1. Introduction

The Kohn–Sham formulation [1] of density functional theory (DFT) [2] is widely used to model the electronic structures of bulk solids, surfaces, and molecules [3–5]. With a reasonable trade-off between accuracy and computational complexity, the possibility to treat the electronic structures of solids as Bloch states of the Kohn–Sham single-particle-like Hamiltonian [6] renders DFT and extensions the de-facto standard for electronic structure simulation in computational materials science [7–11].

Generalized gradient approximations (GGA) and meta-GGAs to electronic exchange and correlation (XC) in Kohn–Sham DFT only depend locally on the charge density, its gradient and Laplacian, and the Kohn–Sham kinetic energy density and yield reasonable cohesive energies with typical errors in comparison to experiment of <0.2 eV per atom for those solids [12, 13] where self-interaction errors due to these semi-local approximations are not too large. In materials with strong electron localization, these errors become large, and semi-local DFT thus typically yields significantly larger per atom errors for relative stabilities of e.g. transition metal (TM) oxide phases [14]. The failure of DFT also manifests itself in a qualitatively wrong description of the electronic structure of Mott insulators, where the semi-local, mean-field treatment of XC incorrectly yields a metallic ground state (see figure 1).

One approach to reducing self-interaction errors is to admix a fraction of Hartree–Fock exchange energies computed from the Kohn–Sham orbitals to the semi-local XC functional [15]. The computational cost for evaluating the non-local Fock-operator generally increases the cost for this hybrid DFT method by a factor of the
system size in comparison to semi-local DFT. In contrast to unscreened Hartree–Fock, which unphysically predicts diverging Fermi velocity and vanishing density of states at the Fermi level for the homogeneous electron gas [16], screened hybrid DFT [17] can treat metallic phases. Magnetic metals are however described worse, e.g., with the widely used HSE06 screened hybrid functional [18] than with semi-local DFT, with overestimated magnetic moments of Fe, Co, and Ni [19, 20], incorrectly predicting bcc and hcp ground state structures for Co and Ni [21], respectively, and describing surface reaction energetics worse in particular for magnetic TM surfaces [22]. HSE06 simulations were nevertheless shown to improve on transition metal oxide formation energies over GGA-DFT [23], showing the importance of the reduction of self-interaction errors due to localized electrons in the oxides.

Another, computationally more efficient approach to addressing the self-interaction errors is the GGA+U approach [24–26], which does not lead to an overall increase in computational complexity over semi-local DFT. In the GGA+U method, the Kohn–Sham Hamiltonian is supplemented with a Hubbard-like term, which depends on the local projections of the Kohn–Sham single-particle density matrix at ionic sites with strong electron localization. The Hubbard term depends thus explicitly on the Kohn–Sham occupation numbers and can e.g. open up a particle–hole excitation gap in the case of Mott insulators. The strength of the Hubbard term is given by the parameter U, which can either be considered as an adjustable, to be fitted parameter or be computed from first principles [27–29]. In recent approaches, machine learning techniques were employed to find U-parameters to mimic hybrid DFT bandstructures and bandgaps with GGA+U [30] or to improve agreement with experimental lattice constants, magnetic moments, and bandgaps [31].

A drawback of the GGA+U method is that the ground state energies of Hamiltonians with different Hubbard parameters U at a given ionic site cannot be compared to each other. The compromise of choosing average U-parameters for each transition metal ionic species (irrespective of e.g. differences in oxidation state) leads to an improvement of oxidation energies of transition metal oxides over GGA-DFT [14, 32, 33]. For molecules, meaningful total energy differences in the case of first-principles, reactant-dependent U-values can be computed by integrating the forces acting on the ions along a reaction path [34]. For the prediction of oxidation energies of bulk TM oxides with first-principles U-parameters, the unit cell volume-dependence of the U-parameters can be computed, and integrals over the volume changes required to match the U-values of different oxides can then be used to estimate reaction energies with the oxides at their respective equilibrium volumes [35]. Oxidation energies were found to be further improved using this method rather than using average U-parameters, while oxide formation enthalpies could not be improved [35]. A general problem for either an average U-parameter approach or relying on volume change integrals is that first-principles methods to computing U yield nonzero values also for metallic phases. Delocalized metallic states are however better described without supplemental Hubbard terms (see figure 1 for an example of the unphysical effects due to application of a U-term to a metallic bandstructure).

**Figure 1.** Bandstructures and densities of states of Fe (top) and V$_2$O$_3$ (bottom) computed using GGA and GGA+U. $U_\text{F}$ is the system-specific Hubbard parameter obtained from linear response ($U_\text{F} \approx 7.2$ eV and $U_\text{F} \approx 6.9$ eV; details in main text). Already at half this parameter strength, Fe shows an incorrect suppression of majority spin density of states at the Fermi level $E_F$, and tendency towards spin saturation (the magnetic moment is increased by about 30% compared to the U = 0 case). At full $U_F$ strength, the magnetic moment increase is 40%. For V$_2$O$_3$, on the other hand, GGA incorrectly predicts a ferromagnetic metallic phase as the ground state (bandstructure plot folded into the smaller irreducible Brillouin zone of the antiferromagnetic phase), while GGA+U correctly predicts an antiferromagnetic Mott-insulating phase.
Improvement over TM oxide formation energy predictions from screened hybrid DFT was achieved by treating the oxide phases at the GGA\(+U\) level with average $U$-parameters for the TM species and the metallic phases within the GGA [36]. For each Hubbard-corrected TM species, an energy offset was fitted to experimental formation enthalpies, so that GGA\(+U\) total energies plus offset can be compared to the GGA total energies for the metallic phases. These fitted offsets allow the resulting corrected total energy differences between GGA\(+U\) and GGA calculations to be interpreted as reaction energies. In a similar approach, elemental reference energies were fitted directly against experimental formation enthalpies with TM oxides treated at the GGA\(+U\) level with average $U$-parameters, and formation energies predicted with this approach were used to fit a multivariate linear regression model of formation energies exclusively against atomic properties [38]. With the availability of sufficient experimental reference data, the approach of combining GGA and GGA\(+U\) total energies can be further refined by fitting offset energies not only for different TM species, but also for different oxidation states and different ligand coordination [39]. While experimental calorimetry reference data for thermodynamics of stoichiometric bulk TM oxides can be used here to parametrize these element- or even system-specific corrections, determining the required offset energies for Hubbard-corrected TM sites near e.g. point defects or interfaces will be difficult.

Therefore, a newly developed method free of adjustable parameters is presented here that combines the treatment of $d$-electrons of varying degrees of localization within the GGA\(+U\) approach with site-specific, first-principles $U$-parameters from linear response for TM oxides with GGA calculations optimal for metallic phases. To correct for the offset between GGA\(+U\) total energies with different $U$-parameters and GGA (i.e. $U = 0$) total energies, genetic programming with experimental TM oxide formation energies as references is applied to find a mathematical model that does not explicitly depend on the Hubbard-corrected TM species. The resulting, relatively simple mathematical model does thus not have adjustable, system-specific parameters. Examples of the applicability of the model for TM oxides with Hubbard-corrected TM species not contained in the training data are shown.

2. Method

The method consists of computing DFT and DFT\(+U\) total energies for metal and oxide phases, respectively, as described below and using a genetic programming model for finding an energy offset for the Hubbard-corrected oxide total energies, such that physically meaningful energy differences between metals and oxides can be obtained (figure 2).

2.1. DFT calculations

GGA and GGA\(+U\) simulations are performed with the VASP code [40] using plane-wave basis sets with a cut-off energy of 600 eV and representing ionic cores by projector-augmented wave (PAW) [41] frozen-core pseudopotentials [42]. Brillouin zones are sampled with $\Gamma$-point centered, regular grids with a spacing of at most 0.02 Å$^{-1}$, and Gaussian smearing with a width of 0.05 eV is used to determine Kohn–Sham occupation numbers. The widely used PBE functional [43] is employed for the GGA to electronic exchange and correlation.
Forces acting on ions and stress on the unit cells are simultaneously relaxed, with residual forces of less than 0.02 eV/Å per atom and residual pressures of less than 0.02 GPa, respectively.

The total energy of the O₂ molecule is described poorly by GGA approaches, which is typical for multiply bonded molecules [27,44]. Empirical approaches fit the total energy of O₂ to experimental reaction energies, such as oxidation energies of solids [14] or the formation of H₂O [45]. Here, instead of empirically fitting to the experimental formation energy of H₂O, atomization energies as computed with the first-principles Weizmann-4 computational thermochemistry protocol (W4) [46] are referenced. Using the difference of the atomization energies of O₂ and H₂O

\[ E_{\text{atom}}(O_2) - 2E_{\text{atom}}(H_2O) = E(O_2) - 2E(H_2O) + 4E(H), \]

the total energy of O₂ used in this study is computed as

\[ E(O_2) = E_{\text{W4}}(O_2) - 2E_{\text{W4}}(H_2O) + 2E_{\text{GGA}}(H_2O) - 4E_{\text{GGA}}(H). \]

2.2. Hubbard corrections

2.2.1. DFT+U energy functional

In the DFT+U approach, the DFT energy functional \( E_{\text{DFT}}[n(\mathbf{r})] \) (considered here to employ a GGA) of the electron charge density \( n(\mathbf{r}) \) is supplemented with a Hubbard-like term:

\[ E_{\text{DFT}+U}[n(\mathbf{r})], \rho_\ell] = E_{\text{DFT}}[n(\mathbf{r})] + E_{\text{Hub}}[\rho_\ell] - E_{\text{DC}}[\rho_\ell]. \]

\( \rho_\ell \) are projections of the Kohn–Sham density matrix onto atomic-like (i.e., 3d, 4d or 5d) orbitals centered on the transition metal ions \( I \). \( E_{\text{DC}}[\rho_\ell] \) is a term only depending on the total projected spin occupations at each site \( I \) and corrects for partially double-counting Coulomb interactions in \( E_{\text{DFT}}[n(\mathbf{r})] \) and the Hubbard-term \( E_{\text{Hub}}[\rho_\ell] \).

Here, the approach of Anisimov et al [24] and Liechtenstein et al [26] is followed: \( E_{\text{DC}}[\rho_\ell] \) is chosen such that it cancels \( E_{\text{Hub}}[\rho_\ell] \) in case of integer eigenvalues of the projected density matrices \( \rho_\ell \), i.e., in the limit of isolated ions. For \( E_{\text{Hub}}[\rho_\ell] \), the spherically averaged expression introduced by Dudarev et al [17] is chosen:

\[ E_{\text{Hub}}[\rho_\ell] = \sum_I \left( \frac{U_I}{2} N_I^\ell - \frac{J_I}{2} \sum_{\sigma \in \{\uparrow, \downarrow\}} (N_I^\ell)^2 - \frac{U_I}{2} \sum_{\sigma \in \{\uparrow, \downarrow\}} \text{Tr} (\rho_\ell^\sigma)^2 \right), \]

and

\[ E_{\text{DC}}[\rho_\ell] = \sum_I \left( \frac{U_I}{2} N_I^\ell (N_I - 1) - \frac{J_I}{2} \sum_{\sigma \in \{\uparrow, \downarrow\}} N_I^\sigma (N_I^\sigma - 1) \right), \]

such that

\[ E_U[\rho_\ell] = E_{\text{Hub}}[\rho_\ell] - E_{\text{DC}}[\rho_\ell] = \sum_I \left( \frac{U_I}{2} - \frac{J_I}{2} \sum_{\sigma \in \{\uparrow, \downarrow\}} \text{Tr} \rho_\ell^\sigma - \text{Tr} (\rho_\ell^\sigma)^2 \right). \]

\( \text{Tr} \rho_\ell^\sigma = \sum_{m=-\infty}^{\infty} \rho_\ell^{\sigma m} \) is the trace over the orbital degrees of freedom of the projected density matrices \( 2(2\ell + 1) = 5 \) in the case of \( d \)-orbitals. \( \sigma \) denotes the electron spin degree of freedom (spins assumed to be collinear in this study), and \( N_I = \sum_{\sigma} N_I^\sigma = \sum_{\sigma} \text{Tr} \rho_\ell^\sigma \) is the total projected occupation at site \( I \). \( \hat{U}_I \) and \( \hat{J}_I \) are the spherically averaged screened on-site Coulomb repulsion and exchange interactions at site \( I \), respectively. Since equation (6) depends on the difference between these interactions at each site \( I \), it is convenient to introduce an effective Hubbard parameter \( U_I \) with

\[ U_I = \hat{U}_I - \hat{J}_I. \]

The Kohn–Sham system is exposed to the following additional potential due to the Hubbard correction at site \( I \):

\[ V_{\text{Hub}}^{\ell,\sigma} = \frac{\partial E_U}{\partial \rho^{\ell,\sigma m}_m} = U_I \left( \frac{1}{2} \delta_{nm} - \rho^{\ell,\sigma m}_m \right). \]

The Hubbard corrections (6) effectively constitute an energy penalty for non-integer eigenvalues of \( \rho^{\ell,\sigma} \). Considering equation (8) in the eigenbasis of \( \rho^{\ell,\sigma} \), the potential is positive for projected occupation eigenvalues less than 1/2, favoring a lowering of the eigenvalue towards zero. For eigenvalues larger than 1/2, the potential is negative thus favoring increasing the eigenvalue towards one. Accordingly, the Hubbard corrections with their
explicit occupation number dependence can open a Mott-Hubbard gap in the Kohn–Sham spectrum of an otherwise metallic system in a single-particle bandstructure picture.

### 2.2.2. Hubbard-U from first principles

Rather than treating the effective Hubbard parameters $U_I$ as empirical fitting parameters, the linear response approach to determining $U_I$ by Cococcioni and de Gironcoli [28] was implemented here on top of the existing DFT+$U$ code in VASP. DFT total energies spuriously show a non-linear dependence on fractional electron numbers (fractional particle numbers correspond to quantum mechanical ensemble averages, and the total energies should show linear behavior between pairs of adjacent integer particle numbers). The $U_I$ can be chosen to compensate spurious curvature with respect to changes in the Hubbard site occupations $N_I$. Instead of constraining Hubbard site occupations to determine these curvatures (which is difficult with a plane-wave basis set), the linear response approach adds perturbations to the potentials (8): $V_{n t, \sigma}^{I, \alpha} = V_{n t, \sigma}^{I, 0} + \alpha \delta_{m, m'}$. The functional (3) is Legendre transformed such that the energy is minimized with respect to the perturbation $\alpha$ instead of the Hubbard-site occupation numbers, and the Hubbard parameters $U_I$ are then determined as [28]:

$$U_I = (\chi_0^{-1} - \chi^{-1})_I. \tag{9}$$

$\chi_{\alpha} = \partial N_I / \partial \alpha$ is a density response function of the total occupations $N_I$ of each Hubbard site $I$ to perturbations from the $\alpha$ for any of the Hubbard sites $I$ and is computed here with finite differences in $\alpha$ letting the electronic structure self-consistently adjust to the perturbed Hubbard potentials. Not letting the Kohn–Sham potentials self-consistently adjust to the perturbed Hubbard potentials, but only computing the Kohn–Sham bands and thus the Kohn–Sham density matrix at the fixed perturbed potentials leads to the perturbed Hubbard site occupations $N_I^0$. The finite difference $\chi_{\alpha I}^0 = \Delta N_I^0 / \Delta \alpha$ is used to compute the corresponding bare response function.

While the updated Kohn–Sham potentials can self-consistently screen the effect of the perturbed Hubbard potentials, the bare response accounts for single-particle bandstructure-only rehybridization between the Hubbard sites and their surroundings. The resulting curvature in total energy due to these latter rehybridizations is not a measure of a self-interaction error to be corrected for with a Hubbard-term, and equation (9) thus removes this curvature from the computation of the $U_I$.

#### 2.2.3. Ab initio U-parameters in practice

In the present implementation, the unperturbed electronic structure is computed first, then the perturbation due to $\alpha_I$ is added, and the Kohn–Sham bands are updated yielding the bare response $\chi_{\alpha I}^0$ for all sites $I$, and finally this same simulation is continued allowing the Kohn–Sham potentials to adjust yielding the response $\chi_{\alpha I}$ for all sites $I$. The Spglib code [48] is used to identify symmetrically equivalent Hubbard sites and the corresponding symmetry operations, so that simulations only need to be performed for perturbations at nonequivalent sites and the full $\chi$ and $\chi^0$ response matrices can be obtained using symmetry operations.

To decouple the effect of the perturbation from its periodic images, the TM oxide unit cells are repeated here into supercells separating the periodic site images by at least $\sim 8 \text{ Å}$. Cococcioni and de Gironcoli [28] find that extending the Hubbard sites by a fictitious background site for charge neutrality of the perturbations improves convergence with respect to supercell size. The response function matrices $\chi$ and $\chi^0$ are thus extended by one row and one column, and the additional matrix elements are determined by the charge neutrality requirement of the sums over each row and each column of the matrices yielding zero. With this construction, the extended response matrices now have a zero singular value and cannot be inverted. This zero singular value is due to the fact that adding a constant offset to all Hubbard site potentials and the background does not lead to any total energy curvature. Here, the Moore–Penrose pseudo inverse is used instead of the matrix inverses in equation (9), such that this zero singular value leads to a zero singular value for the pseudo inverse, correctly not affecting the computed $U_I$.

After the $U_I$ have been computed in supercells, GGA+$U$ optimizations of the primitive cells with respect to atomic forces and unit cell stress are performed. Based on the optimized structures new supercells are created to re-compute the $U_I$ and then to re-optimize the primitive cells again.Performing this relaxation and $U_I$ determination cycle twice led to mean average TM oxide lattice constant changes $< 10^{-2} \text{ Å}$, and thus no further structural-dependence self-consistency cycles of the $U_I$ have been performed.

Generally, the $U_I$ are computed here from GGA + $U$ calculations with all Hubbard parameters set to zero, yielding the spurious energy curvatures to correct for. Some of the considered TM oxides are incorrectly predicted as metallic at the GGA level. For these oxides, GGA + $U$ simulations are performed at a range of Hubbard parameters $U_{\text{input}}$ large enough to open a bandgap, and, following the idea of [49], the computed $U_I$ values are linearly extrapolated to $U_{\text{input}} = 0$ (see Table S1 available online at stacks.iop.org/JPCO/6/035009/mmedia for a list of all computed $U_I$ values, indicating which TM oxide cases are based on this extrapolation technique).
The values of the Hubbard parameters \( U \) determined via equation (9) have a dependence on the choice of the projectors of the Kohn–Sham density matrix on the Hubbard sites. The VASP code uses the projections of the Kohn–Sham pseudo wavefunctions onto the PAW projector functions, which are non-zero only inside the augmentation spheres around the ions [50, 51]. The DFT+U implementation in the Quantum Espresso plane-wave pseudopotential DFT code, e.g., offers further choices for the definition of the projection onto the Hubbard sites, such as the \( d \)-orbitals of free atoms with or without orthogonalization with respect to the atomic orbital overlap between Hubbard sites [52]. Using the density functional perturbation implementation [53] in the Quantum Espresso suite to computing \( \chi \) and \( \chi^0 \) with ultrasoft pseudopotentials [54] from the GBRV dataset [55], the computed \( U \) for bulk Fe is about 7.0 eV and 5.5 eV with and without projector orthogonalization, respectively. Similarly, \( U \)-parameters obtained from constrained local occupations were found to depend strongly on TM ion muffin-tin sphere radii [56]. These dependencies show the importance of determining the parameters \( U \) with the same projectors as are being used in the DFT+U simulation.

Similarly large \( U \)-parameters for bulk Fe based on constrained DFT are reported in [57], and also the VASP PAW implementation (without inter-site projector overlap due to non-overlapping PAW augmentation spheres) finds a large value of about 7.2 eV. Even only applying half this Hubbard parameter strength in a GGA+U simulation of bulk Fe leads to unphysical suppression of majority spin density of states around the Fermi level (see figure 1) and the magnetization is spuriously driven towards saturation. As the corresponding \( d \)-band shifts would deteriorate the description of, e.g., transition metal surface chemistry [22], this example highlights that GGA+U simulations of metallic systems with first-principles \( U \)-parameters should be avoided because of potentially unphysical predictions.

A promising direction for (at least partial) compensation of unphysical effects on metallic electronic structures in DFT+U is the DFT+U+V approach [58], where also inter-site Hubbard corrections, determined from the off-diagonal elements in (9), are applied. The DFT+U+V method is implemented in Quantum Espresso, and for the above orthogonalized projector case with an on-site Hubbard parameter of 7.0 eV, the \( V \)-parameter between nearest neighboring sites is only \( \sim 0.3 \) eV, barely affecting the electronic structure qualitatively and both GGA+U and GGA+U+V calculations yielding a spuriously saturated, integer magnetization of \( \sim 3 \) Bohr magnetons per Fe atom. Generally, the DFT+U+V method can be extended to such off-diagonal Hubbard terms between \( d \)- and \( p \)-projections, also on the same site [58], which could improve suppression of unphysical effects of Hubbard-U on metallic systems, but here only on-site Hubbard-corrections based on \( d \)-projections are considered and implemented. All results presented in the following treat metallic phases without Hubbard corrections.

2.3. Reference data

2.3.1. TM oxide and metallic phases

To find a model enabling the energetic comparison of products and reactants with first-principles \( U \)-parameters rather than average \( U \)-values, experimental formation enthalpies of transition metal oxides requiring Hubbard corrections are used as training data with corrections as described below. The formation enthalpies of 66 such TM oxides containing the TM ions V, Cr, Mn, Fe, Co, Ni, and Mo are taken from [59]. A Debye model fitting technique introduced and benchmarked in [60] is used to determine \( T = 0 \) K enthalpies from the experimental values at 298 K. Neglecting thermal expansion of the solid phases, the heat capacity at constant pressure

\[
c_p(T) = y \cdot \left( \frac{T}{T_D} \right)^3 \int_0^{T/T_D} \frac{u^4 \exp(u)}{\{1 - \exp(u)\}^2} \, du
\]

and the entropy

\[
S(T) = \int_0^T c_p(T') \, dT'
\]

are fitted numerically to match the \( T = 298 \) K entropies and heat capacities provided in [59] (\( c_p(298 \) K) and \( S(298 \) K) for ZnFe\(_2\)O\(_4\) taken from [61]) due to unphysical \( c_p(298 \) K) in [59]), both for the TM oxides and the metallic reference phases, considering the Debye temperature \( T_D \) and the pre-factor \( y \) as fitting parameters.

Here, the fitted \( T_D \) is furthermore employed to provide an experimental estimate of the phonon zero-point energy \( E_{ZP} \) per formula unit using that the Debye model density of states is quadratic in the phonon frequency \( \omega \) below the Debye cut-off frequency \( \omega_D = k_B T_D / \hbar \).

\[
E_{ZP} = \frac{\hbar}{2} \frac{9N_\chi}{2\omega_D} \int_0^{\omega_D} \omega^3 d\omega = \frac{9}{8} N_\chi k_B T_D.
\]

\( N_\chi \) is the number of atoms per formula unit, \( \hbar \) the reduced Planck constant, and \( k_B \) Boltzmann’s constant. The respective enthalpy changes \( \int_0^{298 \) K} \( c_p(T) \, dT \) and the relatively small \( E_{ZP} \) are subtracted from the formation
enthalpies for each of the solid product and reactant phases (for the metallic reactant phases considered here, the enthalpy of formation at 298 K is zero by definition).

2.3.2. $O_2$ reference

From the zero enthalpy reference of $O_2$ at standard conditions the integral $\int_{0\,K}^{298\,K} c_p(T) \,dT$ is subtracted (heat capacity data valid in this temperature range taken from [62]). For the remaining integral from 0 K to 90 K to be subtracted, a constant $c_p = 5R/2$ ($R$: molar gas constant) is assumed. Furthermore, the zero-point energy of $O_2$ as computed with the W4 quantum chemistry protocol is subtracted [46].

Assuming some degree of cancellation of pressure times volume ($pV$) terms between solid reactants and products [60], $pV$ terms are neglected for the solid phases. The formation energy training targets are thus given as the experimental enthalpy differences between the TM oxide and metallic phases and $O_2$ with heat capacity integrals and zero-point energies subtracted for all phases, as described above.

2.3.3. Ru oxide benchmarks

As test cases outside the training and validation data, two correlated Ru oxides with available experimental thermodynamic data are considered: $Ca_2RuO_4$ and $Y_2Ru_2O_7$. The experimental enthalpy of $Ca_2RuO_4$ at 1140 K is determined from the curvature in low-temperature measurements in [64] to be $T_D \approx 418$ K. The entropy of $Ca_2RuO_4$ is computed as $S(298\,K) = \int_{0\,K}^{298\,K} c_p(T) \,dT$, where $c_p(T)$ is taken from [65] for $Ca_2Ru_{1-x}M_2O_4$ with $M \in \{Mn, Fe\}$. The estimated entropy at 298 K is about 147 JK$^{-1}$mol$^{-1}$, irrespective of different small values of $x \lesssim 1\%$. Given the estimated $T_D$, the parameter $y$ is fitted to yield the entropy at 298 K. In analogy to the other solid phases, the integrated heat capacity and the Debye-model zero-point energy are subtracted from the enthalpy of $Ca_2RuO_4$.

The experimental enthalpy of formation of $Y_2Ru_2O_7$ at 298 K is taken from [66]. Debye model fits for the corrections of $Y$ and $Ru$ are based on data in [59], and also $O_2$ is corrected analogously to the other oxide heats of formation. A low-temperature heat capacity fit yielding a Debye temperature of 449 K for $Ca_2RuO_4$ is taken from [67]. Low and moderate-temperature heat capacity data from [68] and [66], respectively, is used in the integral (11) to determine the entropy at 298 K of $\sim 216$ JK$^{-1}$mol$^{-1}$ and fit the parameter $y$ of the Debye model for $T = 0$ extrapolation and subtraction of $E_{GP}$. The resulting benchmark Ru oxide formation energies and the 3d-TM and Mo oxide formation energies used for the model search are listed in Table S2 in the SI.

2.4. Model search

The ansatz for oxide formation energies in the model search used here is based on an energy offset $E_{\text{off}}(U_i, \rho_i)$ for every correlated site $I$:

$$E^{\text{form}} = E_{\text{GGA}} + U(TM \text{ oxide}) - \sum_I E_{\text{off}}(U_i, \rho_i)$$

$$= \sum_{\nu=1}^{N_{\text{metals}}} \beta_\nu E_{\text{GGA}}(\text{metal number } \nu) - \frac{\gamma}{2} E(O_2),$$

(13)

where the sum over $\nu$ accounts for the $N_{\text{metals}}$ metallic references calculated at the GGA level of theory and the factor $\beta_\nu$ for the stoichiometric occurrence of these metals in the oxide. $\gamma$ accounts for the stoichiometry of O in the oxide, and $E(O_2)$ is the corrected value from equation (2). The offset at site $I$ is chosen to be of the form

$$E_{\text{off}}(U_i, \rho_i) = \frac{U_i}{2} f_{\text{GP}} \left( \sum_\sigma \text{Tr} \rho_\sigma^\nu, \sum_\sigma \text{Tr} (\rho_\sigma^\nu)^2 \right).$$

(14)

$f_{\text{GP}}$ is a function of the sum of the eigenvalues of the projected density matrix for site $I$ and of the sum of the squares of these eigenvalues, ensuring rotational invariance of the model. Sums of higher eigenvalue powers as additional arguments for $f_{\text{GP}}$ did not lead to improved models, and models depending on higher eigenvalue powers hence are not presented in the following.

As $f_{\text{GP}}$ has no explicit site-dependence, site indices $I$ are dropped in the following from the density matrices and Hubbard parameters for simplicity, noting that evaluating a model always corresponds to a sum over all correlated sites $I$ in the oxide. Starting with the initial guess
corresponding to $E_{\text{tot}}(U, \rho) = \frac{1}{2} \log(2\pi\sigma^2) + 1$, the Akaike information criterion (AIC) [72] is implemented:

$$\text{AIC} = -2\Sigma + 2\Phi,$$

with

$$\Sigma = \frac{n}{2} \log(2\pi\sigma^2) + 1,$$

$$\Phi = \sum_{i=1}^{n} \left( \frac{a_i - b_i}{\sigma} \right)^2.$$

$b_i$ are the model predictions. Choosing a different energy unit than eV for normalization in equation (18) merely leads to a constant offset of the AIC, not affecting the relative ranking of the models. With decreasing $\sigma^2$, i.e., with increasing goodness of fit, the term $-2\Sigma$ in the AIC(16) decreases. The term $2\Phi$, on the other hand, is an overfit penalty for the number of parameters of the model. The selection strategy is thus to prefer models with small AIC.

The observations are experimental oxide formation energies and the computed values from equation (13) the model predictions, both normalized for each oxide as described in the following. The more atoms there are per formula unit, the larger typically the error in the predicted formation energy of the compound. Cohesive energies of solids that are well described within the GGA, e.g., have typical errors of 0.1 eV–0.2 eV per atom [12, 13]. The considered oxides have a relatively wide range of ratios $\kappa$ of number of atoms $N_A$ per formula unit to Hubbard–corrected sites $N_H$ per formula unit ranging from 2 to 8, and oxides with larger $\kappa$ typically will have a larger GGA contribution to the error in the predicted formation energy per formula unit. If $a_i$ and $b_i$ were normalized by dividing by $N_H$ or by number of formula units, residual errors for oxides with larger ratio $\kappa$ or larger formula units would be weighted more strongly in equation (18), respectively. The $f_{\text{GP}}^z$ could potentially be trained to compensate for relatively large GGA contributions to the errors, which is undesired. If $a_i$ and $b_i$ instead included a factor $1/N_A$, the $E_{\text{tot}}(U, \rho)$ would have a small relative weight for oxides with large ratio $\kappa$. As a compromise, the experimental and computed oxide formation energies per formula unit are divided by $\sqrt{N_A}$, ensuring that all $E_{\text{tot}}(U, \rho)$ have comparable, balanced relative weight in the residuals $a_i - b_i$. With these normalized residuals $a_i - b_i$, the corresponding normalized mean-absolute error (MAE) is defined:

$$\Delta E_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^{n} |a_i - b_i|,$$

which will be presented in addition to the model AIC in section 3.

The complexity term $2\Phi$ in equation (16) is generalized to not only account for and penalize the number of numerical parameters, but also mathematical operations:

$$\Phi = \sum_{j} w_j \varphi_j.$$

$w_j$ is a weight for a mathematical operation, variable, or numerical coefficient (see table 1) and $\varphi_j$ the corresponding occurrence in the considered model (or program) $f_{\text{GP}}^z$.

Using the resulting AIC as a fitness function evaluated on a random selection for each of the programs of half of the 66 TM oxide reference data, $k$-tournament selection [74] is employed as implemented in the gplearn code [71]. In this procedure, $k$ programs out of the current generation (here $k=25$ out of a total population size of 100 per generation) are randomly selected, from which the program with lowest AIC is selected for crossover and mutation operations (operations named in analogy to biological genetic evolution). With 70% probability, the winner from a second $k$-tournament is used for crossover of terms between the two programs, or with 25% probability, random mutations are performed as described in [71], and the resulting new program becomes part of the next generation. With 5% probability, the tournament winner is promoted to the next generation unchanged. The initial generation consists of 30 copies of program (15) and 70 randomly generated $f_{\text{GP}}^z$, and the programs are evolved for $10^5$ generations.
Beginning from the last generation, $10^5$ further generations are evolved, where now all programs in each generation are optimized numerically (using the Nelder-Mead simplex algorithm \[^{75, 76}\] as implemented in \textit{scipy} \[^{77}\]) and simplified symbolically (using the \textit{sympy} code \[^{78}\]) in addition to the genetic operations.

Fractions in programs are normalized to equivalent fractions so that the coefficient of the lowest power in the variables in the denominator is one. To constrain the search space for the final $10^5$ generations, a numerically large penalty is added to the AIC for models that do not vanish for integer density matrix eigenvalues, i.e., where $E_U$ vanishes, the DFT and DFT + $U$ energy functionals coincide, and thus no correction term for total energy comparison is required. A modified version of the \textit{gplearn} code with unique weights for mathematical operations, coefficients, and variables and integrated symbolic and numerical optimization can be found online \[^{79}\].

### 3. Results and discussion

From the constrained genetic programming search for models vanishing for integer density matrix eigenvalues, the programs with lowest AIC at a given complexity from the $10^5$ generations are shown in figure 3. All models found satisfying this constraint contain the factor $\text{Tr} \text{Tr} \rho_2$, which vanishes for integer eigenvalues of $\rho$.

The simplest program is to subtract $E_U$ from the GGA + $U$ total energy, yielding a relatively high $\Delta E_{\text{MAE}}$ of about 0.5 eV. Multiplying this model with a pre-factor increases the complexity $\Phi$ to 6. The resulting model $1.17 \cdot (\sum_\sigma \text{Tr} \rho^\sigma - \text{Tr} (\rho^\sigma)^2)$ lowers $\Delta E_{\text{MAE}}$ by only 6 meV, while the AIC increases by almost 6, thus identifying this model as an overfit. Significant improvements of the AIC and $\Delta E_{\text{MAE}}$ are achieved by programs further multiplying $E_{\text{off}}$ by $\sqrt{\sum_\sigma \text{Tr} (\rho^\sigma)^2}$ or $\sum_\sigma \text{Tr} (\rho^\sigma)^2 / \sum_\sigma \text{Tr} \rho^\sigma$, respectively.

The program with lowest AIC and hence optimal goodness of fit \textit{versus} complexity trade-off yields the correction term

| Complexity | Model $E_{\text{off}}/(\frac{1}{2}U)$ |
|------------|----------------------------------|
| 3          | $\text{Tr} \rho - \text{Tr} \rho^2$ |
| 8          | $0.76(\text{Tr} \rho - \text{Tr} \rho^2) \cdot (\text{Tr} \rho^2)^{\frac{1}{2}}$ |
| 11         | $1.88(\text{Tr} \rho - \text{Tr} \rho^2) \cdot \frac{\text{Tr} \rho^2}{\text{Tr} \rho}$ |
| 13         | $3.72(\text{Tr} \rho - \text{Tr} \rho^2) \cdot \frac{1}{1 + 2(\text{Tr} \rho - \text{Tr} \rho^2)}$ |
| 15         | $3.70(\text{Tr} \rho - \text{Tr} \rho^2) \cdot \frac{1}{1 + 1.99(\text{Tr} \rho - \text{Tr} \rho^2)}$ |
| 18         | $4.22(\text{Tr} \rho - \text{Tr} \rho^2) \cdot \frac{1}{1 + 2.33 \text{Tr} \rho - 2.31 \text{Tr} \rho^2}$ |

Beginning from the last generation, $10^5$ further generations are evolved, where now all programs in each generation are optimized numerically (using the Nelder-Mead simplex algorithm \[^{75, 76}\] as implemented in \textit{scipy} \[^{77}\]) and simplified symbolically (using the \textit{sympy} code \[^{78}\]) in addition to the genetic operations.

Fractions in programs are normalized to equivalent fractions so that the coefficient of the lowest power in the variables in the denominator is one. To constrain the search space for the final $10^5$ generations, a numerically large penalty is added to the AIC for models that do not vanish for integer density matrix eigenvalues, i.e., where $E_U$ vanishes, the DFT and DFT + $U$ energy functionals coincide, and thus no correction term for total energy comparison is required. A modified version of the \textit{gplearn} code with unique weights for mathematical operations, coefficients, and variables and integrated symbolic and numerical optimization can be found online \[^{79}\].
\[ E_{\text{off}}^{\text{opt}}(U_i, \rho_j) = 1.86 U_i \frac{\sum_{i} \Delta \rho_i - \sum_{i} \rho_i^2}{1 + 2\sum_{i} \rho_i^2}. \]  

where the site index \( i \) is not suppressed to emphasize that this model with site-specific \( U_i \) Hubbard parameter strength is to be subtracted from GGA+\( U \) total energies for every Hubbard-corrected site \( i \) in the system.

The next model in figure 3 has a non-integer and numerically optimized coefficient in the denominator (leading to a complexity increase of 2), but the optimized coefficient is almost an integer, and the model is approximately equal to (21), and the AIC increases. The most complex model in figure 3 further has slightly different numerical coefficients for \( \sum \Delta \rho \) and \( \sum \rho^2 \) in the denominator, which increases the AIC, identifying a tendency towards overfitting. Equation (21) is thus the model identified as optimal for computing offsets to GGA+\( U \) total energies for comparison at different \( U \)-values, including \( U = 0 \) GGA total energies.

\( E_{\text{off}}^{\text{opt}} \) only depends on the density matrices through the difference \( \Delta = \sum \Delta \rho - \sum \rho^2 \), and this dependence is depicted in figure 4(a). The model can be considered a rescaling of \( E_{\text{opt}} \). For differences \( 0 < \Delta < 1.36 \), \( E_{\text{off}}^{\text{opt}} \) is increased over \( E_{\text{opt}} \); while for larger differences, it is decreased relative to \( E_{\text{opt}} \).

In figure 4(b), the performance of the model is compared to GGA+\( U \) calculations with average, fitted \( U \)-values. The average \( U \)-values are taken from the Materials Project [9] and are listed in Tab. S3. Constant energy offsets for each TM species are fitted to the binary 3d and Mo oxide heats of formation, following the procedure in [36]. With the seven different TM ions considered and with the empirical, fitted \( U \)-values and energy offsets, there are thus effectively 14 fitting parameters. \( E_{\text{off}}^{\text{opt}} \) with the first-principles, site-dependent \( U \)-values and thus significantly fewer degrees of freedom (through its functional form and numerical coefficient) leads to similar residual errors in comparison to the experimental benchmarks, showing generally good agreement (experimental and computed TM oxide heats of formation listed in Tab. S2). With the similar quality of prediction for these bulk oxide formation benchmark cases, the method presented here holds the promise of enabling predictions of TM oxide formation energies and other reactions energies between transition metals and their oxides where experimental benchmark data for fitting is lacking.

The heats of formation of the ruthenates \( \text{Ca}_3\text{Ru}_2\text{O}_7 \) and \( \text{Y}_2\text{Ru}_2\text{O}_7 \) serve here as well-defined benchmark cases; other Ru oxides, such as \( \text{RuO}_2 \), \( \text{CaRuO}_3 \), \( \text{SrRuO}_3 \), and \( \text{Sr}_2\text{RuO}_4 \) are not considered here, as these oxides are metallic [80] \( \text{Sr}_2\text{RuO}_4 \) is metallic at temperatures below \( \sim 130K \) and superconducting below \( \sim 1K \) [81] and should thus not be treated at the GGA+\( U \) level of theory. As there are only two Hubbard-corrected Ru oxides considered here, no energy offset nor average \( U \)-value is fitted for Ru, but only the first-principles \( U \)-value approach based on \( E_{\text{off}}^{\text{opt}} \) is applied to compute the Ru oxide heats of formation. Calculation of the experimental benchmark values is explained in section 2.3.3.
Ca$_2$RuO$_4$ is an insulator at ambient conditions [82]. With powder diffraction-based structural refinements of the Pbca-phase at 295 K from [83] as a starting guess, the structure of Ca$_2$RuO$_4$ is optimized iteratively at the GGA+$U$ level with self-consistent $U$-parameters as described in section 2.2.3. Antiferromagnetic order with a propagation vector of (1, 0, 0) (corresponding to the A-centered mode [83]) is considered (the energy of the B-centered mode with propagation vector (0, 1, 0) is found to be within $\sim 10^{-4}$ eV/atom and is thus energetically almost indistinguishable from the A-centered phase within the accuracy of the simulations, allowing the lattice to relax for both the A- and B-centered modes). A Hubbard parameter strength $U_{I\text{Ru}}^{K}\approx4.5$ eV is calculated with equation (9). With the corresponding GGA+$U$ total energy of Ca$_2$RuO$_4$ and GGA total energies for the reactants and applying corrections (2) and (21), a heat of formation of $-2.29$ eV/atom is predicted, which is in good agreement with the experimental result of $-2.35$ eV/atom.

Y$_2$Ru$_2$O$_7$ has pyrochlore structure (spacegroup Fd$3m$ [84]) and is an insulator with magnetic frustration and Curie-Weiss-like behavior above $\sim 100$ K [85]. Below $\sim 76$ K, Y$_2$Ru$_2$O$_7$ shows antiferromagnetic, non-collinear spin ordering [86]. Here, the magnetic structure of Y$_2$Ru$_2$O$_7$ is approximated with collinear spins, and different spin orderings in a supercell containing eight formula units as depicted in figure 5 were compared energetically at the GGA+$U$ level. The low-energy antiferromagnetic phase used in the following is depicted in figure 5 (dashed lines indicate the primitive magnetic unit cell). After iterative structural optimization and computation of the first-principles Hubbard parameter strength, a parameter $U_{I\text{Ru}}^{K}\approx4.3$ eV is found. The computed heat of formation is $-2.36$ eV/atom, comparing well to the experimental estimate of $-2.44$ eV/atom. This $\leq0.1$ eV/atom error is within the range of errors found in [14] for the heat of formation of oxides not containing transition metal ions (and thus not requiring Hubbard corrections) with a fitted O$_2$ gas reference, and within the range of GGA reaction energy errors involving metallic TM phases in general [12, 13, 22]. While oxidation state-specific TM references fitted to experiments can reduce per atom oxide reaction energy errors significantly [39], having established here an approach to correcting mixed GGA and GGA+$U$ total energy differences with site-specific, first-principles Hubbard parameters, which leads to errors typical for GGA-only estimates for reaction energies without localization errors, will enable oxide reaction energy predictions in the absence of experimental reference data.

4. Conclusion

The presented approach enables the prediction of transition metal oxide formation energies and other reaction energies involving oxide and metallic phases with Hubbard corrections computed from first principles. Comparison of total energies from Hamiltonians with different Hubbard parameter strengths including $U=0$
is facilitated with a site-projected density matrix-dependent model, that was found through genetic programming against experimental transition metal oxide formation energies. This mathematically simple model does not involve fitting parameters to be adjusted to transition metal ion types or oxidation states. Unlike previous Hubbard-corrected approaches to TM oxide heat of formation prediction requiring experimental benchmark data to fit energetic corrections and average Hubbard strengths, the method presented here can thus also be used to compute transition metal oxide reaction energies when such experimental fitting targets are lacking, which will be of particular importance for oxide interfaces and defect chemistry.

Acknowledgments

This research was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division, Catalysis Science Program to the SUNCAT Center for Interface Science and Catalysis.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

J Voss https://orcid.org/0000-0001-7740-8811

References

[1] Kohn W and Sham L J 1965 Phys. Rev. 140 A1133
[2] Hohenberg P and Kohn W 1964 Phys. Rev. 136 B864
[3] Burke K 2012 J. Chem. Phys. 136 150901
[4] Becke A D 2014 J. Chem. Phys. 140 18A301
[5] Jones R O 2015 Rev. Mod. Phys. 87 897
[6] Payne M C, Teter M P, Allan D C, Arias T A and Joannopoulos J D 1992 Rev. Mod. Phys. 64 1045
[7] Curtarolo S et al 2012 Comput. Mater. Sci. 58 218
[8] Saal J, Kirklin S, Aykol M, Meredig B and Wolverton C 2013 JOM 65 1501
[9] Jain A et al 2013 APL Mater. 1 011002
[10] Ghi Ringhelli L, Carbogno C, Levchenko S, Mohamed F, Huhs G, Lüders M, Oliveira M and Scheffler M 2017 Npj Comput. Mater 3 46
[11] Talirz L et al 2020 Sci. Data 7 299
[12] Zhang G X, Reilly A M, Tkatchenko A and Scheffler M 2018 New J. Phys. 20 065020
[13] Brown K, Maimaiti Y, Trepte K, Bligaard T and Voss J 2021 J. Phys. Chem. 124 2004
[14] Wang L, Maxisch T and Ceder G 2006 Phys. Rev. B 73 195107
[15] Becke A D 1993 J. Chem. Phys. 98 1372
[16] Fulde P 1995 Electron Correlations in Molecules and Solids III edn (Berlin: Springer)
[17] Heyd J, Scuseria G E and Ernzerhof M 2003 J. Chem. Phys. 118 8207
[18] Heyd J, Scuseria G E and Ernzerhof M 2006 J. Chem. Phys. 124 219906
[19] Janthon P, Luo S A, Kozlov S M, Vines F, Limtrakul J, Truhlar D G and Illas F 2014 J. Chem. Theory Comput. 10 3832
[20] Gao W, Abete A, Cai T, Sun Y Y, Zhang S and Zhang P 2016 Solid State Commun. 234-235 10
[21] Jang Y R and Deok Y B 2012 J. Phys. Soc. Japan 81 114715
[22] Mallikarjun Sharada S, Karlsson R K B, Maimaiti Y, Voss J and Bligaard T 2019 Phys. Rev. B 100 035439
[23] Chevrier V L, Ong S P, Armentino R, Chan M K Y and Ceder G 2010 Phys. Rev. B 82 075122
[24] Anisimov V I, Zaanen J and Andersen O K 1991 Phys. Rev. B 44 943
[25] Czyżyk M T and Sawatzky G A 1994 Phys. Rev. B 49 14211
[26] Liechtenstein A I, Anisimov V I and Zaanen J 1995 Phys. Rev. B 52 R5467
[27] Gunnarsson O, Andersen O K, Jepsen O and Zaanen J 1989 Phys. Rev. B 39 1708
[28] Cococcioni M and de Gironcoli S 2005 Phys. Rev. B 71 035105
[29] Miyake T and Aryasetiawan F 2008 Phys. Rev. B 77 085122
[30] Yu M, Yang S, Wu C and Marom N 2020 Npj Comput. Mater. 6 180
[31] Tavazde P, Boucher R, Avendaño-Franco G, Kocan K X, Singh S, Dovale-Farello V, Ibarra-Hernández W, Johnson M B, Mebane D S and Romero A H 2021 Npj Comput. Mater. 7 182
[32] Sai Gautam G and Carter E A 2018 Phys. Rev. Materials 2 095401
[33] Long O Y, Sai Gautam G and Carter E A 2020 Phys. Rev. Materials 4 045401
[34] Kulik H J and Marzari N 2011 J. Chem. Phys. 135 194105
[35] Xu Z, Joshi Y V, Raman S and Kitchin J R 2015 J. Chem. Phys. 142 144701
[36] Jain A, Hautier G, Ong S P, Moore C J, Fischer C C, Persson K A and Ceder G 2011 Phys. Rev. B 84 045115
[37] Stevanović V, Lany S, Zhang X and Zunger A 2012 Phys. Rev. B 85 115104
[38] Deml A M, O’Hayre R, Wolverton C and Stevanović V 2016 Phys. Rev. B 93 085142
[39] Aykol M and Wolverton C 2014 Phys. Rev. B 90 115105
[40] Kresse G and Furthmüller J 1996 Phys. Rev. B 54 11169
[41] Blöchl P E 1994 Phys. Rev. B 50 17953
[42] Kresse G and Joubert D 1999 Phys. Rev. B 59 1758
[43] Perdew J P, Burke K and Ernzerhof M 1996 Phys. Rev. Lett. 77 3865
[44] Ernzerhof M, Perdew J P and Burke K 1997 Int. J. Quant. Chem. 64 285
[45] Nerskov K J, Rossmeisl J, Logadottir A, Lindqvist L, Kitchin J R, Bilgaard T and Jónsson H 2004 J. Phys. Chem. B 108 17886
[46] Karton A, Sylvestsy N and Martin J M L 2017 J. Comput. Chem. 38 2063
[47] Dudarev S L, Botton G A, Savrasov S Y, Humphreys C J and Sutton A P 1998 Phys. Rev. B 57 120
[48] Togo A and Tanaka I 2018 SpqLib: a software library for crystal symmetry search (arXiv:1808.01590)
[49] Kulik H J, Cococcioni M, Scherlis D A and Marzari N 2006 Phys. Rev. Lett. 97 103001
[50] Rohrbach A, Hafner J and Kresse G 2003 J. Phys.: Condens. Matter 15 9797
[51] Bengone O, Alouani M, Bloch P and Hugel J 2000 Phys. Rev. B 62 16392
[52] Giannozzi P et al 2009 J. Phys.: Condens. Matter 21 395502
[53] Timrov I, Marzari N and Cococcioni M 2021 Phys. Rev. B 103 045114
[54] Vanderbilt D 1990 Phys. Rev. B 41 7892
[55] Garrity K F, Bennett J W, Rabe K M and Vanderbilt D 2014 Comput. Mater. Sci. 81 446
[56] Nawa K, Akiyama T, Ito T, Nakamura K, Oguchi T and Weinert M 2018 Phys. Rev. B 97 035117
[57] Anisimov V I and Gunnarsson O 1991 Phys. Rev. B 43 7570
[58] Leiria Campo V Jr and Cococcioni M 2010 J. Phys.: Condens. Matter 22 055602
[59] Kubaschewski O, Spencer P J and Alcock C B 1993 Materials Thermochemistry VI edn (Oxford: Pergamon)
[60] Hautier G, Ono S P, Jain A, Moore C J and Ceder G 2012 Phys. Rev. B 85 155208
[61] Westrum E F and Grimes D 1957 J. Phys. Chem. Solids 3 44
[62] Chase M W 1998 NIST-JANAF Thermochemical Tables IV edn (Woodbury, NY: American Institute of Physics)
[63] Jacob K T, Lwin K T and Waseda Y 2003 J. Electrochem. Soc. 150 E227
[64] Nakatsuji S 2000 Quasi-two-dimensional Mott transition system Ca$_2$-$_x$Sr$_x$RuO$_4$ PhD Thesis Kyoto University
[65] Qi T F, Korneta O B, Parkin S, Hu J and Cao G 2012 Phys. Rev. B 85 165143
[66] Banerjee A 2019 J. Solid State Electrochem. 23 1749
[67] Blacklock K, White H W and Gürmen E 1980 J. Chem. Phys. 73 1966
[68] Taira N, Wakeshima K and Inatsui 2000 J. Solid State Chem. 152 441
[69] Koza JR 1992 Genetic Programming: On The Programming of Computers by Means of Natural Selection (Cambridge, MA: MIT Press)
[70] Koza JR 1994 Stat. Comput. 4 87
[71] Stephens T 2019 Genetic programming in Python, with a scikit-learn inspired API (https://gplearn.readthedocs.io/) (Accessed on 12/6/2021)
[72] Akaikze H 1973 Information theory and an extension of the maximum likelihood principle II International Symposium on Information Theory ed B N Petrov and F Csaki (Budapest: Akadémiai Kiadó) pp 267–81
[73] Konishi S and Kitagawa G 2008 Information Criteria and Statistical Modeling (New York, NY: Springer)
[74] Affenzeller M, Winkler S, Wagner S and Beham A 2009 Genetic Algorithms and Genetic Programming: Modern Concepts and Practical Applications (Boca Raton, FL: CRC Press)
[75] Nelder J A and Mead R 1965 Comput. J. 7 308
[76] Gao F and Han L 2012 Comput. Optim. Appl. 51 259
[77] Virtanen P et al (SciPy 10 Contributors) 2020 Nat. Methods 17 261
[78] Meurer A et al 2017 PeerJ Comput. Sci. 3 e103
[79] GPeeLearn (Web page accessed on 12/6/2021) (https://github.com/vossgroup/gplearn/)
[80] Longo J M, Raccah P M and Goodenough J B 1968 J. Appl. Phys. 39 1327
[81] Maeno Y, Hashimoto H, Yoshida K, Nishizaki S, Fujita T, Bednorz J G and Lichtenberg F 1994 Nature 372 532
[82] Nakatsuji S, Ikeda S and Maeno Y 1997 J. Phys. Soc. Japan 66 1868
[83] Braden M, André G, Nakatsuji S and Maeno Y 1998 Phys. Rev. B 58 847
[84] Kanno R, Takeda Y, Yamamoto T, Kawamoto Y and Yamamoto O 1993 J. Solid State Chem. 102 106
[85] Yoshii S and Sato M 1999 J. Phys. Soc. Japan 68 3034
[86] Blundell S J, Lancaster T, Baker P J, Hayes W, Pratt F L, Atake T, Rana D S and Malik S K 2008 Phys. Rev. B 77 094424
[87] Momka K and Izumi F 2011 J. Appl. Crystallogr. 44 1272