Beyond two-centre tight binding: Models for Mg and Zr

R. M. Fogarty, J. Smutna, M. R. Wenman, and A. P. Horsfield*

Department of Materials and Thomas Young Centre,

Imperial College London, South Kensington Campus,

London SW7 2AZ, United Kingdom

(Dated: December 1, 2020)

Abstract

We describe a systematic approach to building *ab initio* tight binding models and apply this to hexagonal metals Mg and Zr. Our models contain three approximations to plane-wave density function theory (DFT): i) we use a small basis set, ii) we approximate self-consistency and iii) we approximate many-centre exchange and correlation effects. We test a range of approximations for many-centre exchange-correlation and self-consistency to gauge the accuracy of each in isolation. This systematic approach also allows us to combine multiple approximations in the optimal manner for our final tight binding models. Furthermore, the breakdown of errors into those from individual approximations is expected to be a useful guide for which approximations to include in other tight binding models. We attempt to correct any remaining errors in our models by fitting a pair-potential. Our final tight binding model for Mg shows excellent agreement with plane-wave results for a wide range of properties (*e.g.* errors below 10% for self-interstitial energies and below 3% for equilibrium volumes) and is expected to be highly transferable due to the minimal amount of fitting. Calculations with our Zr model also show good agreement with plane-wave results (*e.g.* errors below 2% for equilibrium volumes) except for properties where self-consistency is important, such as self-interstitial formation energies. However, for these properties we are able to generate a tight binding model which shows excellent agreement with non self-consistent DFT with a small basis set (*i.e.* many-centre effects are captured accurately by our approximations). As we understand the source of remaining errors in our Zr model we are able to outline the methods required to build upon it to describe the remaining properties with greater accuracy.

* a.horsfield@imperial.ac.uk

I. INTRODUCTION

In our previous work [1] we focused on building tight-binding models for Mg and Zr (hexagonal metals) for which only pairwise terms were calculated (*i.e.* two-centre models). The resultant models worked reasonably well for near-equilibrium structures to which they were fit, but led to unphysical results for structures with large numbers of many-centre interactions (*e.g.* unrelaxed self-interstitials were predicted to be have negative formation energies). A major problem was the use of a pairwise approximation for calculating the hopping integrals; these include integrals of the form $\langle I\alpha|\sum_K V_K|J\beta\rangle$ where $\phi_{I\alpha}(\vec{r}) = \langle\vec{r}\,|\,I\alpha\rangle$ is an orbital on atom I, $\phi_{J\beta}(\vec{r})$is an orbital on atom J ($I \neq J$) and V_K is the potential from atom K. Within our pairwise approximation we set $\sum_K V_K \approx V_I + V_J$, a standard approximation made in tight-binding models, which means the effect of the local environment is not accounted for in this integral. Thus, in the current work we focus on improving our previous tight-binding models by accounting for many-centre effects more accurately.

Previous attempts to account for many-centre effects on hopping integrals in tight binding models can be divided into three main groups: i) implicitly accounting for the effects in two-centre functions;[2–6] ii) scaling two-centre integrals by an environmentally dependent term;[7–9] iii) applying *ab initio* corrections for many-centre effects, notably exchange and correlation effects.[10–13] In our previous work we attempted to use method i) by fitting two-centre hopping integrals to the ideal ones for a series of near-equilibrium perfect crystal structures.[1] Unfortunately we found that the ideal hopping integral, for a given distance, varied significantly between defect and equilibrium structures; thus we concluded that it was necessary to use corrections with an explicit environmental dependence. Method ii) provides an explicit environmental dependence but relies on the assumption that many-centre effects can be approximated as a fraction of a two-centre interaction. However, our *ab initio* two-centre hopping integrals contain nodes, often at interatomic separations near equilibrium crystal bond lengths; thus, many-centre effects would spuriously be predicted as zero for two atoms separated by this distance. It is therefore unclear how to apply method ii) to the current case. We therefore chose to apply method iii) in the current work.

In the current work we will deal with many-centre electrostatic contributions without approximation by explicitly calculating up to three-centre integrals. These integrals can be done analytically by representing atomic orbitals and potentials as sums of Gaussian

functions, the implementation used in this work, or from interpolation of pre-tabulated two-dimensional values, as used in the FIREBALL code.[10] Unlike electrostatic terms, many-centre exchange-correlation contributions can not be split into a sum of three-centre terms due to the non-additivity of the exchange-correlation potential (ν_{xc}). Furthermore we cannot analytically calculate integrals involving ν_{xc}. Thus, for exchange and correlation we treat up to two-centre terms exactly by using pre-tabulated integrals and approximate the remaining many-centre effects using the techniques described in the methods section.

The choice of hexagonal metals, Mg and Zr, in this work is motivated by our interest in modelling and understanding corrosion processes in both. Mg is highly abundant, biodegradable and the lightest of the structural metals. These properties mean Mg-alloys are useful materials for applications such as biomedical implants and automotive parts as lighter parts means better fuel efficiency.[14, 15] However, the use of Mg-alloys is currently limited due to their generally poor resistance to aqueous corrosion.[16] Zirconium alloys are commonly used for the cladding of nuclear fuel rods. In operation, the rods are cooled by pressurized water or steam, leading to buildup of an oxide layer and hydrogen pickup by the alloy. This corrosion limits the safe operational life of the fuel, and therefore needs to be reduced when designing improved alloys.[17]

In order to study corrosion using tight binding, we need models that are accurate for structures that are far from equilibrium. Studying aqueous corrosion of Mg will require, in particular, a model which describes surfaces (the site of corrosion) well and leads to accurate electronic structure as the metal will become charged during corrosion processes. For Zr we need to be able to describe radiation induced defect structures accurately, notably point defects and dislocations, and how they interact with each other. With these considerations we have chosen to test the ability of our models to predict a wide range of properties including formation energies of self-interstitials and vacancies, the hcp (0001) surface energy, near-equilibrium structural properties, and the bulk structure density of states.

We will first give a brief description of our methodology; a more detailed description of many aspects can be found in our previous paper.[1] Our focus will then be on ways of correcting many-centre exchange-correlation effects as we examine a range of corrections in isolation to assess the accuracy of each approximation which goes into our model. This allows us to combine not just the most accurate individual approximations, but those which have favorable error-cancellation properties. We will subsequently focus on corrections for

the large self-consistency effects found for Zr. Finally, we present results from the tight-binding models we developed in this work; focusing on accuracy, limitations and possible improvements to the models.

II. METHODS

A. The Harris-Foulkes functional

The Harris-Foulkes functional represents the starting point for our tight-binding calculations.[18, 19] This functional approximates the Kohn-Sham functional with a Taylor series expansion around a reference density; this is most easily seen in the Hohenberg-Kohn [20] formulation:

$$\begin{aligned} E[n] =& \underbrace{E[n^{(0)}]}_{E_0} + \underbrace{\int \left.\frac{\delta E}{\delta n(\vec{r})}\right|_{n^{(0)}} q(\vec{r})\mathrm{d}\vec{r}}_{E_1} + \\ & \underbrace{\frac{1}{2}\int\int \left.\frac{\delta^2 E}{\delta n(\vec{r})\delta n(\vec{r}')}\right|_{n^{(0)}} q(\vec{r})q(\vec{r}')\mathrm{d}\vec{r}\mathrm{d}\vec{r}'}_{E_2} + \ldots \\ =& E_0 + E_1 + E_2 + \ldots \end{aligned} \tag{1}$$

where $n^{(0)}$ is the reference density and $q(\vec{r})$ is the difference between the actual density and reference density: $q(\vec{r}) = n(\vec{r}) - n^{(0)}(\vec{r})$. In our case E_0 is the energy generated by the sum of atomic densities. Only the kinetic energy and non-local pseudopotential terms require explicit knowledge of the atomic orbitals, and can be pre-calculated. The remaining terms in E_0 can be calculated without approximation without reference to orbitals and most can be pre-calculated: a single integral over the density is needed, however, for exchange and correlation. E_1 represents the energy from the first order electronic relaxation; its calculation requires the building of a Hamiltonian and density matrix based on the potential generated by the reference density. Higher order terms account for contributions from self-consistency.

In the current work we generally truncate the expansion at E_1 (non self-consistent models), but also explore the use of approximations for E_2 (self-consistent models) for Zr.

B. Approximating E_0

E_0 can be split up into contributions from the isolated atoms (E_I^{atom}), a pairwise sum of potentials (representing attraction or repulsion between them) and a correction for many-centre exchange and correlation effects (E_{xc}^{MB}):

$$E_0 = \sum_I E_I^{atom} + \frac{1}{2}\sum_{I\neq J} V_{IJ}(R_{IJ}) + E_{xc}^{MB}[n^{(0)}] \tag{2}$$

where I and J are atom indices, V_{IJ} is the pair potential and R_{IJ} is the distance between two atoms. The pair-potential can be calculated and tabulated in advance, leading to rapid evaluation. The sum over atomic energies will cancel out when comparing energies of different structures: as a result, we generally work with the cohesive energy, defined as the difference between the energy of the system we are interested in and the free atoms. We can obtain this by subtracting $\sum_I E_I^{atom}$ from E_0, leaving just pairwise and multi-centre contributions. E_{xc}^{MB} can either be calculated using numerical integration, which is expensive, or be approximated.

We approximate E_{xc}^{MB} contributions by using a Taylor expansion of the exchange-correlation energy density (ϵ_{xc}) around an effective density which differs for each atom ($\bar{n}_I$). It is unclear how to do a similar analysis to include gradient corrections, hence this is an approximation to the many body LDA E_{xc} value. The total E_{xc} can be calculated this way as:

$$E_{xc} \approx \tilde{E}_{xc} = \sum_I \int n_I(\vec{r})\ \{\epsilon_{xc}(\bar{n}_I) + \epsilon'_{xc}(\bar{n}_I)(n(\vec{r}) - \bar{n}_I) + 0.5\epsilon''_{xc}(\bar{n}_I)(n(\vec{r}) - \bar{n}_I)^2\}\mathrm{d}\vec{r} \tag{3}$$

The first order terms can be eliminated by an appropriate choice of effective density,

$$\bar{n}_I = \sum_J \frac{\int n_I(\vec{r})\, n_J(\vec{r})\,\mathrm{d}\vec{r}}{\int n_I(\vec{r})\,\mathrm{d}\vec{r}},$$

leaving:

$$\begin{aligned}\tilde{E}_{xc} = &\sum_I \epsilon_{xc}(\bar{n}_I)\int n_I(\vec{r})\,\mathrm{d}\vec{r} \\ &+ \frac{1}{2}\epsilon''_{xc}(\bar{n}_I)\left[\int n(\vec{r})\, n(\vec{r})\, n_I(\vec{r})\,\mathrm{d}\vec{r}\ - 2\bar{n}_I\int n(\vec{r})\, n_I(\vec{r})\,\mathrm{d}\vec{r} + \bar{n}_I^2\ \int n_I(\vec{r})\,\mathrm{d}\vec{r}\right]\end{aligned} \tag{4}$$

In order to get the many-centre correction within this approximation we need to subtract off the one-body and two-body terms:

$$\Delta\tilde{E}_{xc}^{MB} = \tilde{E}_{xc}\left[\sum_I n_I\right] - \sum_I \tilde{E}_{xc}[n_I] - \frac{1}{2}\sum_{I\neq J}\left(\tilde{E}_{xc}[n_I+n_J] - \tilde{E}_{xc}[n_I] - \tilde{E}_{xc}[n_J]\right) \quad (5)$$

If we truncate equation 4 at zeroth order we recover the uniform density approximation we introduced in our previous paper.[1] This approximation requires only two-centre integrals to be calculated (to get the effective densities), and is thus highly efficient. However, we can calculate up to second order (the full expression in equation 4) with three-centre integrals; we test both zeroth-order and second-order approximations in the current work.

C. Approximating E_1

E_1 approximations can be split into those for hopping integrals and crystal-field integrals ($\langle I\alpha|\sum_{K\neq I} V_K + \nu_{xc}|I\beta\rangle$). For hopping integrals we approximate many-centre exchange and correlation effects using the McWEDA method. This is described in detail in reference [11]; we note that within this approximation two-centre exchange-correlation is treated exactly for hopping integrals while many-centre effects are calculated as an approximation to the local density approximation (LDA) many-centre effects in a similar fashion to our approximate E_0 corrections. In terms of efficiency, this requires three-centre integrals over the density; these have a similar computational cost to the integrals used to calculate electrostatic terms. For crystal-field terms we test both the McWEDA method and the Sankey-Niklewski (SN) many-centre correction with two-centre terms treated exactly (the whole method, two-centre and correction, is referred to as just the SN correction below).[1, 13] In both cases many-centre exchange and correlation effects are approximations to the LDA result. There are two main differences between these corrections: i) the SN correction treats up to two-centre exchange and correlation exactly, while the McWEDA crystal-field correction treats all exchange and correlation approximately and ii) the SN correction ignores off-diagonal terms while the McWEDA correction does not.

D. Approximating E_2

We approximate self-consistency within a tight binding framework by using the methods described in reference [21]. Briefly, this involves representing the charge density in terms of charge density moments. Solving the Harris-Foulkes equations for a given set of input charge density moments allows us to obtain a new set of moments from the density matrix. We vary the input charge density moments until we obtain a set of output moments that is sufficiently close to the input moments used to calculate them; *i.e.* we solve the Harris-Foulkes functional self-consistently with respect to the charge density moments.

The charge density moments enter the Harris-Foulkes functional through the E_2 term for which we approximate only the Hartree contribution.

$$E_2 \approx \delta E_{Ha} = \frac{1}{2} \int \frac{\delta n(\vec{r}) \delta n(\vec{r}')}{|\vec{r} - \vec{r}'|} \mathrm{d}\vec{r} \mathrm{d}\vec{r}' \tag{6}$$

We represent these changes in density in terms of atom-centred charge density moments:

$$\delta n(\vec{r}) = \sum_{Ilm} Q_{Ilm} g_{Ilm}(\vec{r}) \tag{7}$$

where Q_{Ilm} are charge density moments, g_{Ilm} are atom-centred multipoles, the index I runs over atoms and the combined lm indices run over multipoles. Each multipole function is expressed as:

$$g_{Ilm}(\vec{r}) = e^{-\alpha_I (|\vec{r} - \vec{R}_I|)^2} K_{Ilm}(\vec{r} - \vec{R}_I) \tag{8}$$

where K_{Ilm} is a spherical harmonic function representing the angular dependence while the radial dependence is represented by a single Gaussian function, which has been found to be sufficient in previous studies.[21, 22] We use spherical harmonics with l=0,1,2 in the current work, which correspond to monopoles, dipoles and quadrupoles respectively. The parameter α_I varies between elements and is derived from the Hubbard parameter.

$$\alpha_I = \frac{\pi}{8} U_I^2 \tag{9}$$

This is calculated *ab initio* in the current work. With the charge density moments defined, the approximate E_2 energy can be written as:

$$
\begin{aligned}
E_2 &\approx \frac{1}{2} \sum_{Ilm,Jl'm'} Q_{Ilm} Q_{Jl'm'} \int \frac{g_{Ilm}(\vec{r}) g_{Jl'm'}(\vec{r}' - \vec{R_{IJ}})}{|\vec{r} - \vec{r}'|} \, \mathrm{d}\vec{r}\mathrm{d}\vec{r}' \\
&\approx \frac{1}{2} \sum_{Ilm,Jl'm'} Q_{Ilm} Q_{Jl'm'} \; B_{Ilm,Jl'm'}(\vec{R_{IJ}})
\end{aligned}
\tag{10}
$$

This leads to the following corrections to the Hamiltonian being required:

$$
\begin{aligned}
\delta H_{I\alpha J\beta} = \frac{1}{4} \sum_{Klm,l'm'} Q_{Klm} \; (& M^{l'm'}_{J\beta I\alpha}(B_{Jl'm'Klm} + B_{Kl'm'Jlm}) \\
& + M^{l'm'}_{I\alpha J\beta}(B_{Il'm'Klm} + B_{Kl'm'Ilm}))
\end{aligned}
\tag{11}
$$

$$
M^{lm}_{I\alpha J\beta}(\vec{R_{IJ}}) = \int \phi_{I\alpha}(\vec{r}) K_{lm}(\vec{r}) \phi_{J\beta}(\vec{r} - \vec{R_{IJ}}) d\vec{r} \tag{12}
$$

E. Conventions for naming methods

We employ the same naming convention for methods as our previous paper. Firstly, LCAO is the name given to the method whereby we calculate E_0 and E_1 terms in the Harris-Foulkes functional without approximation. The scf-LCAO method involves using the Kohn-Sham functional with our small basis sets, and is thus fully self-consistent. Other method names are based on the ways in which they approximate the LCAO method. Approximation to LCAO is divided into three parts; hopping integral approximations (H), crystal-field integral approximations (XT) and E_0 approximations (PP). The overall method name is created by writing each label followed by any approximations to LCAO used. For example, the method name H_Ex_XT_Mc_PP_2b means hopping terms are treated exactly(Ex), crystal-field terms are treated with the McWEDA approximation (Mc) and E_0 terms are treated exactly up to two-body(2b). For self-consistent models an extra scf label is added to the method name followed by a number representing the order of the multipole expansion used; scf0,scf1 and scf2 represent methods using up to monopoles (zeroth-order), dipoles and quadrupoles respectively. For example, H_Ex_XT_Ex_PP_Ex_scf0 represents the LCAO method with self-consistency approximated by a zeroth-order multipole expansion.

The hopping integral methods used in this paper are two-body exchange-correlation (2bxc), the McWEDA approximation (Mc) and exact (Ex). The crystal field methods

used are two-body (2b), two-body plus the SN correction (2b_mbSN), McWEDA(Mc) and exact(Ex). For E_0 terms the methods we use are two-body (2b), a fitted two-body pair-potential (fit), the uniform density approximation to zeroth-order (uden0), to second order(uden2) and exact(Ex). Multiple approximations can follow each main label; for example, in the "H_Mc_XT_Mc_PP_fit_uden0" method E_0 is calculated using a combination of the fitted pair-potential and the zeroth-order uniform density approximation. Note that the 2b approximation for crystal-field and pair-potential terms is equivalent to the 2bxc approximation.

F. Fitting the pair potential

Generating our tight binding models involves replacing the *ab initio* pair potential in E_0 with a fitted pair potential. This is a common method in tight-binding models to account for errors introduced by the various approximations (*e.g.* small basis sets used, and neglect of many-centre effects). Our starting point is to fit an analytical form to the *ab initio* pair-potential. We use four free parameters (A_i) and use a functional form based on the work of Krishnapriyan *et al.*[2] and described in more detail in our previous paper.[1]

$$\begin{aligned} f(r) &= f_0 \zeta(r) N(r) T(r) \\ \zeta(r) &= \exp\left(\sum_{i=1}^{p} A_i (r - r_0)^i\right) \\ N(r) &= \prod_{n=1}^{n_{nodes}} \frac{r - r_n}{r_0 - r_n} \\ T(r) &= \exp\left(\frac{d}{r - r_c}\right) \end{aligned} \tag{13}$$

where d is set to 0.5 a_0, r_0 is a reference bond length, $f(0)$ is the value of the (initial) pair-potential at r_0, r_c is a cutoff distance and r_n is the position of a node. The number of nodes was fixed to that of the initial pair-potential, but their positions were allowed to vary in the fit. Note that, in the current case, each pair-potential had only a single node. Also note that the tail function ($T(r)$) ensures that the fitted potential decays to zero outside the radius of each atom (all basis functions also decay to zero by this point).

We judge the goodness of fit (GOF) to a given property, P, as GOF $= w\frac{\sqrt{(P-P_0)^2}}{P_0}$ where P_0 is the reference value of the property and w is a weight (larger meaning more important).

When fitting to dissociation curves (where GOF is evaluated at each dimer separation) the total GOF was taken as the average GOF for each structure; this means that adding more structures will not automatically increase the GOF value. The total objective function for fitting is simply the sum of all individual GOF values.

We fit the Mg pair-potential to reproduce plane-wave results for hcp equilibrium volume (V_0), bulk modulus (B_0) and LCAO dissociation energies between 2.0 a_0 and 5.0 a_0. The fitting to short-ranged dissociation energies is to ensure that the pair-potential remains sufficiently repulsive at short-distances which would not otherwise be represented in the objective function. The weights for V_0, B_0 and the dissociation energy curve were 50, 5 and 10 respectively. We fit the Zr pair-potential to reproduce plane-wave results for hcp V_0, hcp B_0, bcc V_0, the separation between hcp and bcc energies at equilibrium (ΔE^{bcc}) and LCAO dissociation energies between 2.0 a_0 and 5.0 a_0. The weights for Zr were 50, 5, 50, 10, 115 for hcp V_0, hcp B_0, bcc V_0, ΔE^{bcc} and the dissociation curve respectively. In both cases weights were chosen such that the relative error of V_0 was 10 times as important as B_0, this simply reflects that we believe a 1% error in V_0 is as acceptable as a 10% error in B_0. The weights chosen for the dissociation energy curves were simply those found to be large enough to affect the outcome of the fitting procedure; the effects of systematically tuning these weights were not investigated.

We found that our Mg model predicted an unphysical maximum at ≈ 2 a_0 in the dissociation energy curve (see ESI [23] for the relevent figure). Though these separations are unlikely to occur in any simulations we nonetheless fit a repulsive exponential to dissociation energies in this range; this was multiplied by a tail function with a cutoff of 2 a_0 such that the potential could not affect interactions at any larger separations, thus this fit could be carried out after the main fit to the pair-potential.

G. Computational details

Plane-wave calculations were carried out in CASTEP, while all other calculations were carried out using the PLATO tight binding code.[24, 25] We used the PBE functional for all calculations unless explicitly stated otherwise.[26] For plane-wave calculations we used cut-off energies of 450 eV for Zr and 500 eV for Mg; for all other calculations we used the same single-zeta *spd* basis sets as in reference [1]. Note that our *spd* basis was tuned to reproduce

Method	Hcp V_0 (a_0^3 per atom)	Unrelaxed Vacancy E_F (eV)	Relaxed Interstitial E_F (eV)	hcp(0001) Surface E (J m^{-2})
H_Ex_XT_Ex_PP_2b	137.8	1.34	2.84	0.94
H_Ex_XT_Ex_PP_uden0	158.5	1.08	2.52	0.84
H_Ex_XT_Ex_PP_uden2	164.2	0.92	2.00	0.78
LCAO	155.9	1.08	2.33	0.84

TABLE I. Effects of various E_0 approximations on selected structural and defect properties for Mg. The interstitial energy is for the octahedral site. The surface energy is for the unrelaxed surface. LCAO involves calculating E_0 without approximation; the accuracy of other methods can be judged by their agreement with the LCAO method

properties of the solid system. For plane-wave calculations we used pseudopotentials with core charges of 10 and 12 for Mg and Zr respectively. For all other calculations we used Goedecker type pseudopotentials with core charges of 2 and 4 for Mg and Zr respectively; these are the same pseudopotentials commonly used in the CP2K code.[27, 28] Note that we used a Zr pseudopotential that was optimised for an LDA functional, due to the lack of availability of a suitable Goedecker pseudopotential optimised for the PBE functional. All relaxed structures were optimized using plane-wave calculations; only single-point energy calculations were carried out with the tight binding models. Further methodological details can be found in the ESI.[23]

III. RESULTS

A. Approximating E_0

The effects of our E_0 approximations can be seen by comparing LCAO results to those of H_Ex_XT_Ex_PP_2b, H_Ex_XT_Ex_PP_uden0 and H_Ex_XT_Ex_PP_uden2 (Table I for Mg, ESI[23] for Zr). The pairwise approximation for E_0 is exact except for its neglect of many-centre exchange-correlation terms.

The zeroth-order uniform density approximation is generally found to dramatically improve results compared to the pairwise approximation; equilibrium volumes are very close to

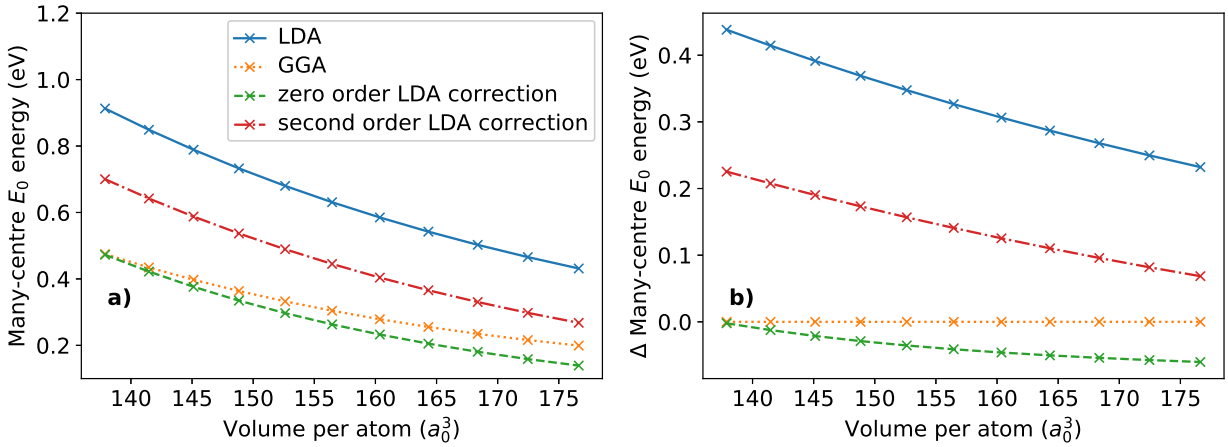


FIG. 1. Many-centre (three-centre and above) contributions to exchange-correlation E_0 for hcp Mg using various approximations. a) Absolute energies, b) energies relative to the GGA value. LDA corrections refer to the uniform density approximation. Note that energies are for the whole (2-atom) unit-cell. The calculations used to obtain these curves are described in the ESI.[23]

the LCAO values and defect energies are generally greatly improved (the exceptions being the Zr relaxed octahedral and split self-interstitials). Surprisingly, including the second-order term in the uniform density approximation leads to less accurate results than truncating after the zeroth order term. The reasons for this can be seen in Figure 1 for Mg (ESI[23] for Zr), which shows the many-centre exchange-correlation contributions to E_0 under various approximations. Focusing on the Mg results, it can be seen that the LDA approximation leads to far larger many-centre effects than the generalized gradient approximation (GGA). Our uniform density correction is an approximation to these (larger than required) LDA many-centre terms. When we take only the zeroth-order term we obtain a relatively small fraction of the LDA many-centre correction and are left with a correction which approximates the GGA many-centre energy well. If we go to a higher order correction we obtain a greater fraction of LDA many-centre energy, meaning we overcorrect with respect to the GGA many-centre energy. The second-order uniform density approximation overcorrects an increasing amount as volume is decreased and thus the correction spuriously stabilises large volume structures. To summarise, the zeroth-order uniform density approximation works well for Mg due to error cancellation; this term underestimates the repulsive LDA many-centre effects by an amount similar to the attractive many-centre gradient corrections. For Zr, similar results are obtained except that the absolute error for the zeroth order correction

Method	Hcp V_0 (a_0^3 per atom)	Unrelaxed Vacancy E_F (eV)	Relaxed Interstitial E_F (eV)	Unrelaxed Interstitial E_F (eV)	hcp(0001) Surface E (J m^{-2})
H_2bxc_XT_Ex_PP_Ex	-	0.09	-12.84	-743.32	0.37
H_Mc_XT_Ex_PP_Ex	161.7	0.64	1.86	5.26	0.66
H_Ex_XT_2b_PP_Ex	164.5	0.92	1.96	-230.58	0.78
H_Ex_XT_2b_MbSN_PP_Ex	155.0	1.09	2.15	5.47	0.85
H_Ex_XT_Mc_PP_Ex	149.6	1.29	2.74	6.87	0.94
H_Mc_XT_2b_MbSN_PP_Ex	160.9	0.66	1.70	-0.06	0.67
H_Mc_XT_Mc_PP_Ex	154.2	0.82	2.25	6.34	0.75
LCAO	155.9	1.08	2.33	6.58	0.84

TABLE II. Effects of various E_1 approximations on selected structural and defect properties for Mg. The interstitial energies are for the octahedral site. LCAO involves calculating E_1 without approximation; the accuracy of other methods can be judged by their agreement with the LCAO method.

is larger; in this case the important factor is that the error in zeroth order energy has a smaller volume dependence (rather than absolute value) than the second order energy.

These results suggest that the zeroth order uniform density approximation will likely be best, in terms of accuracy and speed, for models attempting to approximate DFT with GGA functionals. For models built as approximations to LDA functionals, the second order correction will provide significantly greater accuracy. For models which already include computing three-centre terms, the extra cost of the second order E_0 correction will generally be a small fraction of the total calculation cost. Its use may even be feasible for models which are otherwise two-centre, since its maximum scaling is only K^3, where K is the number of atoms, compared to N^2K for three-centre matrix elements, where N is the number of basis functions.

Method	Hcp V_0 (a_0^3 per atom)	Unrelaxed Vacancy E_F (eV)	Relaxed Interstitial E_F (eV)	Unrelaxed Interstitial E_F (eV)	hcp(0001) Surface E (J m^{-2})
H_2bxc_XT_Ex_PP_Ex	-	96.00	-47.51	-288.41	24.77
H_Mc_XT_Ex_PP_Ex	-	2.04	1.01	2.02	1.37
H_Ex_XT_2b_PP_Ex	-	37.47	-4.73	-123.20	7.74
H_Ex_XT_2b_MbSN_PP_Ex	152.5	2.54	1.62	3.18	1.60
H_Ex_XT_Mc_PP_Ex	152.6	3.34	2.69	4.47	2.10
H_Mc_XT_2b_MbSN_PP_Ex	146.3	1.99	0.92	1.51	1.36
H_Mc_XT_Mc_PP_Ex	151.1	2.75	2.13	3.32	1.85
LCAO	156.9	2.56	1.97	3.53	1.60

TABLE III. Effects of various E_1 approximations on selected structural and defect properties for Zr. The interstitial energies are for the octahedral site. LCAO involves calculating E_1 without approximation; the accuracy of other methods can be judged by their agreement with the LCAO method.

B. Approximating E_1

Approximations for E_1 can be split into two parts; hopping integral corrections and crystal field integral corrections. For hopping integrals, the McWEDA method can be seen to greatly improve results compared to the pairwise treatment: compare H_Mc_XT_Ex_PP_Ex to H_2bxc_XT_Ex_PP_Ex (Tables II and III). In particular, for Mg it leads to stable energy vs volume curves. Nonetheless, when this is the only approximation applied we could not find a minimum in the energy-volume curve for Zr. The reason for this unphysical result can be demonstrated by studying the density of states for a compressed Zr structure (Figure 2). High-energy states in the LCAO calculations are spuriously lowered in energy when using McWEDA hopping integrals; once a certain level of compression is reached these states enter the valence band and spuriously stabilise the structures. Compared to two-centre exchange-correlation, the effect is significantly smaller when using the McWEDA correction (see ESI[23] for a comparison of the density of states) but still leads to spurious stabilisation of small volume structures.

As with hopping integrals, our corrections for crystal-field terms lead to significantly

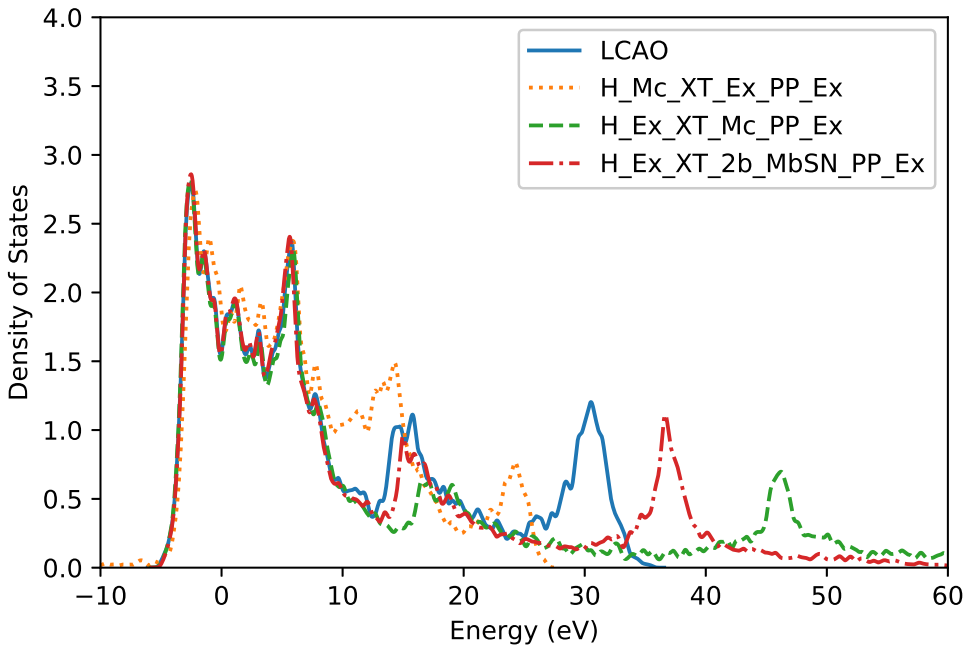


FIG. 2. Effects of E_1 corrections on the density of states for a compressed (115 a_0^3 per atom) Zr hcp structure. The Fermi level is located at 0 eV.

improved results for Mg and Zr compared to the pairwise case (compare H_Ex_XT_2b_PP_Ex, H_Ex_XT_2b_MbSN_PP_Ex and H_Ex_XT_Mc_PP_Ex with LCAO, these methods only differ in how crystal-field terms are treated). Of the two corrections, the SN method leads to significantly more accurate calculated properties. In terms of electronic structure, both corrections lead to qualitatively good agreement with LCAO results near the Fermi level (Figure 2). Both corrections also lead to spurious destabilization of high-energy conduction band states, with the effect being greater for the McWEDA correction. This contrasts with the effect of using the McWEDA correction for hopping integrals, whereby high energy states are spuriously stabilised. As a result, mixing the McWEDA hopping correction with either crystal-field correction leads to a stable model despite the fact that the McWEDA hopping correction alone led to an unstable model. The useful error-cancellation upon mixing these approximations can be seen from the density of states plot in Figure 3; the states at the bottom of the valence band in H_Mc_XT_Ex_PP_Ex are shifted into the conduction band when either crystal-field approximation is added.

In isolation, the McWEDA hopping correction and the SN crystal field corrections are the most accurate approximations for the hopping and crystal field terms respectively. However,

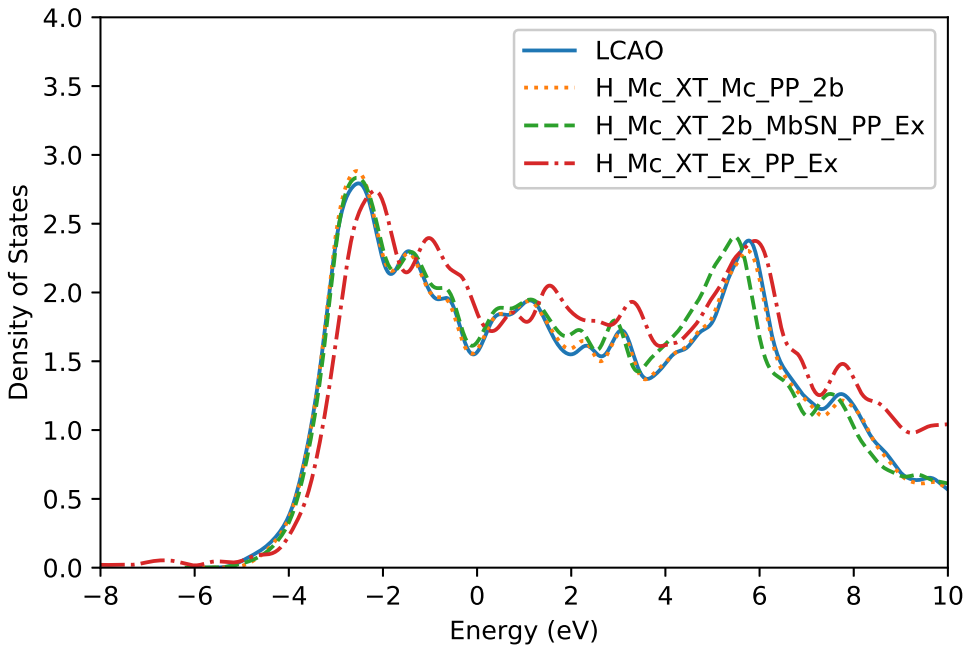


FIG. 3. Effects of combining different E_1 corrections on the density of states for a compressed (115 a_0^3 per atom) Zr hcp structure. The Fermi level is located at 0 eV.

in creating a tight binding model we need to use the corrections which lead to the best results when combined. Focusing on E_1, this is achieved by combining McWEDA hopping and crystal field corrections (compare H_Mc_XT_Mc_PP_Ex and H_Mc_XT_2b_MbSN_PP_Ex). This is a result of error cancellations between crystal field and hopping terms within the McWEDA approximation. This useful error cancellation can be observed in the effect of each correction (hopping and crystal field) on the Mg equilibrium volume, V_0. Using the McWEDA hopping correction alone leads to a Mg V_0 value which is $\approx 6\ a_0^3$ per atom larger than the LCAO result, whilst the McWEDA crystal field correction in isolation leads to V_0 being $\approx 6\ a_0^3$ per atom smaller than the LCAO result; thus combining these approximations leads to a result very close to the LCAO value. We will therefore approximate the E_1 energy using the McWEDA correction for both crystal field and hopping terms. We note that to improve upon this would require a better description of both crystal field and hopping terms simultaneously; improving the description of either in isolation will only lead to useful error cancellations being lost.

While we use the McWEDA approximation for both crystal field and hopping integrals in this work, the greater accuracy of the SN approximation for crystal field terms suggests

Property	LCAO	LCAO_scf0	LCAO_scf1	LCAO_scf2	scf-LCAO
Basal octahedral E_F (eV)	1.13	1.50	1.90	2.07	2.61
Basal tetrahedral E_F (eV)	2.86	2.35	2.72	3.14	4.11
Crowdion E_F (eV)	2.08	1.28	2.11	1.88	3.00
Octahedral E_F (eV)	1.97	2.44	2.53	3.22	2.71
Split E_F (eV)	1.28	2.70	3.05	2.78	2.45
Average self-interstitial E_F error (%)	39	33	24	23	0
Hcp0001 surface energy (J m^{-2})	1.60	8.50	-1.64	9.70	2.23
Stable I_1 stacking fault (mJ m^{-2})	85	1036	-1496	305	101
Stable T_2 stacking fault (mJ m^{-2})	220	845	424	-30	249
Unstable I_2 stacking fault (mJ m^{-2})	-	905	362	394	180

TABLE IV. Effects of various E_2 approximations on selected properties for Zr. All energies except the hcp0001 surface energy and unstable I_2 stacking fault are for the relaxed structures. scf-LCAO includes all self-consistency effects; the accuracy of other methods can be judged by their agreement with this method. The value "-" for the unstable I_2 stacking fault with the LCAO method indicates that the value could not be calculated, due to absence of a maximum in a reasonable position on the relevant displacement curve.

it may be superior for models which use different approximations for the hopping integrals. It is likely to be particularly useful for systems where two-centre hopping integrals are sufficient, but many-centre interactions still play an important role. One possible example is organic materials under pressure; reasonable results have been obtained for these systems using environmentally screened or fitted hopping integrals.[3, 7, 29, 30] Addition of explicit, fitted many-centre crystal field corrections has previously been found to improve results for these systems suggesting these interactions are important.[7, 30] Results in the current paper suggest the SN correction may offer a simpler method (requiring no fitting) to account for these interactions in similar systems.

C. Approximating E_2

We have previously found that the E_2 term (*i.e.* self-consistency) is highly important in describing properties such as self-interstitial energies in Zr but not Mg.[1] In the current work, we have also found that unstable stacking fault energies are highly sensitive to self-consistency effects for Zr but not Mg. We will therefore focus on Zr in this section, as $E_2 = 0$ appears to be a sufficiently accurate approximation for Mg.

Large self-consistency effects suggest that the reference electronic density (built from linear combinations of atomic densities) can be far from the true, relaxed electron density for a system. Thus, we tested effects of using different atomic densities when building the reference density; the methodology and results are described in full in the ESI,[23] but we will briefly summarise the main points here. We built our varying atomic densities from atomic calculations with different enforced orbital occupancies; d^2s^2 (which we use for all results in the main paper), d^3s^1 and d^4s^0 configurations were used. We note in passing that the choice of atomic configuration also has an effect on the basis set: see the ESI for more information. For the I_2 stacking fault along the [10-10] displacement pathway we found that errors in stacking fault energies (but not absolute energies) from neglecting self-consistency increased significantly in the order $d^4s^0 < d^3s^1 < d^2s^2$, suggesting that using the d^4s^0 configuration would be the best way to reduce stacking-fault errors in our models. In contrast, errors from self-consistency in energy-volume curves increased in the opposite order, $d^2s^2 < d^3s^1 < d^4s^0$. Thus, perfect crystals appear to be best described as a linear combination of Zr d^2s^2 atomic densities. In terms of Zr model development, these results show that an approximation for self-consistency (rather than a change in reference atomic density) is required for an accurate description of both energy-volume curves and stacking faults using the same model.

Table IV shows the effect of various E_2 approximations (including LCAO, with E_2=0) on Zr properties sensitive to self-consistency. The simplest, explicit approximation (LCAO_scf0) involves treating atomic charge density moments as spherical and is very similar to the standard SCC-DFTB approach to self-consistency introduced in reference [31]. The scf0 correction leads to a slight increase in the overall accuracy for self-interstitial energies, though it leads to E_F shifting in the wrong direction for 2 out of 5 cases. For all other properties the scf0 approximation leads to a shift in the correct direction but, generally, with a magnitude

that is too large. The inclusion of dipole (scf1) and quadrupole (scf2) terms to describe changes in charge densities leads to further large shifts, compared to scf0, in defect formation energies. This demonstrates that the use of charge monopoles is too restrictive to reasonably describe the environmental-dependent changes in atomic charge density for Zr. However, while the scf1 and scf2 corrections improve the accuracy (relative to scf0) of many properties they also lead to unphysically negative stacking fault formation energies.

The tested approximations to E_2 were found to be insufficient to accurately describe Zr properties sensitive to self-consistency effects. Nonetheless, the use of the scf0 approximation, by overcorrecting for scf effects, led to all defect formation energies being positive, in contrast to the LCAO method and the use of higher order multipole terms. Therefore we will use the scf0 approximation for our final Zr model. While this will not lead to high quantitative accuracy for properties sensitive to self-consistency, it is at least expected to lead to a stable model.

D. The final TB models

We build our final Mg and Zr models using the same corrections for E_0 and E_1 terms; E_0 is calculated using a pair potential with an empirically fitted correction plus the zeroth-order uniform density approximation while E_1 terms are calculated using the McWEDA method for both hopping and crystal field terms. We completely neglect E_2 contributions (*i.e.* self-consistency) for the Mg model, while we use the scf0 correction for the Zr model. Thus, these models can be labelled as "H_Mc_XT_Mc_PP_fit_uden0" for Mg and "H_Mc_XT_Mc_PP_fit_uden0_scf0" for Zr. Note that, while we add the scf0 correction for Zr, the pair-potential was fit such that the "H_Mc_XT_Mc_PP_fit_uden0" model reproduced the relevant results (described in more detail in Section II F).

We report slightly different properties for Zr and Mg, based on the desired use cases. For Zr we report results for a wider range of self-interstitial structures, while for Mg we report multiple surface energies (compared to one for Zr). Furthermore, for Mg we report results for the model before and after fitting a pair-potential in the main paper, while this is deferred to the ESI[23] for Zr so that differences in the model with and without the scf0 approximation can be discussed more clearly.

Property	Final model (pre-PP fit)	Final model (post-PP fit)	LCAO	scf-LCAO	Plane-wave
hcp V_0 (a_0^3 per atom)	157.1	154.5	155.9	158.7	154.5
hcp B_0 (GPa)	41.3	36.5	41.8	37.7	36.0
Vacancy E_F (eV)	0.80	0.46	1.04	1.03	0.78
Unrelaxed octahedral interstitial E_F (eV)	6.76	6.96	6.58	6.73	6.17
Octahedral interstitial E_F (eV)	2.44	2.14	2.33	2.35	2.25
Split interstitial E_F (eV)	2.50	2.36	2.37	2.36	2.33
I_1 stable stacking fault (mJ m^{-2})	10.6	11.1	18.4	21.3	20.4
I_2 stable stacking fault (mJ m^{-2})	12.2	15.2	27.8	34.3	33.0
T_2 stable stacking fault (mJ m^{-2})	26.7	26.7	39.1	46.6	42.2
I_2 unstable stacking fault (mJ m^{-2})	108	96	101	106	105
(0001) surface energy (J m^{-2})	0.75	0.60	0.84	0.84	0.55
$(10\text{-}10)_s$ surface energy (J m^{-2})	0.83	0.66	0.91	0.93	0.62
$(10\text{-}10)_l$ surface energy (J m^{-2})	1.15	0.95	1.27	1.31	0.88

TABLE V. Selected properties calculated with our final Mg model compared to reference values. The final model (pre-PP fit) corresponds to the H_Mc_XT_Mc_PP_uden0 method, the final model (post-PP fit) corresponds to the H_Mc_XT_Mc_PP_fit_uden0 method. The unstable stacking fault energy is calculated using unrelaxed structures. Self-interstitial, vacancy, stable stacking fault and surface energies are calculated with structures relaxed at the plane-wave level unless otherwise stated. $(10\text{-}10)_s$ and $(10\text{-}10)_l$ refer to the "short" and "long" terminated surfaces, see ESI[23] for more details.

1. Mg

Figure 4 shows the DoS for hcp Mg using our final model compared to plane-wave results. We find an excellent agreement between tight binding and plane-wave results, with only slight differences near the Fermi energy. The good agreement up to 1 eV above the Fermi energy is especially encouraging, as this suggests that adding charge to the system should lead to the correct extra states being populated. The small remaining errors in the DoS for our model are due to the approximations we make for calculating E_1; this can be deduced

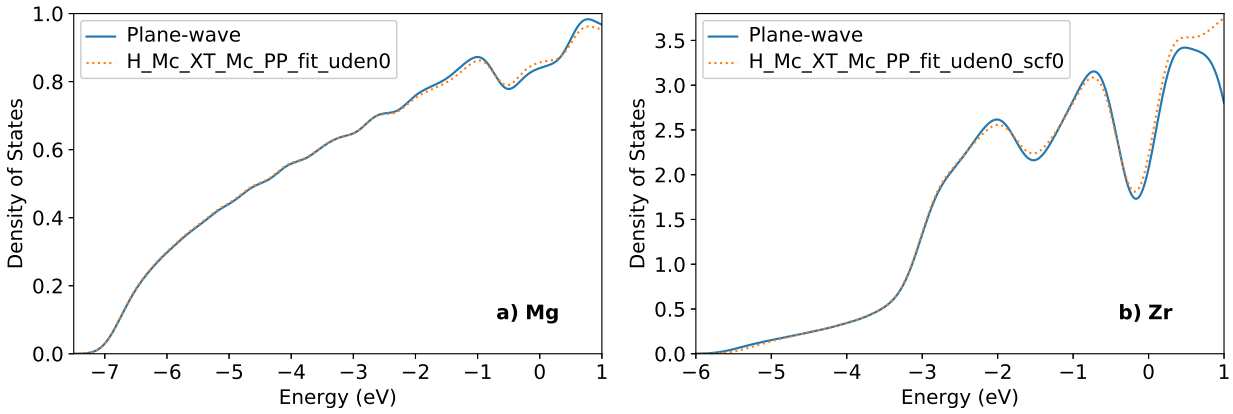


FIG. 4. Density of states for the hcp experimental geometry for a) Mg and b) Zr calculated using our final tight binding models and plane-wave DFT. Note the Fermi level is at 0 eV.

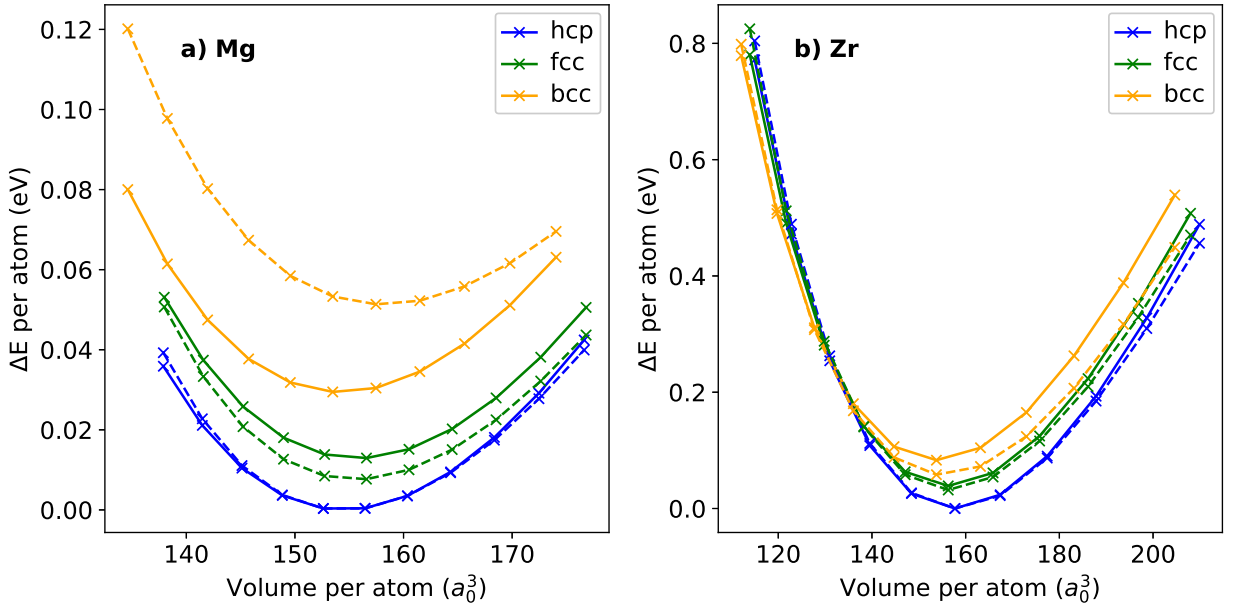


FIG. 5. Energy-volume curves for a) Mg and b) Zr using our final tight binding models (dashed lines) and plane-wave calculations (solid lines). These models correspond to labels H_Mc_XT_Mc_PP_fit_uden0 for Mg and H_Mc_XT_Mc_PP_fit_uden0_scf0 for Zr.

from the slightly improved agreement of plane-wave DoS with the LCAO and scf-LCAO methods, which is shown in the ESI[23]

Figure 5a shows energy vs volume curves for the final Mg model compared to plane-wave results. An excellent agreement between our model and plane-wave results is observed for both the hcp (which was fit to) and fcc (which was not fit to) structures. The agreement

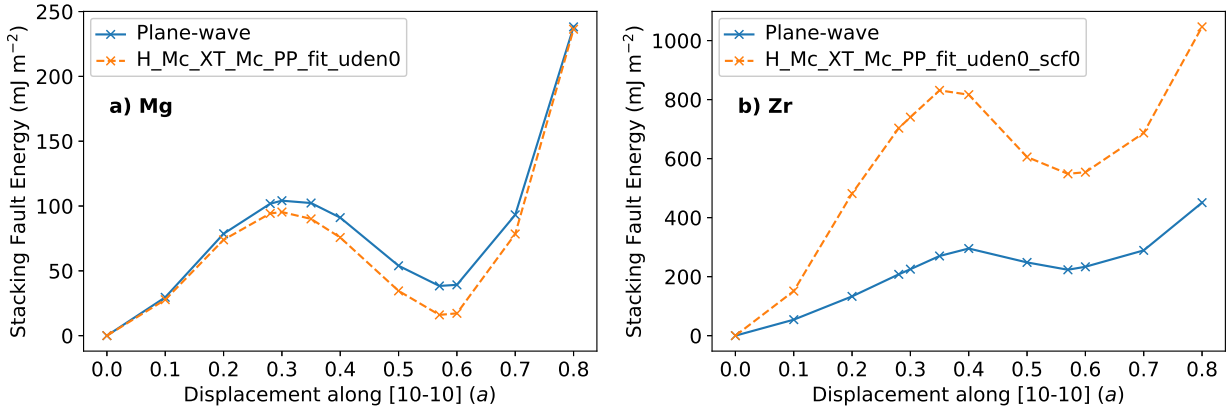


FIG. 6. Stacking fault energies for varying magnitude displacements along the [10-10] direction for the I_2 stacking fault from plane-wave calculations compared to our final models for a) Mg and b) Zr. Displacement magnitude is in terms of the lattice parameter a; a value of ≈1.7 corresponds to a displacement back to the original structure. Note that only the initial structure (displacement=0) was relaxed.

for bcc is significantly worse, though still reasonable. We note that the bcc structure is relatively unimportant for most applications since it will not be the most stable except under conditions of high pressure [32]. The calculated unrelaxed elastic constants (reported in the ESI[23]) show an average error of 16% compared to plane-wave results; this contrasts favorably with both our previous two-centre model (29% error) and a previous model from Gotsis *et al* (38% error for relaxed elastic constants compared to experiment).[33]

Self-interstitials may be expected to present the most difficult test for our models, due to the large number of many-centre effects present which we approximate. Despite this, our final Mg model leads to accurate self-interstitial energies for both split and octahedral sites, even for an unrelaxed structure. Our final model is slightly less successful for a vacancy, with E_F underestimated by ≈ 40%. However, a large portion of this error comes from our fitting of the pair potential. Thus, in applications where this error in vacancy E_F is important it may be possible to use the H_Mc_XT_Mc_PP_uden0 method at the cost of less accurate structural parameters.

The calculated surface energies from our model are in very good agreement with plane-wave results, despite these quantities being overestimated by both LCAO and scf-LCAO methods. Thus, this close quantitative agreement is due to cancellation of errors from the

limited basis set with errors from the approximate treatment of integrals. However, all methods in Table V show approximately the same ratios between the three surface energies, suggesting that the essential differences between interactions in the varying surfaces are captured even without fitting the empirical pair-potential correction. In terms of surface energies, our model contrasts favorably with both embedded atom models (EAMs), which can have difficulties separating the (0001) and (10-10) surface energies and a previous tight binding model where absolute values of surface energies were overestimated by 40-60%.[33–35]

Our Mg model significantly underestimates the absolute values of the basal plane stable stacking fault energies, with values ≈50% smaller than plane-wave results. The majority of this error comes from our approximate treatment of integrals, as can be seen from comparing our final model with the LCAO results which lead to significantly more accurate values. Despite errors in absolute stacking fault energies, our model successfully captures the differences between I_2, I_1 and T_2 stacking faults well; the ratio between I_1,I_2 and T_2 stable stacking fault energies is 1.0:1.4:2.4 for our model compared to 1.0:1.6:2.1 for plane-wave results. Furthermore, Figure 6 demonstrates that our model provides a qualitatively correct description of the I_2 displacement along the [10-10] direction, which is expected to be a good approximation to the minimum energy pathway, with the unstable stacking fault energy within 10% of the plane-wave value. In terms of absolute errors for the I_2 stacking fault our model is similar to multiple others; previous tight binding and EAM models which were not directly fit to the I_2 stable stacking fault generally show errors between 30% and 50%, though lower errors can be obtained with fitting.[33, 36–38] Our models I_2 unstable stacking fault error (<10%) is slightly better than is found when comparing EAM directly to DFT results; for example, the model of Pei *et al.* led to an unstable fault of 72 mJm^{-2} compared to 91 mJm^{-2} from DFT calculations under the same conditions.[36, 38] More importantly, our model can reasonably capture differences in the I_1,I_2 and T_2 stacking faults. In contrast, EAM models are expected to lead to a ratio of 1:2:2 for the I_1:I_2:T_2 stable stacking fault energies, as is found for the case of Zr models.[39, 40] In summary, our Mg model provides a good overall description of the stacking faults investigated despite the significant errors in the absolute values of stable stacking fault energies.

Overall, our final Mg model leads to results that are generally close to plane-wave accuracy for a wide range of properties. This is despite only fitting a pairwise potential to hcp

Property	Final model (no scf)	Final model (scf0)	LCAO	scf-LCAO	Plane-wave
hcp V_0 (a_0^3 per atom)	158.1	158.1	156.9	155.3	157.8
hcp B_0 (GPa)	88.6	88.6	90.7	99.4	90.7
Vacancy E_F (eV)	1.32	1.31	2.42	2.48	2.03
Unrelaxed octahedral interstitial E_F (eV)	3.99	10.60	3.53	6.96	8.86
Octahedral interstitial E_F (eV)	1.93	2.45	1.97	2.71	3.02
Split interstitial E_F (eV)	1.53	3.14	1.28	2.45	3.24
Crowdion interstitial E_F (eV)	2.05	1.24	2.08	3.00	3.49
Basal tetrahedral interstitial E_F (eV)	2.81	2.38	2.86	4.11	4.57
Basal octahedral interstitial E_F (eV)	1.36	1.76	1.13	2.61	3.09
I_1 stable stacking fault (mJ m^{-2})	94.1	995.9	84.6	101.3	140.4
I_2 stable stacking fault (mJ m^{-2})	134.5	569.0	132.9	150.6	211.6
T_2 stable stacking fault (mJ m^{-2})	235.5	806.8	219.9	249.2	325.3
I_2 unstable stacking fault (mJ m^{-2})	-	831	-	180	295
(0001) Unrelaxed surface energy (J m^{-2})	1.14	6.93	1.60	2.23	1.70

TABLE VI. Selected properties calculated with our final Zr model compared to reference values. The final model (no scf) corresponds to the H_Mc_XT_Mc_PP_fit_uden0 method, the final model (scf0) corresponds to the H_Mc_XT_Mc_PP_fit_uden0_scf0 method. Self-interstitial, vacancy and stable stacking fault energies are calculated with structures relaxed at the plane-wave level unless otherwise stated. The unstable stacking fault energy is calculated using unrelaxed structures. The value "-" for the unstable stacking fault indicates that no maximum was found in the expected range of displacements.

structural properties and an LCAO dimer dissociation energy curve. We therefore expect this model to have good transferability, due to the minimal amount of fitting combined with the model accurately describing a wide range of properties.

2. Zr

Figure 4 shows the electronic structure for Zr in our final model compared to plane-wave results. A good agreement is seen between the DoS from plane-wave calculations and our model up to, and slightly above, the Fermi level. The small remaining errors are due to the use of a finite basis set, as is demonstrated by the similar agreement between scf-LCAO and plane-wave results (shown in the ESI[23]).

Figure 5b shows that our final model accurately captures energy-volume curves for Zr: this is expected as these properties were included in the fitting process. While the bcc structure is slightly stabilised relative to hcp by our model, the agreement with plane-wave results is better than for the LCAO method. In particular, the change from hcp to bcc as the ground state happens at a similar volume per atom for plane wave and tight-binding results ($\approx 135\ a_0^3$ and $\approx 130\ a_0^3$ for tight binding and plane wave respectively, see ESI[23]). The unrelaxed elastic constants show an average error of 17% compared to plane-wave results; this is a similar, though slightly smaller, error compared to our previous two-centre model (25% error) and a previous Zr model from Schnell *et al* (23% error in relaxed elastic constants compared to experiment).[41] Overall, our Zr tight binding model appears to accurately capture near-equilibrium structural properties well.

Our Zr model is not able to accurately reproduce plane-wave self-interstitial energies; for example, the E_F for the crowdion interstitial is approximately a third of the plane-wave value. Part of these errors can be attributed to finite basis set effects, as is shown by scf-LCAO E_F being consistently smaller than the plane-wave values. This underestimation of E_F can be rationalised by the dissociation energy curve of Zr-Zr, which has a minimum at $\approx 3\ a_0$ for scf-LCAO compared to $\approx 4.5\ a_0$ for plane-wave calculations (see ESI[23]). Errors in these short-ranged interactions likely result from the inadequate pseudopotential, but it is unclear whether this is due to the pseudopotential being optimised for LDA or semi-core states being significantly perturbed. Regardless, the majority of the error in calculating self-interstitial energies with our model comes from our approximate description of self-consistency. This can be seen by comparing results for LCAO with those for our model without the scf0 approximation, whereby differences are small. Thus, our model accurately captures the non self-consistent terms in the Harris-Foulkes functional, but the low accuracy of our self-consistency correction means the model cannot accurately describe self-interstitial

structures.

Figure 6 shows the I_2 stacking fault energies for displacements along the [10-10] direction for our final model. This leads to the correct shaped curve, but stacking fault energies are approximately three times higher than plane-wave results. Similarly, the T_2 and I_1 stable stacking faults are also overestimated by our model. As for self-interstitial E_F, the majority of these errors come from inaccuracies in the scf0 approximation; this can be seen from the good agreement between LCAO results and our model without the scf0 correction. Basis set errors also cause a reduction in the absolute values of stable stacking fault energies, as can be seen from differences in scf-LCAO and plane-wave results. However, the scf-LCAO method (which contains only basis-set errors) still provides the correct ordering of stable stacking fault energies ($I_1 < I_2 < T_2$) and contains a maximum in a reasonable position for the I_2 [10-10] displacement curve.

In summary, our Zr model provides a good description of electronic and structural properties for near-equilibrium geometries but cannot accurately describe self-interstitial E_F or stacking faults. We have shown that these remaining errors come from the inaccuracy of the scf0 correction for self-consistency. Thus, use of a more accurate self-consistency approximation is required in order to yield a model capable of describing stacking faults and self-interstitial E_F with high accuracy.

IV. CONCLUSIONS

We have presented a methodology for building tight binding models capable of describing systems containing large numbers of many-centre interactions. Our models contain three main approximations to plane-wave DFT calculations: i) a smaller basis set, ii) approximate self-consistency and iii) approximate many-centre exchange-correlation. We have quantified the errors introduced by each of these approximations, with the main focus being on isolating errors from a range of many-centre exchange-correlation approximations. This quantification of individual errors allowed us to combine approximations in the most pragmatic manner, to understand the origins of remaining errors and should help others to judge which approximations may be useful in their own model development.

We were able to find the most accurate approximations for E_0 (uniform density to zeroth-order), crystal-field (the SN correction) and hopping terms (the McWEDA correction) by

looking at each approximation individually. Understanding errors from individual approximations allowed us to combine them in a pragmatic manner such that error-cancellation between approximations could be exploited. This is demonstrated by our use of the McWEDA correction for crystal-field integrals, despite the SN method being more accurate, due to favorable error cancellation found between McWEDA crystal-field and hopping terms. While this combination of approximations is likely to be useful for systems with large many-centre effects, other combinations may be more useful in more open systems. Of the approximations we tested in this work, only the McWEDA hopping and second-order uniform density approximations required up to three-centre terms to be included. Thus, any other approximation could be included in two-centre *ab initio* models without adding significant computational expense. Our quantification of the accuracy of each individual approximation indicates that the SN correction for crystal field terms and the uniform density approximation for E_0 could be inexpensive additions to improve accuracy and transferability of models where the use of optimised two-centre hopping integrals is sufficient; for example models for studying organic materials or biomolecules.[3, 42, 43]

Standard self-consistency approximations, with changes in the charge density represented as a multipole expansion, were found to be insufficient to accurately describe the large self-consistency effects for Zr. We found the results varied significantly between scf0,scf1 and scf2 methods, which suggests a large degree of error likely comes from these expansions being too inflexible to accurately describe the environmental-dependent changes in charge densities. It is also possible that some of the remaining errors come from neglect of exchange-correlation contributions to E_2. Regardless, we have found that more refined approximations to E_2 are necessary to account for self-consistency for describing defects in metallic Zr; there are various ways in which this might be accomplished. For example, the use of multiple Gaussian functions to represent the radial dependence of each multipole moment may lead to a sufficiently flexible basis for describing changes in charge density; this is similar to the use of multiple Hubbard parameters proposed in reference [44], where it was suggested that this may be particularly important for systems with d-electrons. One obvious way to implement this would be to have a charge distribution that has the right shape for the s-orbitals, and a second one with the correct shape for the d-orbitals; this would allow the model to describe electron transfer between these orbitals. A second possibility would be to include corrections for anisotropic exchange-correlation contribution to E_2, as is used in

reference [45]. A different approach would be to add fitted integrals or an empirical potential to try to reproduce plane-wave results for properties sensitive to self-consistency, similar to the approach taken for many empirical tight binding models which involve fitting to self-consistent DFT calculations (and therefore implicitly include self-consistency effects).[5, 8, 9, 46] Applying this to our model would have the advantage of being more transparent than previously used implicit methods, but it is unclear how well an implicit treatment could handle large self-consistency effects.

We have presented a Zr model which provides a good description of electronic structure and structural properties for near-equilibrium structures. However, we were not able to create a model which provides an accurate description of defect energies due to the large self-consistency effects in Zr. Nonetheless, our final model (without the scf0 correction) leads to good agreement with LCAO results which demonstrates that we are able to capture the many-centre contributions to E_0 and E_1 accurately. More refined corrections for self-consistency will need to be added to the model in order to approximate the plane-wave results well; the model in this paper (without the scf0 correction) provides a strong starting point for this.

Our final Mg model accurately reproduces results close to plane-wave values for a range of electronic, structural and defect properties; this includes self-interstitial energies,surface energies and equilibrium lattice volumes. This was accomplished with only a minimal amount of fitting, with even the unfitted version of the final model providing reasonably accurate results. This high accuracy for a wide range of properties demonstrates the transferability of this model and gives us confidence that it captures the essential aspects of Mg-Mg interactions in the metallic phase.

ACKNOWLEDGMENTS

We thank Adrian Sutton for many important conversations on TB models. This work has in part been performed within the framework of the international MUZIC (Mechanistic Understanding of Zirconium Corrosion) program. The authors gratefully acknowledge the industrial support from EDF, EPRI, Naval Nuclear Laboratory, Rolls-Royce, Westinghouse and Wood. R.M.F. gratefully acknowledges funding from EPSRC under grant EP/R005419/1. J.S. gratefully acknowledges support through a studentship in the Centre

for Doctoral Training on the Theory and Simulation of Materials at Imperial College London, funded by EPSRC under Grant No. EP/L015579/1, and by Rolls-Royce. We acknowledge support from the Thomas Young Centre under grant TYC-101.

---

[1] J. Smutna, R. M. Fogarty, M. R. Wenman, and A. P. Horsfield, Physical Review Materials **4**, 043801 (2020).

[2] A. Krishnapriyan, P. Yang, A. M. N. Niklasson, and M. J. Cawkwell, Journal of Chemical Theory and Computation , acs.jctc.7b00762 (2017).

[3] M. J. Cawkwell and R. Perriot, The Journal of Chemical Physics **150**, 024107 (2019).

[4] T. J. Sheppard, A. Y. Lozovoi, D. L. Pashov, J. J. Kohanoff, and A. T. Paxton, Journal of Chemical Physics **141** (2014), 10.1063/1.4887095.

[5] M. J. Mehl and D. A. Papaconstantopoulos, Physical Review B - Condensed Matter and Materials Physics **54**, 4519 (1996), arXiv:9602003 [mtrl-th].

[6] S. Silayi, D. A. Papaconstantopoulos, and M. J. Mehl, Computational Materials Science **146**, 278 (2018).

[7] N. Goldman, Chemical Physics Letters **622**, 128 (2015).

[8] P. F. Li and B. C. Pan, Journal of Physics Condensed Matter **24** (2012), 10.1088/0953-8984/24/30/305802, arXiv:9410002 [chem-ph].

[9] M. Tang, C. Wang, and C. Chan, Physical Review B - Condensed Matter and Materials Physics **53**, 979 (1996).

[10] J. P. Lewis, P. Jelínek, J. Ortega, A. A. Demkov, D. G. Trabada, B. Haycock, H. Wang, G. Adams, J. K. Tomfohr, E. Abad, H. Wang, and D. A. Drabold, Physica Status Solidi (B) Basic Research **248**, 1989 (2011).

[11] P. Jelínek, H. Wang, J. P. Lewis, O. F. Sankey, and J. Ortega, Physical Review B - Condensed Matter and Materials Physics **71**, 1 (2005), arXiv:0409509v2 [arXiv:cond-mat].

[12] A. P. Horsfield, Physical Review B **56**, 6594 (1997).

[13] O. F. Sankey and D. J. Niklewski, Physical Review B **40**, 3979 (1989).

[14] N. Li and Y. Zheng, Journal of Materials Science and Technology **29**, 489 (2013).

[15] T. B. Abbott, CORROSION **71**, 120 (2015).

[16] M. Esmaily, J. E. Svensson, S. Fajardo, N. Birbilis, G. S. Frankel, S. Virtanen, R. Arrabal,

S. Thomas, and L. G. Johansson, Progress in Materials Science **89**, 92 (2017).

[17] A. T. Motta, A. Couet, and R. J. Comstock, Annual Review of Materials Research **45**, 311 (2015).

[18] J. Harris, Phys. Rev. B **31**, 1770 (1985).

[19] W. M. C. Foulkes and R. Haydock, Phys. Rev. B **39**, 12520 (1989).

[20] P. Hohenberg and W. Kohn, Physical review **136**, B864 (1964).

[21] M. Boleininger, A. A. Guilbert, and A. P. Horsfield, The Journal of Chemical Physics **145**, 144103 (2016).

[22] P. Soin, A. P. Horsfield, and D. Nguyen-Manh, Computer Physics Communications **182**, 1350 (2011).

[23] See Supplemental Material [url] for additional methodological details and supporting results, which includes Refs. [47].

[24] S. D. Kenny and A. P. Horsfield, Computer Physics Communications **180**, 2616 (2009).

[25] S. J. Clark, D. S. Matthew, J. P. Chris, J. H. Phil, I. J. P. Matt, R. Keith, and C. P. Mike, Zeitschrift fur Kristallographie **220**, 567 (2005).

[26] J. Perdew, K. Burke, and M. Ernzerhof, Physical review letters Phys. Rev. Lett. (USA), **77**, 3865 (1996).

[27] C. Hartwigsen, S. Goedecker, and J. Hutter, Physical Review B **58**, 3641 (1998).

[28] M. Krack, Theoretical Chemistry Accounts **114**, 145 (2005).

[29] N. Goldman, S. Goverapet Srinivasan, S. Hamel, L. E. Fried, M. Gaus, and M. Elstner, Journal of Physical Chemistry C **117**, 7885 (2013).

[30] W.-C. Lu, C. Z. Wang, L.-Z. Zhao, W. Qin, and K. M. Ho, Physical Review B **92**, 035206 (2015).

[31] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, Physical Review B **58**, 7260 (1998).

[32] H. Olijnyk and W. B. Holzapfel, Physical Review B **31**, 4682 (1985).

[33] H. J. Gotsis, D. A. Papaconstantopoulos, and M. J. Mehl, Physical Review B **65**, 134101 (2002).

[34] Y. M. Kim, N. J. Kim, and B. J. Lee, Calphad: Computer Coupling of Phase Diagrams and Thermochemistry **33**, 650 (2009).

[35] Z. Wu, M. F. Francis, and W. A. Curtin, Modelling and Simulation in Materials Science and

Engineering **23** (2015), 10.1088/0965-0393/23/1/015004.

[36] J. A. Yasi, T. Nogaret, D. R. Trinkle, Y. Qi, L. G. Hector, and W. A. Curtin, Modelling and Simulation in Materials Science and Engineering **17** (2009), 10.1088/0965-0393/17/5/055012.

[37] W. Hu, B. Zhang, B. Huang, F. Gao, and D. J. Bacon, Journal of Physics Condensed Matter **13**, 1193 (2001).

[38] Z. Pei, L. F. Zhu, M. Friák, S. Sandlöbes, J. Von Pezold, H. W. Sheng, C. P. Race, S. Zaefferer, B. Svendsen, D. Raabe, and J. Neugebauer, New Journal of Physics **15** (2013), 10.1088/1367-2630/15/4/043020.

[39] A. Poty, J. M. Raulot, H. Xu, J. Bai, C. Schuman, J. S. Lecomte, M. J. Philippe, and C. Esling, Journal of Applied Physics **110** (2011), 10.1063/1.3599870.

[40] C. Varvenne, O. Mackain, and E. Clouet, Acta Materialia **78**, 65 (2014), arXiv:1407.5154.

[41] I. Schnell, M. D. Jones, S. P. Rudin, and R. C. Albers, Phys. Rev. B **74**, 054104 (2006).

[42] M. Gaus, A. Goez, and M. Elstner, Journal of Chemical Theory and Computation **9**, 338 (2013).

[43] M. Gaus, Q. Cui, and M. Elstner, Wiley Interdisciplinary Reviews: Computational Molecular Science **4**, 49 (2014).

[44] Z. Bodrog and B. Aradi, Physica Status Solidi (B) Basic Research **249**, 259 (2012).

[45] C. Bannwarth, S. Ehlert, and S. Grimme, Journal of Chemical Theory and Computation **15**, 1652 (2019).

[46] C. E. Lekka, N. Bernstein, D. A. Papaconstantopoulos, and M. J. Mehl, Materials Science and Engineering B: Solid-State Materials for Advanced Technology **163**, 8 (2009).

[47] M. Gunde, L. Martin-Samos, S. De Gironcoli, M. Fanetti, D. Orlov, and M. Valant, Physical Review B **100**, 1 (2019).