{d}P_A}{\mathrm{d}V}\Big|_{P_A=0} \tag{2}$$

3. Perform a calculation using a second approach (method B), e.g. hybrid DFT with the screened HSE06 functional[22], using an estimated initial volume; this may be the equilibrium volume (V_0) for method A. Convert the resulting stress tensor to a hydrostatic pressure P_0. (If no analytical stress tensor is available, use a finite difference as in Eqn. 1.)

TABLE I: Equilibrium properties of PbS from the Murnaghan EoS, fitting over a range of functionals: lattice parameter a; unit cell volume V_0; volume difference ϵ_V from experimental value; Murnaghan EoS parameters k_0 and k_0'. k_0 is equivalent to Bulk modulus at zero pressure. The experimental lattice constant was obtained from low-temperature neutron powder diffraction data fitted and extrapolated to zero temperature by K. S. Knight (2014).[23]

XC functional	a / Å	V_0 / Å^3	ϵ_V / %	k_0	k_0'
LDA	5.84	199.01	−3.47	65.71	4.42
PW91	5.99	215.21	4.38	55.26	3.98
PBE	5.98	214.18	3.88	54.64	4.00
PBEsol	5.88	203.43	−1.33	61.13	4.25
TPSS	5.96	211.76	2.71	57.33	4.01
revTPSS	5.94	209.05	1.39	57.85	4.00
PBE+D2	5.84	199.19	−3.39	59.93	5.02
B3LYP	6.06	223.02	8.17	53.20	4.07
HSE06	5.96	210.09	1.90	59.29	4.32
Experiment[23]	5.91	206.17			

4. Estimate the corrected volume for method B:

$$V_1 = V_0 + \frac{P_0}{m} \tag{3}$$

5. Generate a unit cell with volume V_1 (e.g. by interpolating between the previous calculations with method A), and recalculate the desired properties with method B.

6. Iterate steps 4-5 until P is acceptably low:

$$V_{n+1} = V_n + \frac{P_n}{m}, \qquad P_n = f(V_n) \tag{4}$$

III. ERROR ESTIMATION

A. Accuracy of linear approximation

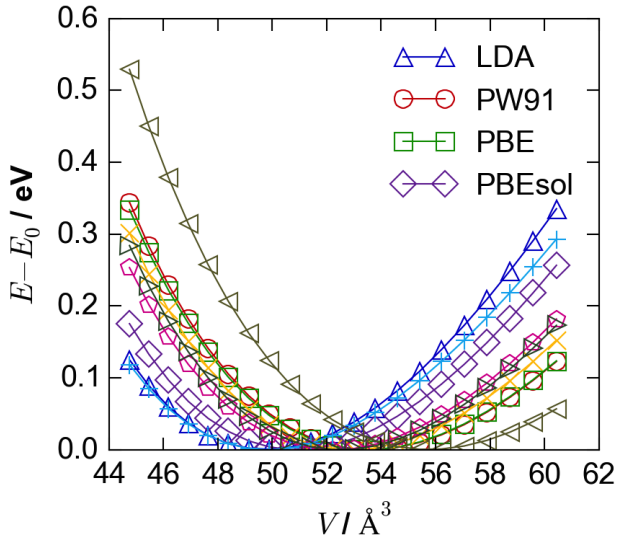
In this approach, a linear fit is used for the pressure-volume relationship:

$$P_{\mathrm{est}} = aV + b \tag{5}$$

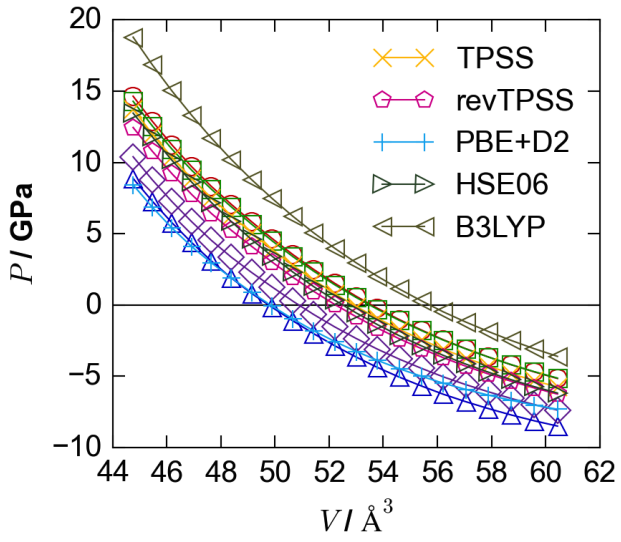
$$\frac{\mathrm{d}P_{\mathrm{est}}}{\mathrm{d}V} = a. \tag{6}$$

This is by no means a conventional equation of state, but may provide a useful approximation when close to the minimum volume. The standard definition of the bulk modulus,

$$B = -V\frac{\mathrm{d}P}{\mathrm{d}V}, \tag{7}$$



(a)



(b)

FIG. 1: Energy-volume and pressure-volume curves computed for PbS using a variety of DFT exchange-correlation functionals. Markers indicate calculated values, while smooth lines are fits to the Murnaghan equation of state.

yields the static bulk modulus B_0 when evaluated at the equilibrium volume V_0.

$$B_0 = -V_0 \cdot \frac{\mathrm{d}P}{\mathrm{d}V}\Big|_{V=V_0} \tag{8}$$

As we have assumed this derivative to be constant, we combine with Eqn. (6) to give a physically meaningful expression of our assumption:

$$\frac{\mathrm{d}P_{\mathrm{est}}}{\mathrm{d}V} = a = -\frac{B_0}{V_0}. \tag{9}$$

It is now straightforward to compare this approximate EoS with a more conventional form for solid materials, estimating an associated error (ϵ). The simplest case is a system with constant bulk modulus.

$$\frac{\mathrm{d}P}{\mathrm{d}V} = -\frac{B_0}{V} \tag{10}$$

$$\epsilon = P - P_{\mathrm{est}} \tag{11}$$

$$\frac{\mathrm{d}\epsilon}{\mathrm{d}V} = \frac{\mathrm{d}P}{\mathrm{d}V} - \frac{\mathrm{d}P_{\mathrm{est}}}{\mathrm{d}V} = \frac{-B_0}{V} - \frac{-B_0}{V_0} = B_0\left(\frac{1}{V_0} - \frac{1}{V}\right) \tag{12}$$

$$\epsilon(V) = \int_{V_0}^{V} \frac{\mathrm{d}\epsilon}{\mathrm{d}V}\mathrm{d}V = B_0\left[\frac{V}{V_0} - \ln V\right]_{V_0}^{V} \tag{13}$$

$$\epsilon(V) = B_0\left(\frac{V}{V_0} - 1 + \ln\left(\frac{V_0}{V}\right)\right) = B_0\left(\frac{V - V_0}{V_0} - \ln\left(\frac{V}{V_0}\right)\right) \tag{14}$$

At a typical volume deviation of 5% (Table I):

$$\epsilon(1.05V_0) = B_0\left(\frac{1.05V_0 - V_0}{V_0} - \ln\left(\frac{1.05V_0}{V_0}\right)\right) \tag{15}$$

$$= B_0(0.05 - \ln(1.05)) \tag{16}$$

where the pressure

$$P = \int_{V_0}^{1.05V_0} -\frac{B_0}{V}\mathrm{d}V = -B_0 \ln\left(\frac{1.05V_0}{V_0}\right) \tag{17}$$

and hence the fractional error $\frac{\epsilon}{P}$ is -2.5%.

Moving to an improved, while still relatively simple, EoS, the Murnaghan EoS adds a parameter, effectively giving a linear volume dependence to B_0.[24] Taking its derivative form (Eqn. 18), we improve our error estimate:

$$\frac{\mathrm{d}P}{\mathrm{d}V} = -\frac{k_0}{V}\left(\frac{V_0}{V}\right)^{k_0'} \tag{18}$$

$$\frac{\mathrm{d}\epsilon}{\mathrm{d}V} = \frac{\mathrm{d}P}{\mathrm{d}V} - \frac{\mathrm{d}P_{\mathrm{est}}}{\mathrm{d}V} = \frac{k_0}{V}\left(\frac{V_0}{V}\right)^{k_0'} - \frac{-B_0}{V_0} \tag{19}$$

We can relate B_0 to the Murnaghan parameters k_0, k_0' by finding the slope at V_0:

$$B_0 = -V_0 \cdot \frac{\mathrm{d}P}{\mathrm{d}V}\Big|_{V=V_0} = \frac{V_0}{V_0}k_0\left(\frac{V_0}{V_0}\right)^{k_0'} = k_0 \tag{20}$$

$$\frac{\mathrm{d}\epsilon}{\mathrm{d}V} = k_0\left(\frac{1}{V_0} - \frac{V_0^{k_0'}}{V^{k_0'+1}}\right) \tag{21}$$

$$\epsilon(V) = \int_{V_0}^{V} k_0\left(\frac{1}{V_0} - \frac{V_0^{k_0'}}{V^{k_0'+1}}\right)\mathrm{d}V \tag{22}$$

$$= k_0\left[\frac{V}{V_0} + \frac{1}{k_0'}\left(\frac{V_0}{V}\right)^{k_0'}\right]_{V_0}^{V} \tag{23}$$

$$\tag{24}$$

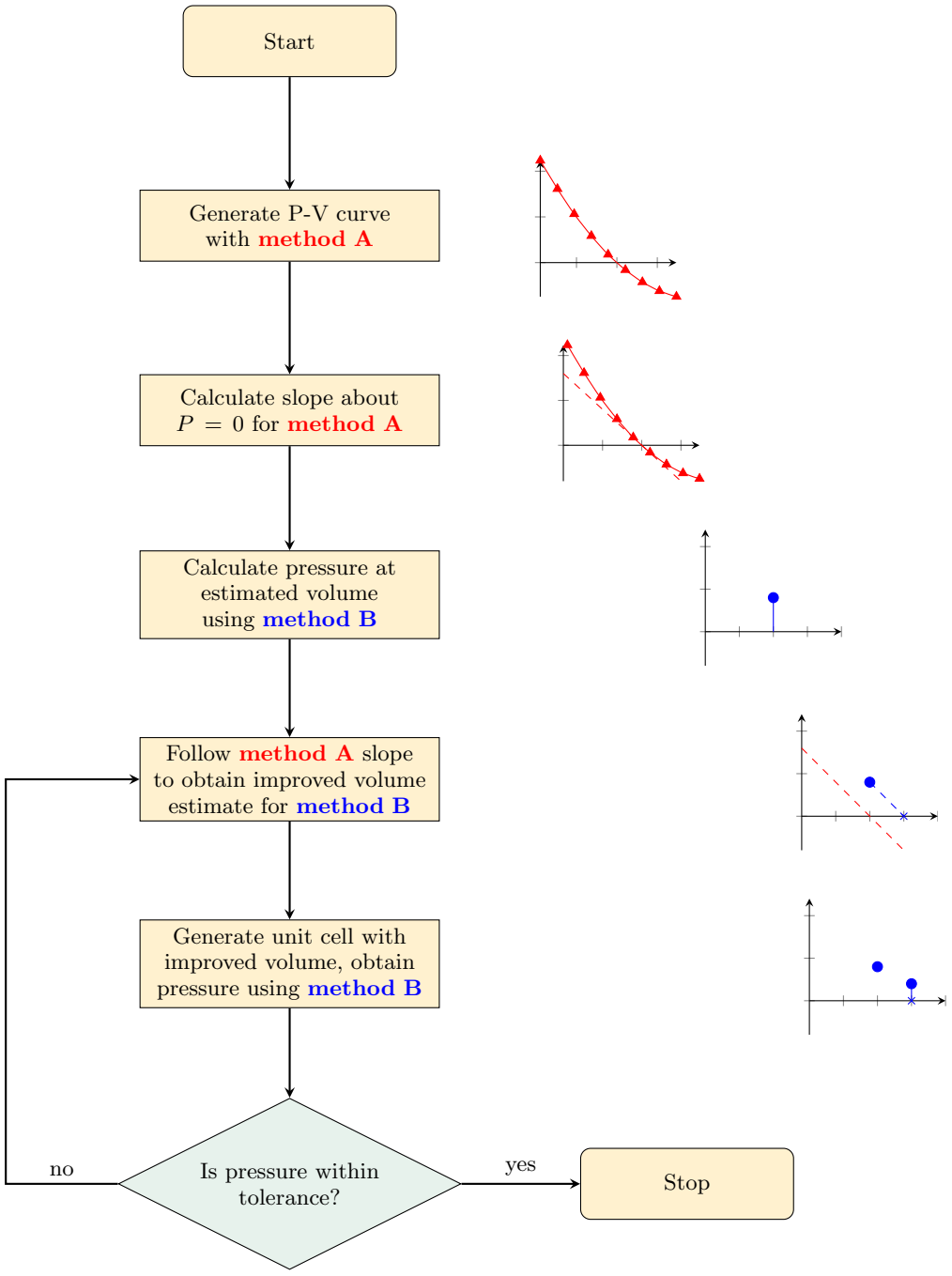


FIG. 2: Flow chart for the rapid volume optimisation procedure.

$$\epsilon(V) = k_0 \left(\frac{V}{V_0} + \frac{1}{k_0'} \left(\frac{V}{V_0} \right)^{-k_0'} - 1 - \frac{1}{k_0'} \right) \tag{25}$$

$$= k_0 \left(\frac{V}{V_0} - 1 \right) + \frac{k_0}{k_0'} \left(\left(\frac{V}{V_0} \right)^{-k_0'} - 1 \right) \tag{26}$$

Plotting these error estimates in Figure 3, we find that the error in pressure is less than 10% of the static bulk modulus for volumes 10% from the optimum, given a material that obeys the Murnaghan EoS with a typical value $k_0' = 5$. For smaller values of k_0' (i.e. closer to the fixed-bulk-modulus model), the errors are greatly reduced. In any case, the linear approximation appears to be sufficiently accurate for a stable optimisation process.

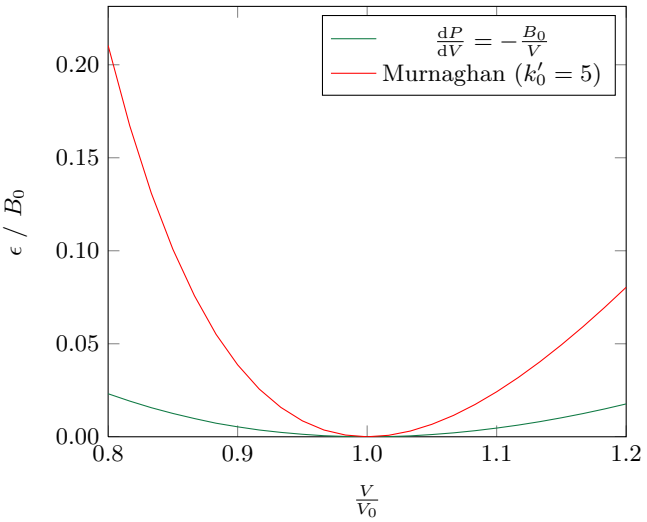


FIG. 3: Divergence of the linear approximation from more complex equations of state. The error $\epsilon = P_{\text{est}} - P_{\text{EOS}}$ is given in units of the static bulk modulus.

B. Dependence on accuracy of EoS

Returning to the simplistic EoS of Eqn. 10

$$P = -\int_{V_0}^{V} \frac{B_0}{V} \mathrm{d}V \tag{27}$$

$$= B_0 \ln\left(\frac{V_0}{V}\right), \tag{28}$$

we examine the residual pressure P_1 following a single step of RVO from an initial volume V_i,

$$V_1 = V_i + \frac{P_B(V_i)}{m_A} = V_i - P_B(V_i)\frac{V_{0,A}}{B_{0,A}} \tag{29}$$

$$= V_i - B_{0,B} \ln\left(\frac{V_{0,B}}{V_i}\right) \cdot \frac{V_{0,A}}{B_{0,A}} \tag{30}$$

$$P_1 = B_{0,B} \ln\left(\frac{V_{0,B}}{V_1}\right) \tag{31}$$

$$\frac{P_1}{B_{0,B}} = -\ln\left[\frac{V_i}{V_{0,B}} - \frac{B_{0,B}V_{0,A}}{B_{0,A}V_{0,B}} \ln\left(\frac{V_{0,B}}{V_i}\right)\right]. \tag{32}$$

We note that $\ln(x) \approx x - 1$ for x close to 1, and hence the residual pressure is approximately linear with respect to the error in initial volume estimate. The term $\frac{B_{0,B}V_{0,A}}{B_{0,A}V_{0,B}}$ indicates a smaller linear dependence on the similarity of the EoS for method A and method B. Moving to the Murnaghan EoS,

$$P_1 = \frac{k_{0,B}}{k'_{0,B}} \left(\left(\frac{V_{0,B}}{V_1}\right)^{k'_{0,B}} - 1\right) \tag{33}$$

$$= \frac{k_{0,B}}{k'_{0,B}} \left[\left(\frac{V_{0,B}}{V_i + P_B(V_i)\frac{V_{0,A}}{-k_{0,A}}}\right)^{k'_{0,B}} - 1\right] \tag{34}$$

$$\frac{P_1}{\frac{k_{0,B}}{k'_{0,B}}} = \left(\frac{V_{0,B}}{V_i - \frac{V_{0,A}}{k_{0,A}} \cdot \frac{k_{0,B}}{k'_{0,B}} \left(\left(\frac{V_{0,B}}{V_i}\right)^{k'_{0,B}} - 1\right)}\right)^{k'_{0,B}} - 1 \tag{35}$$

$$\frac{P_1}{\frac{k_{0,B}}{k'_{0,B}}} = \left(\frac{V_i}{V_{0,B}} - \frac{V_{0,A}k_{0,B}}{V_{0,B}k_{0,A}} \cdot \frac{1}{k'_{B,0}} \left(\left(\frac{V_{0,B}}{V_i}\right)^{k'_{0,B}} - 1\right)\right)^{-k'_{0,B}} - 1 \tag{36}$$

Again, the pressure is dominated by the initial position, with a smaller contribution from the difference between EoS "stiffness" and volume minima. The non-linearity of this relationship follows the non-linearity of the true EoS through the exponent $k'_{0,B}$. We conclude therefore that the performance of the first step is equally sensitive to percentage differences in equilibrium volume and bulk modulus between method A and method B. Convergence is impacted by the non-linearity of the EoS, but not by how accurately this non-linearity is reproduced by method A.

IV. METHODS

A. Electronic Structure Calculations

Studies have been carried out on the binary chalcogenides PbS, PbTe, ZnS, and ZnTe as well as the quaternary semiconductor Cu_2ZnSnS_4 and an organic-inorganic hybrid material HKUST-1.

All DFT calculations on the binary chalcogenides were carried out with VASP[25] using the two-atom primitive face-centred cubic unit cells. We employed projector-augmented wave (PAW) frozen-core potentials[26,27] treating the outermost s and p electrons of S, Te, and Pb and the outermost s, p, and d electrons of Zn explicitly as valence. For consistency, we used the LDA PAW potential set. The PAW potentials are highly transferable and tests showed a very weak dependence of the resulting optimised electronic and crystal structures. A plane-wave kinetic-energy cutoff of 550 eV was employed in all these calculations, and the Brillouin zone was sampled with an

$8\times8\times8$ Γ-centered Monkhorst-Pack mesh.[28] During electronic minimisation, the wavefunctions were optimised to an energy tolerance of 10^{-8} eV. These parameters were found to be sufficient to converge the absolute total energies to within 1 meV atom^{-1}, and the stress tensors to well within 1 kbar (0.1 GPa).

The simplicity of the binary systems allowed us to test a wide range of functionals, spanning different "levels" of approximations to the exchange-correlation potential.[29] As a baseline, we took the local-density approximation (LDA) with the Perdew-Zunger parameterisation of the correlation energy.[30] Calculations using the generalised-gradient (GGA) approximation were performed with the Perdew-Wang 91 (PW91)[31–33] and Perdew-Burke-Enzerhof (PBE)[34] functionals, plus the variant of PBE revised for solids (PBEsol).[19] To complement this set of functionals, we also tested the Grimme D2 dispersion correction to PBE.[35] "Meta-GGA" calculations were carried out using the Tao-Perdew-Staroverov-Scuseria (TPSS) functional[36] and the subsequent revision of Perdew et al. (revTPSS).[37] Finally, we tested two hybrids, *viz.* the popular HSE06[22] and B3LYP[38] functionals. For each material and functional, we calculated an E-V curve by adjusting the lattice parameter to yield 21 volumes about the experimental 300 K lattice parameters reported in Refs. 39 and 40 covering a range of approx. ±5%. We note that, as a result of the high symmetry of these systems, the lattice parameter is the only degree of freedom, and thus relaxation of the cell shape and internal positions was not required.

For Cu_2ZnSnS_4 (Section V B), energy-volume curves were formed from all-electron DFT calculations using the FHI-aims code.[41,42] These calculations employed numerically-tabulated atom-centered basis functions (the 'tight' defaults were used, which correspond to expected convergence of < 10 meV per atom) and evenly-spaced **k**-point grids. Additional hybrid (HSE06) DFT calculations and primitive-cell optimisations used VASP with the PAW-PBE potential set and a 500 eV cutoff for the plane wave basis set. All calculations on Cu_2ZnSnS_4 sampled the Brillouin zone with $6\times6\times6$ Γ-centered **k**-point grids.

For the Cu-based metal-organic framework HKUST-1, calculations were again performed with the VASP code, considering only the point Γ in reciprocal space due to the large size of unit cell. Owing to the presence of the open-shell Cu(II) ion (d^9 configuration) all calculations were spin-polarised, and a range of magnetic structures were tested as discussed in Section V C. The PBEsol and HSEsol functionals were used along with the PAW-PBE potential set. Here 'HSEsol' refers to a modification of HSE06, with PBEsol replacing PBE as the local exchange-correlation functional.[43] Due to the complexity of the crystal structure, only three energy-volume points were included in the EoS and a single iteration of RVO was performed to recover the ground-state HSEsol structure. A slightly different procedure was followed in this case: a quadratic E-V curve was fitted to the three PBEsol points. The initial HSEsol calculation was carried out at the fully-optimised PBEsol point, and the E-V curve was followed assuming a constant pressure offset to estimate the equilibrium volume for HSEsol. (This application was the first chronologically, and led to the development of the Murnaghan fitting procedure.)

B. Implementation

The RVO method was implemented and tested with code written in Python 2.7.3, using the standard library and Numpy/Scipy/Matplotlib.[44–46] (The code is freely available; details in Supporting Information.) Non-linear fitting to the Murnaghan EoS uses the `curve_fit` routine in the Scipy Python library, which accesses Minpack, an open-source Fortran library.[45] This implements least-squares fitting with the Levenberg-Marquardt algorithm.[47] Initial guesses of 50.0 eV Å^{-3} and 5.0 were used for the k' and k'_0 parameters, respectively.

V. RESULTS

A. II-VI Binary Chalcogenide Semiconductors

1. Simulated application across a range of methods

For PbS there is a significant range in equilibrium lattice parameters for different exchange-correlation functionals, corresponding to a maximum volume difference of over 10%, between the LDA and B3LYP calculations (Figure 1). Values are tabulated in Table I, and compared to a recent low-temperature study by K. S. Knight.[23] The computed values are similar, but slightly different, to the computational results reported by Hummer et al.[48] Direct bandgaps were also calculated at each volume expansion, at the special **k**-point X for the lead compounds and at Γ for the zinc compounds. (Note that PbS and PbTe have another, smaller direct bandgap at the L point.) It can be seen in Figure 4 that over the lattice-parameter expansion and contraction of up to 5%, the bandgaps vary by around 1 eV, with the direction of change depending on the structure type and chemistry. In this case, using an LDA-predicted geometry for a 'single-shot' B3LYP calculation would lead to a difference in bandgap of ~ 0.4 eV compared to that at the equilibrium geometry for B3LYP.

An iterative application of the RVO procedure was then simulated from the data. The Murnaghan EoS (Eqn. 37) in its integrated form (Eqn. 38) was fitted to each E-V curve from DFT calculations. This allowed energy and pressure to be calculated for arbitrary volumes

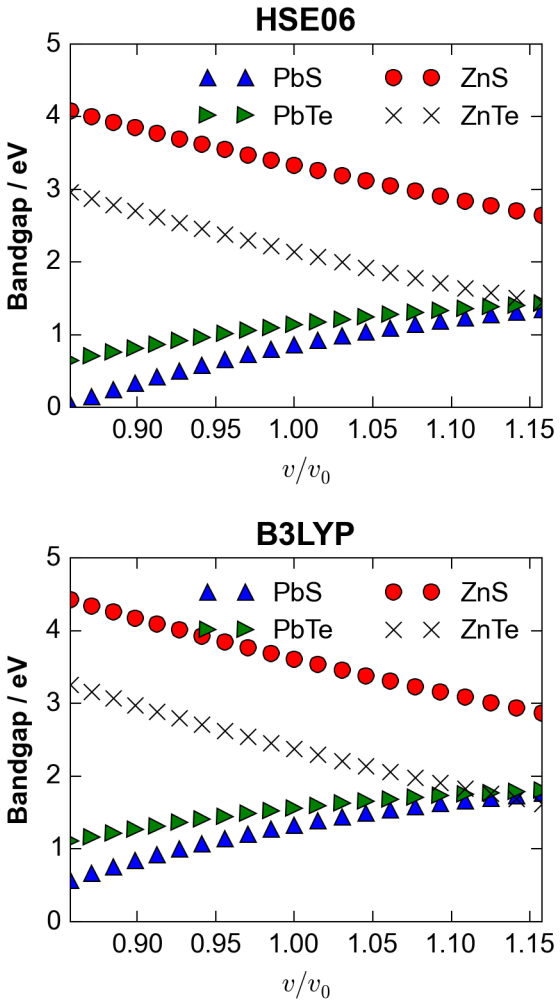


FIG. 4: Volume-dependence of calculated direct bandgaps at Γ (ZnS, ZnTe) and X (PbS, PbTe) with the HSE06 and B3LYP hybrid DFT functionals. Results as a function of volume (temperature) for PbTe have previously been reported in Ref. 49. The behaviour is characteristic of a positive and negative bandgap deformation potential for the Pb and Zn semiconductors, respectively. The volume expansion range of 0.86–1.16 corresponds to lattice parameter expansions and contractions of 5% in each dimension. Markers indicate the calculated values.

without carrying out additional DFT calculations.

$$P = \frac{k_0}{k_0'}\left(\left(\frac{V_0}{V}\right)^{k_0'} - 1\right) \tag{37}$$

$$E = E_0 + k_0 V_0\left(\left(\frac{V}{V_0}\right)^{1-k_0'} \cdot \frac{1}{k_0'(k_0'-1)} + \frac{V}{k_0' V_0} - \frac{1}{k_0'-1}\right) \tag{38}$$

The quality of these fits was sufficient for this exercise, with RMS fitting errors of < 1 meV. Fitting parameters and full data are included in ESI. For each "test functional"-"reference functional" pair, the minimum volume (corresponding to the fitting parameter V_0) of the reference functional was taken as the initial volume guess, and an external pressure calculation modelled by evaluating the pressure at this volume using the EoS for the test functional. This refined pressure and volume was then used as the basis for further iterations. The external pressure over successive iterations is shown for PbS in Figure 5 for each combination of functionals; convergence is rapid with the residual pressure dropping almost logarithmically with subsequent steps, typically by a factor of $\sim 10^3$ in three steps.

2. Comparison with a standard optimisation procedure

In general terms, a direct optimisation with method B will take $N_{\mathrm{opt},B}$ steps, each requiring an average computing time t_B, to converge to the equilibrium volume. Constructing an EoS for RVO using method A requires $N_{\mathrm{EoS},A}$ optimisations, which, as for method B, take $N_{\mathrm{opt},A}$ steps of time t_A. We note that in most cases $N_{\mathrm{opt},B}$ will be larger than $N_{\mathrm{opt},A}$, since the direct optimisation with method B must adjust the internal coordinates, cell shape and volume, while the optimisation with method A needs only to optimise the internal coordinates and the shape. Subsequent application of N_{RVO} iterations of the algorithm then requires $1+N_{\mathrm{RVO}}$ single-point calculations using method B, each again requiring t_B time. RVO is expected to be more efficient than a direct optimisation with method B if the following inequality holds:

$$N_{\mathrm{EoS},A} N_{\mathrm{opt},A} t_A + (1 + N_{\mathrm{RVO}})\, t_B < N_{\mathrm{opt},B} t_B. \tag{39}$$

The cubic systems considered in this section, for which the cell volume is the only degree of freedom, represent a special case where $N_{\mathrm{opt},A} = 1$. We assume that energy gradients are available with method B, and that the optimisation algorithm would converge in three steps, i.e. $N_{\mathrm{opt},B} = 3$. This is reasonable if a good estimate of the starting volume is available, such as a room-temperature lattice constant. The inequality simplifies to

$$N_{\mathrm{EoS},A} t_A + (1 + N_{\mathrm{RVO}})\, t_B < 3 t_B; \tag{40}$$

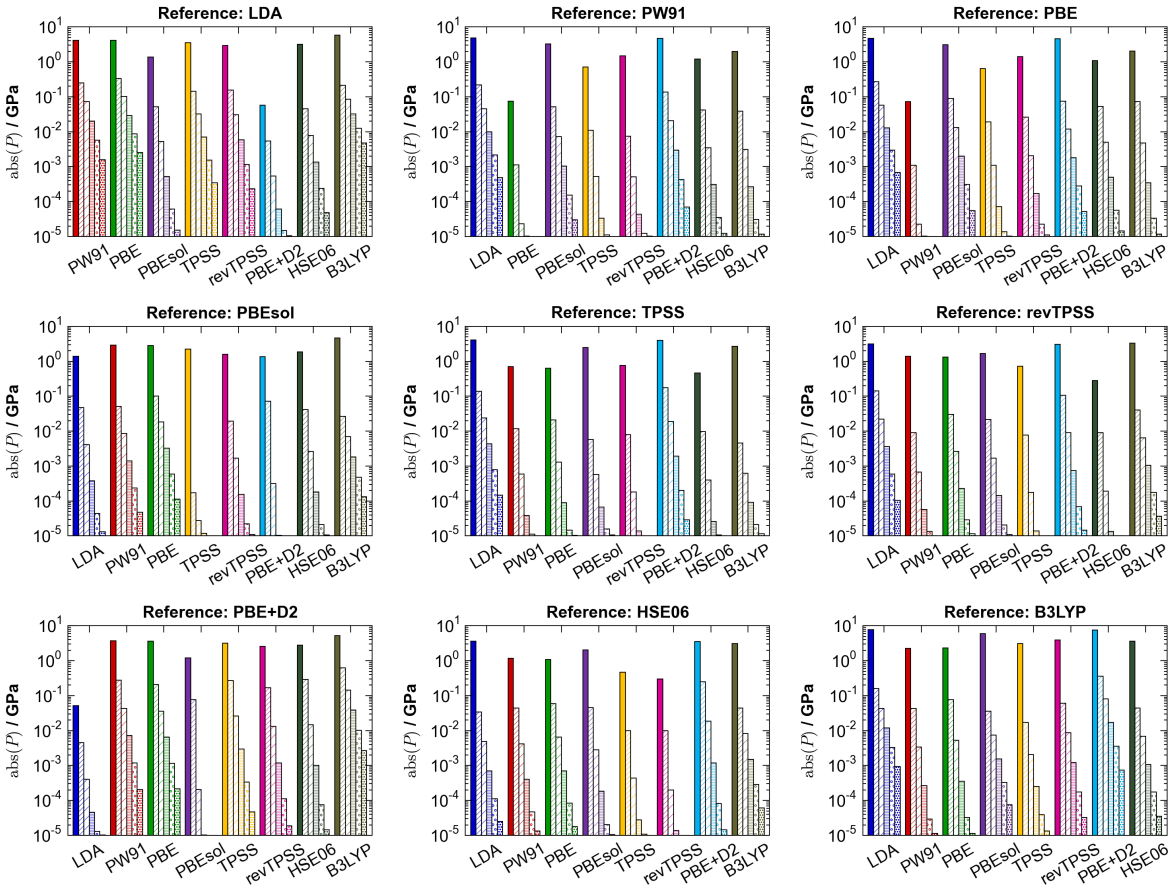


FIG. 5: Residual pressures following six iterations of volume optimisation of PbS using different "reference"/"test" combinations of exchange-correlation functionals. The "reference" refers to the functional used for the EoS ("method A"), while the groups of bars correspond to the functionals used in simulated single-point calculations ("method B"). Note that the pressure is presented on a logarithmic scale.

it can be seen that RVO will outperform a direct optimisation if a suitable pressure is obtained after one iteration while $t_A < \frac{t_B}{N_{\mathrm{EoS},A}}$.

As a concrete example, we compared a direct optimisation of PbS with HSE06 to an optimisation with RVO using PBEsol and HSE06 as method A and method B, respectively. The initial structure for both optimisations was the experimentally-measured room temperature volume, and an eleven-point EoS for RVO was computed about this value using PBEsol. The direct optimisation used a quasi-Newton algorithm as implemented by VASP (with the input tag "IBRION=1"). Both sets of calculations were performed on a dual-CPU Intel Xeon workstation with 12 physical cores and 64 Gb RAM, allowing the timings to be compared directly. The comparison is given in Table II.

In this test, a single-point calculation with HSE06 was on average 150 times more expensive than a calculation with PBEsol; this is both due to the higher complexity of non-local hybrid functionals compared to semi-local GGA methods, and to the different scaling properties with the number of **k**-points used to sample the Brillouin zone. With the force convergence criterion of 10^{-2} eV $\mathring{A}^{-1}$ for direct optimisation, the pressure was reduced to -0.1 kbar from an initial pressure of 3.42 kbar in three steps, taking 4.75 hrs on our test hardware. A single iteration of RVO yielded a pressure of 0.15 kbar in 4.18 hrs, while a second iteration yielded 0.03 kbar in a total time of 6 hrs.

It can be seen from the data in Table II that the direct optimisation takes on average less time per force calculation than RVO; the procedure implemented in VASP re-uses calculated wavefunctions to speed up the convergence of the second and third steps. In this case, we found that one of the conjugate-gradient electronic-minimisation cycles during the first single-point calculation for the RVO algorithm took some 1500s longer than both the other steps in this series and the first step of the quasi-Newton volume optimisation, despite the latter being notionally an identical calculation. This artefact contributes significantly to the cost of the single-iteration RVO calculation.

TABLE II: Comparison of a direct HSE06 volume optimisation and one and two iterations of the RVO algorithm in determining the equilibrium volume of PbS. For each step, the total time for each step is recorded alongside the cell volume and pressure after the cycle where appropriate. For the direct optimisation, the timings of the three steps are printed alongside the total for the complete calculation, so the latter includes additional operations such as setup time and is slightly longer than the sum of the three electronic minimisations.

Algorithm	Step	t / s	V / Å^3	p / kbar
Direct (HSE06)	1	6669	52.21	3.42
	2	5808	52.63	−1.45
	3	4509	52.50	−0.10
	Total	17038		
RVO	PBEsol EoS	47		
	HSE06 1	8234	52.21	3.39
	HSE06 2	6754	52.49	0.15
	Total (1 iter)	15035		
	HSE06 3	6701	52.50	0.03
	Total (2 iters)	21736		

Nonetheless, even for this relatively simple test case, useful savings in computing time could potentially be obtained in practice with RVO. Given the poor scaling of computational cost with system size when using advanced electronic-structure methods, we would expect more substantial savings for more complex unit cells. This would also be true in systems where direct optimisation requires the minimisation of a larger number of degrees of freedom, leading a larger number of steps with method B, which is the subject of the following case studies.

B. Quaternary Sulfide Cu_2ZnSnS_4

Cu_2ZnSnS_4 (CZTS) is an attractive light-absorbing material for thin-film photovoltaics, with a direct bandgap and consisting of earth-abundant components, which has attracted significant experimental[50–54] and computational[55–59] research effort. In the search for new materials for solar energy conversion, the prediction of accurate bandgaps from first-principles is a serious challenge and CZTS represents a suitable case for probing the effect of crystal structure.

An initial structure for CZTS in the kesterite phase, optimised with PBEsol, was obtained during previous work.[60] This was reduced from a conventional 16-atom body-centered-tetragonal cell with $\bar{I}4$ symmetry to the corresponding 8-atom primitive cell using Spglib.[61] A set of seven structures was obtained for both cells by multiplying each lattice vector by a scale factor from 0.97–1.03 and performing a local optimisation of the atomic positions, this forming the "method A" energy-volume curves. In addition to this isotropic scaling, an additional set of structures were calculated including optimisation of the cell shape (i.e. the tetragonal c/a ratio) for each volume point. The iterative RVO procedure was followed in order to minimise the pressure and obtain a more accurate electronic structure using the HSE06 functional. The results are given in Table III; pressure minimisation was rapid in all cases, decreasing logarithmically with each step.

We note that the resulting lattice parameters from these calculations, especially those using the primitive cell, are very close to both the experimental lattice parameter a=5.427 Å and theoretical a=5.448 Å reported by Paier et al. following a conventional optimisation procedure with a variant of the HSE functional.[62] We also note a bandgap shift of almost 0.1 eV when the E-V curve was provided by a non-isotropic set of primitive lattices.

This case also highlights the importance of internal structure optimisation. After two steps of optimisation using the E-V curve further calculations were carried out, employing the HSE06 exchange-correlation functional, where the internal atomic positions were relaxed while fixing the unit cell. As shown in the table, these lead to an increase in the absolute pressure, but also a considerable improvement in the bandgap estimation compared to experimental measurements. Previous electronic structure studies have show that the bandgap of CZTS is highly sensitive to the S positions, which correspond to deviations away from the ideal tetrahedral coordination environment.[56] In the ideal kesterite crystal structure, the metal nuclei all occupy high symmetry Wyckoff positions ($2a$ and $2c$ by Cu, $2d$ by Zn, and $2b$ by Sn). However, the sulfur anions occupy the lower symmetry $8g$ positions, with x, y and z displacement parameters. The change of ~ 0.3 eV in the bandgap following further optimisation (Table III) emphasises the importance of internal relaxations for quantitative studies of electronic structure.

C. Metal-organic Framework HKUST-1

In 1999, Williams and co-workers isolated $Cu_3(btc)_2$ (HKUST-1).[63] Since then this material has been widely studied in the field metal-organic frameworks (MOFs), with possible applications in catalysis, ionic and electrical conductivity, photovoltaics, batteries and gas capture.[64] First-principles calculations of MOF properties have traditionally posed challenges for computational chemists because they combine large unit cells with complex organic and inorganic components.

HKUST-1 features an additional layer of complexity: it is composed of Cu-Cu "paddlewheel" inorganic regions, where each Cu(II) atom is associated with an unpaired electron and each paddlewheel is antiferromagnetically (AFM) coupled in the ground state configuration.[65,66] The magnetic interactions are highly sensitive to the Cu-Cu separation. Moreover, previous studies[67,68] have

TABLE III: Results from application of RVO to Cu_2ZnSnS_4, using the HSE06 functional and a PBEsol-derived E-V curve. The unit of pressure P is kbar (10^8 Pa) and volumes V are given in Å^3. a is the lattice parameter in Å; these are calculated as a mean over the a and b vectors (crystal symmetry is not enforced in FHI-AIMS). E_g is the electronic bandgap in eV taken from the Kohn-Sham eigenvalues at the Γ-point. The methods in parentheses refer to the process by which the E-V curve was generated; isotropic expansion energies with FHI-aims and volume-conserving relaxations with VASP. Iteration "2*" is the data from a final set of calculations, where the internal atomic positions are relaxed while maintaining the unit cell from iteration 2.

Iteration	Conventional cell (Isotropic expansion)				Primitive cell (Isotropic expansion)				Primitive cell (Constrained relaxation)			
	P	V	a	E_g	P	V	a	E_g	P	V	a	E_g
0	22.82	309.12	5.38	1.26	17.49	155.43	5.38	1.23	17.49	155.43	5.38	1.23
1	−1.23	318.03	5.40	1.18	−0.64	158.87	5.42	1.15	−1.35	159.00	5.44	1.13
2	0.00	317.56	5.40	1.19	−0.01	158.74	5.42	1.15	0.02	158.73	5.43	1.14
2∗	7.46	317.56	5.40	1.49	10.17	158.74	5.42	1.48	10.31	158.73	5.43	1.47

TABLE IV: Results from volume optimisation of HKUST-1. Residual pressure P at each step, energies of valence band maximum (VBM) and conduction band minimum (CBM) with respect to the vacuum level, and the bandgap (E_g). All energies are given in eV and the pressures is in kbar (10^8 Pa).

Iteration	P	VBM	CBM	E_g
0	−13.98	−7.5	−3.7	3.8
1	−1.09	−7.0	−3.5	3.5

shown that the ionisation potentials and bandgaps of porous materials are sensitive to cell pressure and volume, similar to some of their inorganic counterparts. HKUST-1 represents an extreme case, where deviations from the equilibrium Cu-Cu distance result in spin flipping and formation of a ferromagnetic (FM) state, which impacts the electronic structure.

Typically, PBEsol-optimised structures agree with low-temperature experimental measurements of MOFs to within 1 %. This is the case here, and PBEsol also reproduces the correct AFM state. However, a single-point HSE06 calculation on the PBEsol structure yields an incorrect FM ground-state, as shown in Fig. 6, with an associated HSE06 cell pressure of -13.98 kbar (Table IV). In order to recover an accurate HSE06 crystal structure (and the associated correct magnetic structure), a single iteration of RVO was required. Notably, there is not only a magnetic difference, but also a significant difference in predicted electronic bandgap and workfunction.

VI. CONCLUSIONS

The rapid volume optimisation (RVO) approach presented here uses information from an inexpensive energy-volume curve to obtain a useful estimate of the optimal unit cell volume for a different level of theory. Our focus was in bridging between different exchange-correlation functionals within density functional theory, but a measured bulk modulus or classical interatomic potential could also be used to construct the reference energy-volume data. In sensitive systems the volume change can lead to qualitative differences in the electronic and magnetic properties. The results depend on the initial volume estimate and are relatively insensitive to the accuracy of the E-V curve. The RVO method is expected to be competitive with conventional optimisation approaches for simple symmetric unit cells, as demonstrated for rocksalt structured PbS. For materials such as Cu_2ZnSnS_4 that are sensitive to the atomic positions within the unit cell, RVO may form part of the optimisation approach but direct optimisation of internal positions is still needed. More significantly, it allows for the improved estimation of properties for large unit cells as demonstrated for HKUST-1, where conventional optimisation methods may be infeasible. As the only property used is the hydrostatic pressure, it is possible to employ calculation methods which return a total energy without analytical gradients by evaluating the energy of a single finite difference. In this case, an improvement over the "single-shot" may be obtained with two additional high-level calculations and an inexpensive E-V curve; a fourth high-level calculation would give an estimate of the convergence. We expect that in many cases this will prove an economical approach for the application of state-of-the-art electronic structure calculations in the solid state.

VII. SUPPLEMENTARY DATA

Python code providing a reference implementation of this method is available at `http://github.com/wmd-bath/rvo`; a snapshot as of this publication is available at DOI:10.5281/zenodo.31940. This includes a program to generate the plots in this publication from the binary chalcogenide data. Calculation data has been deposited online with Figshare at the DOI:

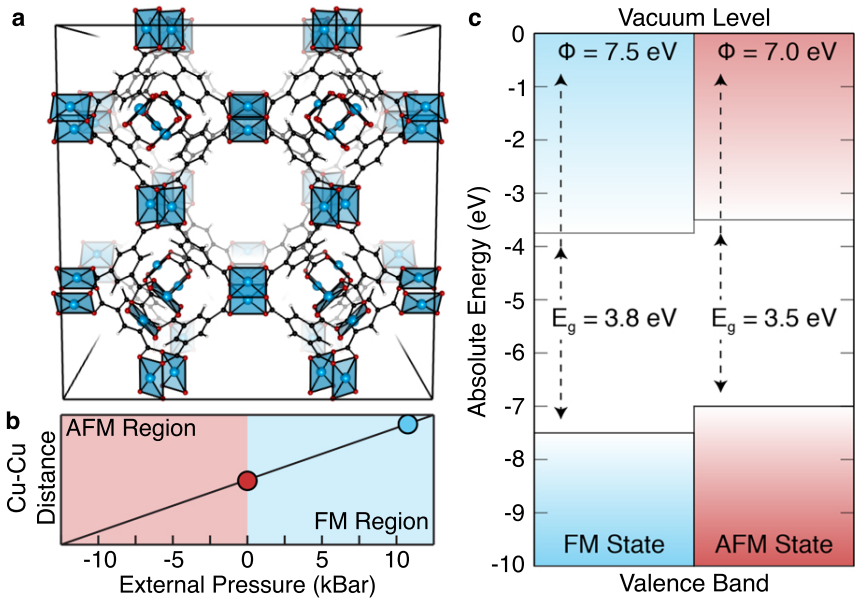


FIG. 6: (a) HKUST-1 features a periodic array of 32 Cu(II)-Cu(II) paddlewheels per crystallographic unit cell. (b) The favoured magnetic structure depends on the Cu-Cu separation: antiferromagnetic (AFM) and ferromagnetic (FM) states are accessible. (c) The valence band energy (ionisation potential) is sensitive to the magnetic structure (calculated using the procedure outlined in Ref. 67). A 'single-shot' HSE06 calculation on the PBEsol structure favours the FM state (blue), whilst the corrected structure favours the experimentally observed AFM state (red).

10.6084/m9.figshare.1468388. Raw output files from the hybrid DFT calculations are made available for CZTS and HKUST-1, and energy-volume curves are available for all the systems. The full set of graphs and fitting parameters for PbS, PbTe, ZnS, and ZnTe are also included as supplementary data with this paper.

ACKNOWLEDGMENTS

The authors thank J. M. Frost for useful discussions. We acknowledge membership of the UK's HPC Materials Chemistry Consortium, which is funded by EPSRC grant EP/L000202. J.M.S. is funded by an EPSRC Programme Grant (EP/K004956/1). A.J.J. is funded by the EPSRC Doctoral Training Centre in Sustainable Chemical Technologies (EP/G03768X/1 and EP/L016354/1). A.W. acknowledges support from the Royal Society and the ERC (grant no. 277757).

[1] C. R. A. Catlow, Z. X. Guo, M. Miskufova, S. A. Shevlin, A. G. H. Smith, A. A. Sokol, A. Walsh, D. J. Wilson, and S. M. Woodley, Philos. Trans. A. **368**, 3379 (2010).
[2] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[3] P. Vinet, J. H. Rose, J. Ferrante, and J. R. Smith, J. Phys.: Condens. Matter **1**, 1941 (1989).
[4] G. P. Francis and M. C. Payne, J. Phys.: Condens. Matter **2**, 4395 (1990).
[5] A. Walsh, J. Da Silva, and S.-H. Wei, Phys. Rev. Lett. **100**, 256401 (2008).
[6] S. Lany and A. Zunger, Model. Simul. Mater. Sci. Eng. **17**, 084002 (2009).
[7] T. Shimazaki and Y. Asai, J. Chem. Phys. **130**, 164702 (2009).
[8] A. Walsh, J. L. F. Da Silva, and S.-H. Wei, Chem. Mater. **21**, 5119 (2009).
[9] I. E. Castelli, D. D. Landis, K. S. Thygesen, S. Dahl, I. Chorkendorff, T. F. Jaramillo, and K. W. Jacobsen, Energy Environ. Science **5**, 9034 (2012).
[10] L. A. Burton and A. Walsh, Appl. Phys. Lett. **132111**, 132111 (2013).
[11] C. H. Hendon, R. X. Yang, L. A. Burton, and A. Walsh, J. Mater. Chem. A **3**, 9067 (2015).
[12] D. Bhachu, D. Scanlon, E. Saban, H. Bronstein, I. Parkin, C. Carmalt, and R. Palgrave, J. Mater. Chem. A **3**, 9071 (2015).
[13] S. Dudarev, L.-M. Peng, S. Savrasov, and J.-M. Zuo, Phys. Rev. B **61**, 2506 (2000).
[14] W. Kohn and L. J. Sham, Phys. Rev. **385**, 1133 (1965).
[15] G. Kresse, M. Marsman, and J. Furthmüller, *VASP the GUIDE* (Universität Wien, Wien, Austria, 2014).
[16] S. Baroni, A. Dal Corso, S. de Gironcoli, and P. Giannozzi, "Pwscf and phonon: Plane-wave pseudo-potential codes," (2001).
[17] R. Dovesi, V. R. Saunders, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N. M. Harrison, I. J. Bush, P. D'Arco, M. Llunell, M. Causà, and Y. Noël, *CRYSTAL14 User's Manual* (University of Torino, Torino, 2014).
[18] J. P. Perdew, J. Chevary, S. Vosko, K. A. Jackson, M. R. Pederson, D. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).
[19] J. Perdew, A. Ruzsinszky, G. Csonka, O. Vydrov, G. Scuseria, L. Constantin, X. Zhou, and K. Burke, Phys. Rev. Lett. **100**, 136406 (2008).
[20] M. De la Pierre, R. Orlando, L. Maschio, K. Doll, P. Ugliengo, and R. Dovesi, J. Comput. Chem. **32**, 1775 (2011).
[21] G. Csonka, J. Perdew, A. Ruzsinszky, P. Philipsen, S. Lebègue, J. Paier, O. Vydrov, and J. Ángyán, Phys. Rev. B **79**, 155107 (2009).
[22] A. V. Krukau, O. a. Vydrov, A. F. Izmaylov, and G. E. Scuseria, J. Chem. Phys. **125** (2006), 10.1063/1.2404663.

[23]K. S. Knight, J. Phys.: Condens. Matter **26**, 385403 (2014).
[24]F. D. Murnaghan, Proc. Natl. Acad. Sci. U. S. A. **30**, 244 (1944).
[25]G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).
[26]P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[27]G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
[28]H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).
[29]J. P. Perdew and K. Schmidt, AIP Conference Proceedings **577**, 1 (2001).
[30]J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[31]Y. Wang and J. P. Perdew, Phys. Rev. B **44**, 13298 (1991).
[32]J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **46**, 6671 (1992).
[33]J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, Phys. Rev. B **48**, 4978 (1993).
[34]J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[35]S. Grimme, J. Comput. Chem. **27**, 1787 (2006).
[36]J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).
[37]J. P. Perdew, A. Ruzsinszky, G. I. Csonka, L. A. Constantin, and J. Sun, Phys. Rev. Lett. **103**, 026403 (2009).
[38]A. D. Becke, J. Chem. Phys. **98**, 1372 (1993).
[39]S. Kastbjerg, N. Bindzus, M. Sndergaard, S. Johnsen, N. Lock, M. Christensen, M. Takata, M. A. Spackman, and B. Brummerstedt Iversen, Advan. Funct. Mater. **23**, 5477 (2013).
[40]G. J. McIntyre, G. Moss, and Z. Barnea, Acta Crystallogr., Sect. A **36**, 482 (1980).
[41]V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, Comput. Phys. Commun. **180**, 2175 (2009).
[42]V. Havu, V. Blum, P. Havu, and M. Scheffler, J. Comput. Phys. **228**, 8367 (2009).
[43]L. Schimka, J. Harl, and G. Kresse, J. Chem. Phys. **134**, 024116 (2011).
[44]T. E. Oliphant, Comput. Sci. Eng. **9**, 10 (2007).
[45]"SciPy: Open source scientific tools for Python," (2001), [Online; accessed 2015-03-06].
[46]J. D. Hunter, Comput. Sci. Eng. **9**, 90 (2007).
[47]D. W. Marquardt, J. Soc. Indust. Appl. Math. **11**, 431 (1963).
[48]K. Hummer, A. Grüneis, and G. Kresse, Phys. Rev. B **75**, 1 (2007).
[49]J. M. Skelton, S. C. Parker, A. Togo, I. Tanaka, and A. Walsh, Phys. Rev. B **89**, 205203 (2014).
[50]S. Schorr, Sol. Energy Mater. Sol. Cells **95**, 1482 (2011).
[51]S. Siebentritt and S. Schorr, Prog. Photovoltaics Res. Appl. **20**, 512 (2012).
[52]J. J. Scragg, P. J. Dale, D. Colombara, and L. M. Peter, ChemPhysChem **13**, 3035 (2012).
[53]O. Gunawan, T. Gokmen, and D. B. Mitzi, J. Appl. Phys. **116**, 084504 (2014).
[54]J. J. S. Scragg, L. Choubrac, A. Lafond, T. Ericson, and C. Platzer-Björkman, Appl. Phys. Lett. **104**, 041911 (2014).
[55]S. Chen, X. G. Gong, A. Walsh, and S.-H. Wei, Phys. Rev. B **79**, 165211 (2009).
[56]S. Botti, D. Kammerlander, and M. a. L. Marques, Appl. Phys. Lett. **98**, 241915 (2011).
[57]C. Persson, J. Appl. Phys. **107** (2010), 10.1063/1.3318468.
[58]A. Walsh, S. Chen, S.-H. Wei, and X.-G. Gong, Adv. Energy Mater. **2**, 400 (2012).
[59]J. M. Skelton, A. J. Jackson, M. Dimitrievska, S. K. Wallace, and A. Walsh, APL Mater. **3**, 041102 (2015).
[60]A. J. Jackson and A. Walsh, Journal of Materials Chemistry A **2**, 7829 (2014).
[61]A. Togo, "Spglib," [Online; accessed 2015-03-06].
[62]J. Paier, R. Asahi, A. Nagoya, and G. Kresse, Phys. Rev. B **79**, 1 (2009).
[63]S. Chui, S. Lo, J. Charmant, A. Orpen, and I. Williams, Science **283**, 1148 (1999).
[64]C. H. Hendon, K. E. Wittering, T.-H. Chen, W. Kaveevivitchai, I. Popov, K. T. Butler, C. C. Wilson, D. L. Cruickshank, O. Š. Miljanić, and A. Walsh, Nano Lett. **15**, 2149 (2015), pMID: 25706577, http://dx.doi.org/10.1021/acs.nanolett.5b00144.
[65]X. X. Zhang, S. S.-Y. Chui, and I. D. Williams, J. Appl. Phys. **87**, 6007 (2000).
[66]A. Pöppl, S. Kunz, D. Himsl, and M. Hartmann, J. Phys. Chem. C **112**, 2678 (2008), http://dx.doi.org/10.1021/jp7100094.
[67]K. T. Butler, C. H. Hendon, and a. Walsh, J. Am. Chem. Soc. **136**, 2703 (2014).
[68]K. T. Butler, C. H. Hendon, and A. Walsh, ACS Appl. Mater. Interfaces **6**, 22044 (2014).